
COMPUTER SCIENCE
DEPARTMENT
nigeria computer society
ncs
1st INTERNATIONAL CONFERENCE
ON
ADVANCED COMPUTERISED SYSTEMS
AND EMERGING TECHNOLOGIES
ICACSET 2014
CONFERENCE PROCEEDINGS
NOVEMBER 30 - DECEMBER 3, 2014
VENUE: BABCOCK UNIVERSITY GUEST HOUSE,
BABCOCK UNIVERSITY, ILISHAN-REMO, OGUN STATE NIGERIA
Edited by
Nicolae Goga Ph.D
Sunday A. Idowu Ph.D, MCPN
Oludele Awolele Ph.D,PNCS,NCPN,PCAI
Folasade Osisanwo MNCS,MCSTA,MCPN
New Horizons
FirstBank
Since 1894
BUCC
eTranzact

1st INTERNATIONAL CONFERENCE OF THE

COMPUTER SCIENCE DEPARTMENT
BABCOCK UNIVERSITY, ILISHAN REMO, OGUN STATE
WITH

THEME: ADVANCED COMPUTERISED SYSTEM AND EMERGING TECHNOLOGIES

CONFERENCE PROCEEDINGS

VOLUME 1
November 30th– December 3rd, 2014

@
BABCOCK UNIVERSITY GUEST HOUSE,
BABCOCK UNIVERSITY, ILISHAN REMO, OGUN STATE

ACKNOWLEDGEMENT

The entire members of staff of the Computer Science Department, Babcock University, Ilishan Remo, Ogun State, appreciates and acknowledges the immense and esteemed support received from our professional bodies: Nigeria Computer Society (NCS), Computer Professionals (Registration Council of Nigeria (CPN)) and other organisations such as: New Horizons, Nigeria, First Bank, Nigeria, eTranzact, Shepherd University, Babcock University Computer Club, Babcock University Guest House and the Babcock University Community at large.

We also recognise the admirable efforts of all our reviewers and others that contributed to the success of this maiden conference.

REVIEWERS:

INTERNAL

PROF. A.C. OGBONNA
PROF. OLUDELE AWODELE
PROF. O.J. OMOTOSHO
PROF. S.A. IDOWU
DR Y.A. ADEKUNLE
DR S.O OKOLIE
DR A. O. ADEBAYO
DR. J.V. JOSHUA
DR. O.D. ALAO
DR. S.O. KUYORO

EXTERNAL

PROF. NICHOLAE GOGA: *UNIVERSITY POLITEHNICA OF BUCHAREST, ROMANIA / UNIVERSITY OF GRONIGEN, NETHERLANDS*

DR. MARIA GOGA: *UNIVERSITY OF CONSTRUCTIONS, BUCHAREST, ROMANIA*

DR. F.A. OLAJIDE: *UNIVERSITY OF PORTSMOUTH, UNITED KINGDOM*

UNIVERSITY OF COMPUTER SCIENCES AND SKILLS, POLAND:
PROF.ANDRZEJ NOWAKOWSKI
PROF. DR HAB. LILIANA BYCZKOWSKA-LIPIŃSKA
DR HAB. JERZY CZERBNIAK
DR MARIUSZ FRYDRYCH
DR HAB. ZBIGNIEW GMYREK
DR WOJCIECH HORZELSKI
DR MARIUSZ JAROCKI
PROF. ZW. DR HAB. EDWARD KĄCKI
DR HAB. ADAM PELIKANT
PROF. DR HAB. PAWEŁ TYMOSZCZUK
DR GRZEGORZ ZWOLIŃSKI

FOREWORD

I am happy to welcome you all to Babcock University, where the maiden International Conference of the Department of Computer Science, Babcock University, Ilishan-Remo, organised in collaboration with Nigerian Computer Society (NCS) is proceeding. The conference is themed "Advanced Computerised Systems and Emerging Technologies" with the intention of looking at various researched subjects on emerging technologies and their positive impacts on our society, by bringing together various intellect from the industry and academia. The divers researched works submitted were peer-reviewed and categorized under different sub themes of the conference such as Artificial Intelligence, Database and Enterprise Systems, Networking and Information Security, Cloud Computing and Services, Mobile and Ubiquitous Applications, Medical Informatics , IT Project Management, ICT-based Education , Cyber-security and Terrorism , Cashless Economy and Sustainable Development, IT based Research Methodology, Software Engineering ,Human Computer Interaction.

Alongside these papers, there would be three different keynote paper presentations, three lead paper presentations by renowned researchers in the field of computing and chains of rich educating activities scheduled to come up during the conference.

I have no doubt that this collection of research material will meet various research needs on advanced computerised systems and emerging technologies. It will be a useful tool for researchers at both undergraduate and postgraduate level

I wish to thank all those who submitted paper. I will like to also use this opportunity to express my appreciation to all our sponsors and the entire members of staff of the Computer Science Department, the local organising committee especially those who worked assiduously to make this conference a success.

Prof. Oludele Awodele *(FNCS, MCPN, FCAI)*
Conference Programme Editor.

TABLE OF CONTENTS

Approach to IT Project Formulation and Design: The Case Study of aflonPLAZA.com

Akin FAPOHUNDA
Aflon Digital Institute
Kuje, Abuja.
080331221004, director@aflon.org

INTRODUCTION

Variously untapped opportunities exist in the electronic commerce segment of the Nigerian business sector. It is quite easy to see notably successful sites such as Jumia and Konga in the marketplace. But the challenge is the identification of specific areas of need and deploying applicable IT software development tools. It obviously takes far more than just a degree or being a Professor in any of Computer Sciences discipline to midwife and deliver IT projects at that level of magnitude. Rather, the variety of skills and skillsets required for project execution are more than mind boggling: **Entrepreneurship**, **Technical Writing**, **Expansive Knowledge Base**, **Computer Graphics**, **Content Management System** (CMS), **PHP/MySQL Programming**, **Shopping Cart Development and Implementation**, **Payment Gateway Installation**, **Customer Relations Management** (CRM) and **E-Mail Marketing**. Additional requirements are **spatial geometry**, **colour theory** and **human-machine interface**. The formulation and development of aflonPLAZA.com as described in this paper represents a model that should provide guidance on how to proceed with IT-based business ventures.

PROJECT DESCRIPTION

The first basic requirement to guide an orderly development is for project articulation and formulation of its specification of features. The intended project is the building of www.aflonPLAZA.com as an electronic **Business to Business** (B2B) website for the creation of an online platform for multiple vendors to sell their products to buyers. All patrons: sellers and buyers are to be registered as members of the site, prior to being able to transact businesses. The first phase of the project is to cover

- Books and Pharmaceuticals: a single market for the publishing sector to facilitate the order and delivery of book by publishers to all educational institutions.
- Pharmaceuticals: a single point of contact between all Manufacturers and Pharmacy Stores owners.

SCOPE OF DESIRED FEATURES

Software programming is better predicated on a statement of requirements and specification of desired features. It is here that the clarity of intensions and purposes are better demonstrated. More often than not this aspect of the project requires far more than core IT qualifications. It needs a deep insight into the workings of the targeted business sector.

After due consideration of required functionalities and available software tools, the following features were outlined for incorporation into the site

- Unlimited vendors to register and sell their products, daily deals and auctions.
- The first objective of the site is to register members as sellers and buyers. Everyone has to register. Non-registered visitors have to be encouraged to register while browsing the site. Visitors can only browse the site. They cannot make selections or make orders.
- Once registration application is received, it has to be processed by the root administrator. Approval will subsequently communicated to the applicant. Buyers can then make others instantaneously. Sellers will also now be able to list their products. Respective USERNAME and PASSWORD will now become active.
- Vendor to have their own login access to manage products, deals, auctions.
- Book Sellers and Book Buyers can only list and buy books respectively. Buyers cannot be sellers and vice versa. They should be unable to deal on pharmaceutical products
- Manufacturers of Pharmaceutical Products and Pharmacy Stores can only list and buy pharmaceuticals respectively. Buyers cannot be sellers and vice versa. They should be unable to deal on Books
- Multi-delivery locations by Multi-vendors
- Separate Catalogue and Branding with company logo for each Vendor (on their pages)
- Unlimited buyers also have their own login access to perform product checkouts from multiple vendors in a single purchase.
- Admin can automate vendor payment processing after deducting admin commission / profit ratio when the customer/user places an order.

- Integrated SEO features, Social Media settings, E-mails, Newsletters are in-built to boost your online store traffic.
- Content Management System (CMS) facility for creating new web pages
- Customer Relations Management (CRM) facility for information dissemination: It will be necessary to have features that will persuade virtually every visitor to register either as a buyer or seller. Once registered, we need to keep track of members though

    - Greetings messages on birthday and other social occasions
    - messages announcing new product additions,
    - Invitations to purchase at discounted prices,
    - Appreciation messages to high volume buyers
    - Forwarding of newsletters, etc

- Messaging, Chatting, E-Mail and Forum facilities: All communications mails and chats are to be made from within the website and logged at the backend. Members are merely to be notified vide their own email addresses that MESSAGES & CHAT requests are pending
- Engagement of customers/ users with real-time feedback, ratings, product comparison, add to wish-list, most viewed products, Ads & more.

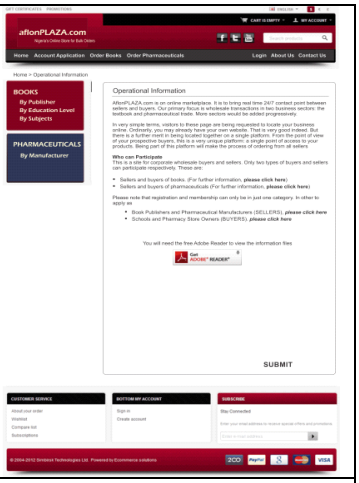
Above are outlines of features in plain English, rather than the mere naming of software development platforms such as JAVA, Django, C++. It is left to the discretion of the developer to work with a suite of applicable IT tools

TECHNICAL WRITING AND PRODUCTION OF GRAPHIC MOCK UPS

The transformation of thoughts and business ideas into implementable project document is another challenge. It is here that technical writing skills are required. Content writing is often quite problematic for IT professionals to handle. Very few web designers pay adequate attention to the need to communicate with web pages. Rather, many resort to animations that are quite problematic for visitors to the website to view or download.

There is an additional need for the production of mock-ups: exact details and specifications of all intended webpages. This requires expert vector and raster graphical handling skills. In the following illustrations are the mock-ups for the website.

1– Front End

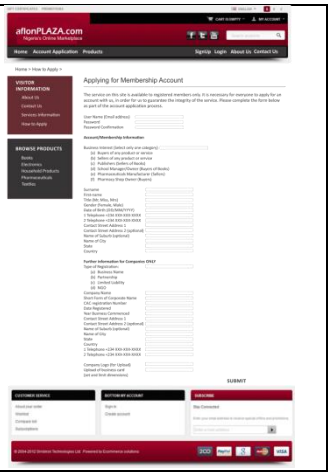


2–Operational Information

3–Account Application (Sellers)

4–Account Application (Buyers)

5-Terms and Conditions

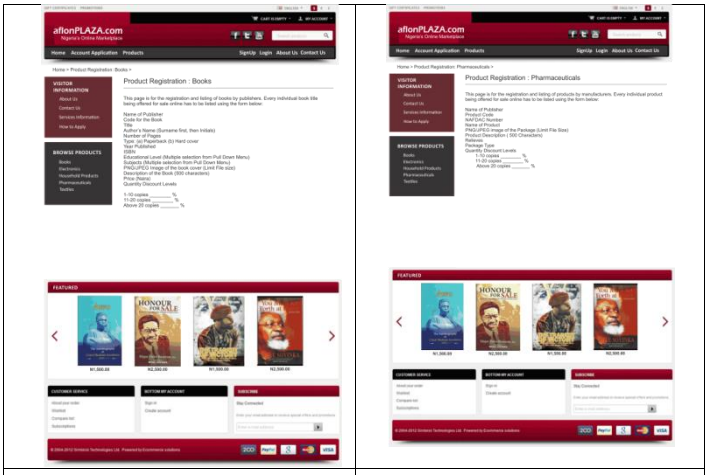


6- Membership Registration Page

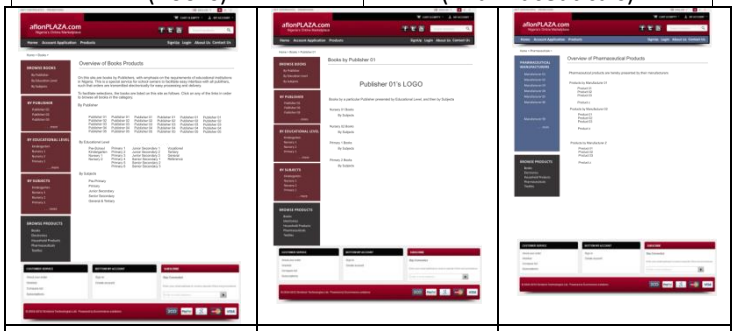
DATABASE MANAGEMENT SYSTEM

The underlying functionality that is basic to all the needs of IT applications by industry is record keeping: data on all daily transactions of clients, suppliers, customers and finances. Should it therefore not have been quite obvious that DBMS technology should be the foundation of all academic courses in the universities? Should the educational sector not have accorded the pride of place to say Visual Basic/MS-Access of PHP/MySQL? That is the right thing to do. Beyond mere data management systems, **data mining** and **data analytics** are emerging as the pathfinder to the future of IT. We need to pay due attention.

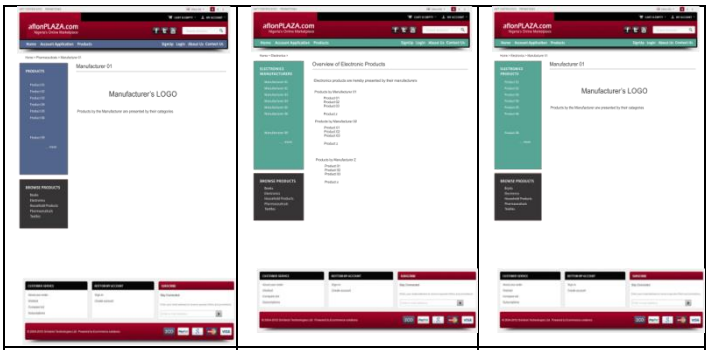
It is better imagined the level of programming required to successfully build the pages specified in the three mock ups in Fig 6 to 14. The first was to facilitate the registration of all potential customers of the proposed site. The next two illustrations are to help prospective sellers list their products on the website. The requirements for the rest illustrations are quite obvious.



7-Product Registration (Books) | 8-Product Registration (Pharmaceuticals)



9 Books Order Main Page | **10 Books Order Publishers** | **11** Pharmaceuticals Order Main



12 Pharmaceuticals Order - Manufacturers | **13** Electronics Order Main | **14** Electronics Manufacturers

APPLICABLE TOOLS

More than anything else, the features specified for the site necessitate several software development skills. Proficiency in a restricted area of specialisation readily makes any professional quite incapable of producing any viable outcome. It is a very big challenge as we are quite unwilling or even unable to work as a team in this country. Yet, all projects require multifaceted skills and abilities

Besides, learning high programming languages such as JAVA, C++, etc. are not directly helpful towards a successful IT project implementation. I venture to state that the practical requirement is for a very deep and practical knowledge of software packages, especially of the Open Source Software (OSS) class, as follows:

> **DRUPAL, PLONE, MAGENTO, OPENCART, SUGARCRM, MAIL-CHIMP, ALFRESCO. MOODLE, BLENDER, DOCEBO, GIMP, LAMP, XAMP, WHMCS, CPANEL, SAKAI,** etc.

It is my opinion that expert knowledge of multiples of the above listed packages will help the academic community break out into the sphere of industry in Nigeria. I would go to the extent of proposing that entire courses be based on these software suites

ENGAGEMENT OF CONSULTANTS

I may be faulted for not looking and search adequately enough around Nigeria, but I had no hesitation in looking beyond the shores of Nigeria for the requisite expertise for the above described project. First I retained a Russian firm to install and produce the theme for the site. The back-end programming was awarded to a freelance consultant that was sourced vide www.odesk.com. The project is completed and now up and running.

CONCLUSION

The website aflonPLAZA.com has been built to solve the problem of Business to Business (B2B) interactions in the selected industrial sectors of publishing and pharmaceuticals. Many more of such sites need to be built. The academic community ought to adopt an entrepreneurial mode of attitude. They need to identify opportunities for solving problems of the society. Quite universally, society rewards its segment that identifies and proffers solutions to her perceived problems. Wealth is created and accumulated in contemporary era by those that use IT knowledge base to clear bottlenecks or just simplify daily living.

But one tract mindedness and individualistic quests for quick wealth will get us nowhere. The academic campuses are best placed to function multidisciplinarily and foster group and team building. Integrated suites of software should be the focus of IT project development and implementation in the campuses.

Surveying Ethical Concerns in Virtualization for Business

Emmanuel C. Ogu[1]
ecoxd1@yahoo.com,
Department of Computer Science,
Babcock University, Ilishan Remo, Ogun State.

Ogbonna A.C.[3]
acogbonna06@yahoo.com,
Department of Computer Science
Babcock University, Ilishan Remo, Ogun State.

Omotunde A.A.[5]
ayo_omotunde@yahoo.com,
Department of Computer Science
Babcock University, Ilishan Remo, Ogun State.

Oyerinde O.D.[2]
oyerindeo@unijos.edu.ng,
Corporate Information Systems, ICT Directorate,
University of Jos
Jos, Plateau State, Nigeria

Michael I. Ogu[4]
ogumyk@gmail.com,
Department of Political Science and Public Administration,
Babcock University, Ilishan Remo, Ogun State.

***Abstract* – Virtualization, the technology that allows multiple guests (clients) to reside on a single host (provider) machine and share the resources of the host machine, is rapidly gaining prominence in the corporate business world of the 21st century. Virtualization is a powerful technology for increasing the efficiency of computing services. However, besides its advantages, its flexibility also raises a number of ethical concerns especially as it forms the bedrock of the revolutionary technology of cloud computing. Research has shown that Cloud Computing directly has the potential to tremendously impact positively the profit margins of leading 21st century businesses but the lack of satisfactory answers to some of these ethical concerns, especially by cloud service providers, has greatly limited the rate of adoption of full virtualization by many governments and organizations. This paper aims to elucidate and classify the questions and threats to virtualization, and suggest possible solutions to them.**

***Keywords* – Cloud computing, Computer Ethics, Virtualization.**

I. INTRODUCTION

It was not until the 1970s before mainframe users saw the first implementation of virtualization and symmetric multiprocessing. Prior to this time, the cost of computers and computing services were very high. Coupled with only sparsely intermittent computational needs, organizations, researchers and academicians, who were the first to welcome computing equipment and services. They found it difficult owning computers because they could not provide justification for investing at high costs in a computer that would sit idle for most of the time. To assuage this situation entrepreneurs came up with the idea of "renting" time; thus making it possible for organizations and users to either own or subscribe to computing resources at much lower costs [1]. It therefore became possible for users to find access to large scale mainframe computer systems using thin clients / terminal machines which were often referred to as "static terminals / machines" because they were used for input and output communications only and had no internal processing capabilities. This technology made the use of expensive mainframe systems more efficient because it allowed multiple users to share both the physical access to the computer from multiple terminals as well as to share the processing and computing resources (CPU time, disk time and space, etc.). This practice was able to curb periods of "no activity" on the mainframes and allowed for greater returns on investment for companies that practiced such around the 1950s [2]. It is this technology that grew through various transformations and related nomenclature over the following time periods: Remote Job Entry in the 1950s [3], Shared and Dedicated Web Hosting (which are forms of Virtual Web Hosting) around 1995 to 1997 [4] [5]; Virtual Private Server (VPS) Hosting around 1998 [6], Grid/Utility Computing [7] [1] to become Cloud Computing (CC) about three decades later.

According to [8], Cloud Computing is "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction; having characteristics of on-demand self-service, broad network access, resource pooling, rapid elasticity and payment per usage of various business models."

Virtualization is the technology that allows multiple virtual machines (also called guest machines) to run on a single physical machine (host machine) and share the resources of the physical machine [9]. This

therefore makes it possible for a single physical server to host many guest virtual machines (VMs), operating systems, and applications without the additional cost and complexity that result from running multiple physical machines [10]. This development became one of the forces that was to revolutionize information technology and cloud computing in the coming years [11].

Cloud computing services are delivered through three standardized service models: the Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and the Software as a Service (SaaS) Models.

The Software as a Service (SaaS) model is the first and topmost model in which clients/consumers are provided access to the applications of a provider which are deployed on a cloud infrastructure. These applications are made accessible to various client devices either through a thin client interface, such as a web browser (e.g., web-based email), or through a program interface. Consumers do not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

The second model is the Platform as a Service (PaaS) in which clients/consumers are allowed to deploy onto cloud infrastructure, consumer-created or acquired applications that are created using programming languages, libraries, services and tools supported by a provider. Consumers do not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but reserve control over deployed applications and possibly the configuration settings for the application-hosting environment.

Infrastructure as a Service (IaaS) is the third and most foundational model. In this model, consumers are provided with processing, storage, networks, and other fundamental computing resources, which enable the consumer is able to deploy and run arbitrary system and/or application software. Consumers do not control the underlying cloud infrastructure (basically hardware), but reserve control over operating systems, storage, and deployed application; possibly with limited control of some networking components (such as host firewalls).

The service models specified above directly define the three layers comprised in the core of most modern cloud computing infrastructure. Each of these layers offer the specified types of services to a particular segment of the consumer market while at the same time paying for the services provided by the preceding layer (except the IaaS layer) [12]. This is shown in figure 1:

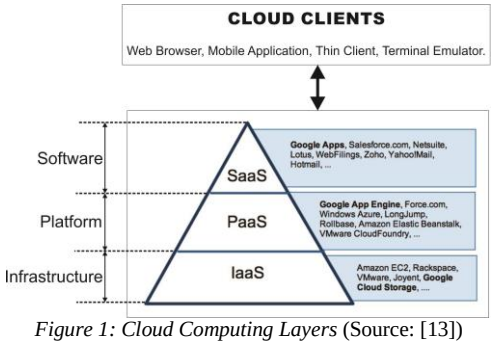


Figure 1: Cloud Computing Layers (Source: [13])

The first / foundational layer of Cloud Computing is the IaaS layer. The products here relate to hardware and associated services such as: general processing, servers, storage devices, database management, and all other hardware related services that are offered as a service to the end user. The next layer is the PaaS layer upon which developers can build and test applications that run on the IaaS, either for the IaaS layer itself or for the SaaS layer above it. The topmost layer is the SaaS, and this deals exclusively with applications for end users [12].

Virtualization is very important to cloud computing. As a matter of fact, it provides the abstraction from hardware state that cloud computing enjoys by taking a physical resource such as a server (computer) and dividing it into virtual resources called virtual machines (multiple computers that can be released to subscribers) to which users (businesses) could subscribe. "This abstraction from the hardware state allows not only multiple operating systems to coexist on the same hardware, but for one VMM to run on multiple different networked physical systems concurrently. By utilizing a VMM to mediate between the OS and the hardware, virtualization changes the one-to-one mapping of OSs to hardware to many-to-many" [14].

II. VIRTUALIZATION: THE TECHNOLOGY

Virtualization is basically a technology that makes it possible for an operating system to run within another operating system as an application [15]. It is a technology that provides for an environment that is equivalent to the operating environment of the host operating system, in which other operating systems can run as guests, but which is decoupled from the hardware state [14]. Although the technology of virtualization is not an entirely new paradigm, it has begun to gain rapid prominence in recent times because most modern system architectures now come with built-in computer functionalities that make it possible for other operating systems to run on top of one major (host) operating system.

Virtualization could be commonly server-based [11] or system-based [14] depending on the level of abstraction from which it is viewed. Usually, in server-based virtualization, a physical resource (such as a server) is used to host other operating systems that are provisioned and delivered over a network

environment, while in system-based virtualization, a host operating system within a particular computer is able to provision its resources to other operating systems that are able to run within it as guests that are decoupled from the hardware state. Other types of virtualization are network virtualization [16], storage virtualization [17], and desktop virtualization [18]. In this study, however, the term "virtualization" would be used to refer to the broad, generic, underlying concepts that power the technology of virtualization and not necessarily any of the specific types.

It should be noted that although various computer organizations and architectures implement virtualization using different implementation arrangements, certain generic components basically make virtualization possible. These include the:

Guest Operating System (OS),
Virtual Hardware,
Virtual Machine Monitor (VMM) or Hypervisor,
Host Hardware System,
Host Operating System,
VMM Drivers, and
Virtual Machine (VM) Applications [14].

These components are organized into a generic structure that is as illustrated figure 2:

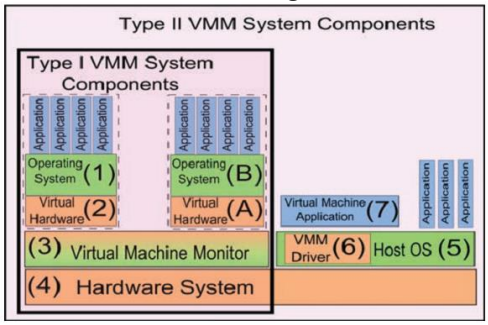


Figure 2: Core Components in Virtualized Architectures (Source: [14])

In virtualized environments, operating systems operate on the hardware as privileged software, and are generally able to perform any operation the hardware supports, whereas programs running inside an operating system are less privileged, and generally cannot perform operations except those that the operating system permits. These privilege levels are often called *rings*, with the lower numbered rings (i.e., *ring 0* or *dom 0*) having higher privileges than those with higher designations [19]. Operating system kernels generally run in the lowest ring, and thus have control over everything running in the lower privilege (higher numbered) rings [14].

Current virtualization solutions can be classified into three main categories. These three categories fall within two main classes: the *hypervisor-based virtualization* and the *non-hypervisor-based virtualization* (also known as operating system or container-based virtualization).

In the hypervisor-based virtualization, virtualization is made possible through a **virtual machine monitor** (VMM), otherwise known as a h**ypervisor,** which would be discussed shortly, while in the non-hypervisor-based categories, virtualization is made possible via operating system containers. The three main categories of virtualization solutions are:

Paravirtualization (PVM). This was one of the first adopted models of virtualization and is still widely used today. This virtualization solution needs no special hardware to achieve virtualization but instead relies on special kernels and drivers. This kernel(s) sends privileged system calls and hardware access requests directly to a hypervisor or virtual machine monitor [VMM]. The VMM then decides how to handle the request. The use of special kernels and drivers limits the window of flexibility in terms of choosing what operating systems to run. PVM must hence use an operating system (OS) that can be modified to work with the hypervisor. This virtualization solution reduces the overhead needed to virtualize privileged operating system calls since no special hardware is needed to intercept them. Examples of paravirtualized solutions include Xen and User Mode Linux. [9]

Hardware Virtual Machine (HVM). This solution stands at the lowest level of virtualization. HVM requires special hardware capabilities to be able to interpret privileged system calls from guest machines. It makes the full virtualization of a machine possible without the need for any special operating systems or drivers on the guest system. Guest machines actually interact through an emulated interface that appears as though they are communicating directly with the host hardware. Most modern processors have HVM capabilities which are often called *virtualization extensions*. These extensions detect when a guest machine tries to make a privileged system call like sending data on a Network Interface Card (NIC). This call is intercepted by the hardware and passed on to the hypervisor to decide how it should be handled. Great flexibility in terms of what OS to run is present. However, HVMs are known to have the highest overheads of all virtualization solutions [20], and usually do not come as the first choice in most situations. Examples of HVM virtualization solutions include the VMware Server, KVM, and Virtual-Box. [9]

Container Virtualization (CV). This non-hypervisor-based solution, also known as OS-level virtualization, creates multiple secure containers on a single operating system which runs different applications. It is based on the fact that a server administrator may want to isolate different applications for security or performance reasons, while maintaining the same OS across each container. It allows a user to share a single OS kernel between multiple containers and have them securely use computer resources with very minimal interference between containers. It is shown to have the lowest overhead of all virtualization solutions [20], but completely lacks flexibility in terms of which OS to run. Examples of CV solutions

include OpenVZ, Linux-VServer and Solaris Zones. [9]

Table 1 gives a brief comparative summary of the categories of virtualization solutions.

Table 1: Comparative Summary of the categories of Virtualization Solutions [Source: authors]

S/N	Virtualization Solution	Performance Overhead	OS Flexibility	Performance Isolation	Hypervisor-Based
1	Paravirtualization (PVM)	Moderate	Minimal	High	Yes
2	Hardware Virtual Machine (HVM)	High	High	Least	
3	Container Virtualization (CV)	Low	None	Highest	No

The **hypervisor** or **virtual machine monitor** (VMM) is a highly privileged piece of software that runs either alongside or underneath the operating system that must be designed in such a manner that reflects an "*efficient, isolated duplicate of the real (or physical) machine*". It is a control program that functions as a *dispatcher* (for assigning tasks to other resident control modules), an *allocator* (for taking final decisions on the allocation of system resources) and an *interpreter* (for interpreting instructions and system calls for all resident guests) [21] [14]. In essence, the VMM is responsible for all tasks relating to the interfacing of guests with host hardware. The exact responsibilities (resource allocation and management, virtual disk space management, processing of request and system calls, etc.) that are borne by operating systems in standalone systems are the same responsibilities borne by the virtual machine monitor / hypervisor in multi-tenant architecture (or systems), where a single server is able to host several virtual machines (subscribers) [21].

There are basically two types of hypervisors. These types are distinguished relative to their position with respect to the host operating system and the host hardware. These types are as follows:

The Classical Type-I Hypervisor. This is also known as the bare-metal hypervisor. This type of hypervisor is installed as the primary boot system on the hardware and as such executes at the highest level of privilege (*ring 0* or *dom 0*). They have full control over all virtual machines that run on the hardware [22] [14]. The position of a Type-I hypervisor in a typical virtualized environment is illustrated in figure 3.

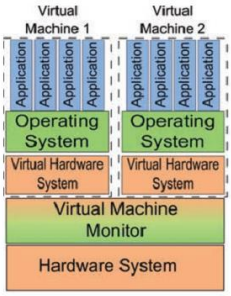


Figure 3: Generalized Architecture for Type-I Virtual Machine Monitors (Source: [22])

The Type-II Hypervisor. This is also known as the hosted hypervisor. This type of hypervisor runs alongside or above a host operating system which sits on an underlying hardware, and may share drivers from the host operating system to handle input / output (I/O) calls. This cooperation results in a VMM system that does not require hardware-specific drivers for VMM I/O operations, and allows virtual machines to coexist within an existing operating system environment. Due to the above conditions, restrictions on the types of VM to use are light because the existing OS does not need to be overwritten or migrated to a multiple boot arrangement before virtualization can occur [14] [22]. The position of a Type-II hypervisor in a typical virtualized environment is illustrated in figure 4.

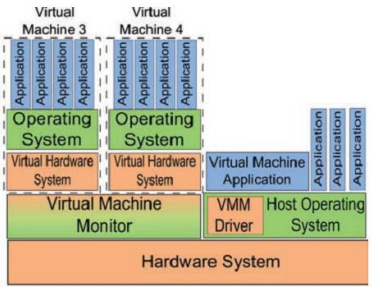


Figure 4: Generalized Architecture for Type-II Virtual Machine Monitors (Source: [22])

Table 2 shows a comparison of some other features of the Types I and II Hypervisors.

Table 2: Comparison of the Type 1 and Type 2 Hypervisors
(Source: [23])

Sl.No	Feature	Type 1	Type 2
1.	**Definition**	Hypervisors run directly on the system hardware.	Hypervisors run on a host operating system.
2.	**Support**	Hardware virtualization.	Operating system virtualization.
3.	**Examples**	VMware ESXi and Citrix XEN Server.	KVM, Virtual Box, VMware Server and Microsoft Virtual PC.
4.	**Efficiency, Availability and Security**	Comparatively better than Type 2.	Though inferior, it is used mainly on systems where support for a broad range of I/O devices is important.
5.	**Performance**	Very high. Resources are not being consumed by a bloated parent operating system.	Steep resource-overhead penalties reduce performance.
6.	**Ease of use**	Fairly easy to install but complicated to configure.	Easy to install, use and maintain.
7.	**High availability**	Yes.	No.
8.	**Reliability**	Yes.	Moderate.
9.	**Virtualization hypervisor management**	More options for management and automation. Centralized consoles to manage large number of hosts and VMs.	Fewer options for management and automation as well as limited VMs can be managed.
10.	**Cost**	Very costly.	Moderate.
11.	**Scalability**	Very high (easily run hundreds of VMs on a single host).	Very limited scalability (in the size of the VMs and the number of VMs that can run on a single host).
12.	**Resource control**	It offers the least amount of resource overhead and advanced resource controls that allow you to guarantee, prioritize and limit VM resource usage.	It has no or limited resource controls, so VMs have to fight each other for resources.
13.	**Size OR Complexity**	Smaller.	Bigger and more complex.

[21], proposed formal requirements for full virtualization of Virtualizable third generation architectures which have formed the foundation and yardsticks against which all hypervisors and hypervisor-based solutions have been measured overtime. Many other succeeding researches have either agreed with or slightly improved on the propositions of these requirements: [24], [25], [26], [23], [14] and [27]. Three properties for virtualized architectures include sameness, equality, and control. [21]

In sameness, "the VMM provides an environment for programs which is essentially identical with the original machine" (i.e. innocuous instructions must be run directly on the processor with interventions only where necessary).

In the case of equality, "programs run in this environment show at worst only minor decreases in speed"

Control denotes the situation where "the VMM is in complete control of system resources"

Two types (or properties) of virtualized instructions are also usually specified namely privilege level and sensitivity. [21]

Privilege level addresses the question of "does this instruction require a process to be highly privileged to call it directly? If the CPU traps and switches control to supervisory software (running in low rings) when the instruction is called from a process running in user mode (high rings) then the instruction is *privileged*, as it requires privilege to be executed [else, it is *non-privileged*]." [14]

Sensitivity on its part tries to answer the question of whether an instruction has the capacity to interfere with something the VMM should have complete control of. A *sensitive* instruction has the capacity to interfere with VMM operation, whereas an *innocuous* one does not possess such capacity. A simple example is that reading VMM program memory will not interfere with VMM behaviour, but writing to it could."

Rationale for Virtualization

Despite the relatively cheap cost of commodity hardware and availability of networks, virtualization still has benefits that put it in the fore for most organizations. Physical systems are associated with other costs which could be operational, physical and/or technical, all in addition to the initial purchasing cost. In addition, every physical machine must contend with requirements of physical space, cabling, energy, cooling, and software administration [14]. These requirements inadvertently leave virtualization as a preferred alternative.

Furthermore, modern commodity operating systems such Windows and Linux are very complex, usually comprising of tens of millions of lines of code (LOC) in the latest desktop versions. This situation results in a much larger vulnerability surface than can be easily or provably secured according to [25] and [24]. The fewer lines of codes in virtual machine monitors thus provide a preferred alternative in the above regard.

Similarly, it remains a fact that operating systems add a single point of failure for everything (data, processes and information) that runs on them. The difficulty in securing this single, complicated point of failure poses a security risk for the system's data and processes, hence accounting for the choice of virtualization [14].

Virtualization makes it possible to test new computer and IT aided business solutions in an environment that is less vulnerable to external vices. Using methods of sandboxing and container operating systems, new, untrusted and potentially vulnerable IT business solutions can actually be tested in a more secure environment where failures and breaches of the new systems or solutions would not hamper the current smooth running of existing deployed solutions.

Legacy business solutions that are already in use and properly functioning can easily be retained and run alongside new solutions deployed on virtualized platforms until these new solutions have been confirmed and smooth migration / transition has been achieved.

An added advantage is that the simulation of networks and independent business units in order to study and observe the cooperation and interactions between them is made easier by virtualization. In the same vein, independent business units can be simulated to run various independent processes that may either be interrelated or not, and the interactions between these can be observed more critically within a closed environment.

Virtualization further makes the previously dreaded tasks of migration / transition, backup and recovery of business systems a "walk-in-the-park", and the administration and management of these tasks can be more centrally done [28].

The implication of removing the dependency of operating systems on a system's physical state through system virtualization, according to [14], is that it allows multiple operating systems to be installed on a VMM, and thus multiple operating system VMs (called guest operating systems) can be installed on each physical system thereby allowing multiple VMs on the same hardware with its many advantages. Near-complete isolation between guest operating systems on the same hardware protects against OSs being a single point of failure; further allowing OS consolidation from different machines as is necessary to reduce system underutilization and maintain efficiency of operation" [14]. The abstraction from the hardware state consequently allows not only multiple operating systems to coexist on the same hardware, but for one VMM to run on multiple different networked physical systems concurrently. By utilizing a VMM to mediate between the OS and the hardware, virtualization changes the one-to-one mapping of OSs to hardware to many-to-many as reported by [14].

However, it must also be noted that the extent of virtualization could also impose extra performance overheads to computing infrastructure as already pointed out in tables 1 and 2.

III. VIRTUALIZATION: BUSINESS AND INDUSTRY PERSPECTIVE

Although virtualization and cloud computing provides ICT executives with high hopes of lasting solutions to business problems and needs, thereby enhancing business growth and performance indices through advancements in virtual collaboration, green IT, enterprise mobility, business intelligence, amongst others [29], it has become necessary to attempt to deal with the many ethical apprehensions slowing down full virtualization of IT infrastructure. These apprehensions, classified as straight-forward questions, would ensure that the journey to finding solutions may involve a more direct, precise and decisive approach to dealing with these issues.

The essential concept behind virtualization is the concept of multi-tenancy. As is the case in other areas of human endeavour, getting two different sets or groups (businesses or enterprises in this case) of people to co-habit the same habitat (be it economic, environmental or technological) generates grave ethical concerns which are aggravated in the case of virtualization by advancements in information and communication technology.

It needs to be noted at this juncture businesses, organizations and governments have various reasons hindering full virtualization of their IT infrastructure, as reported in [11]. A careful examination, however, reveals that the reasons are rooted in deep, pertinent ethical concerns to which cloud service providers usually do not provide satisfactory answers when they are posed as the questions which will be dealt with subsequently.

IV. DRIVERS OF VIRTUALIZATION

The influence of technology on government processes and procedures cannot be overemphasized. In turn, the procedures and processes of government – policy (formation and implementation) and infrastructure – can also influence the efficiency of several technological innovations within the state. Virtualization has been identified as one of the most promising technologies for the enhancement of business performances, leading to sustainability of the business, as well as productivity of both the individual employee and the organization as a whole. It is also important to discover if the environment within which virtualization occurs, the action and inactions – policies – of either the business organization or government and other efforts at training and development, have any direct or indirect implications for the efficacy of virtualization as a technological procedure.

Policy Implications: Policies are, more or less, the major instruments for running businesses or governments. The actions and inactions of government impact greatly on the level of progress experienced by the state, likewise those of business executives influence business outcomes to a large extent. [30], [31], among other authors, has asserted the link between public policy and technology. Policy formulation and implementation must be supportive of technological advancement and use, if such technologies will be effectual within the territories they are intended. Policies formulation without properly implementation, or vice versa, the merits of technology may not be fully accomplished.

Infrastructural Implications: Also as important as policy formulation and implementation is the provision of necessary and adequate infrastructure to ensure the deployment and use of such technologies for either business or government purposes. Several states in Africa face some very peculiar problems of infrastructure, which has made it difficult for the continent to compete favourably, globally. In addition to electricity, educational institutions are also very vital infrastructures that are necessary for ensuring that maximum potentials of technology are exploited to the benefit of businesses and governments alike.

Training and Development: Training and development is another major challenge of technology in many African states. In most cases, technologies exist, even at the very minimum levels, but they are not maximized owing to an inadequate knowledge of operating such

technologies. Virtualization in businesses may not provide optimum benefits if employees are not trained on the use of such technologies and the benefits that maximum utility could provide for the business, as well as the ethical considerations that go into such use.

V. ETHICAL CONCERNS IN VIRTUALIZATION

To the layman, ethics merely refers to moral standards governing behaviour, actions, and choices; or a discipline that studies the rightness or wrongness of behaviours, actions and choices.
Scholars of ethics study ethical principles from the viewpoints of normative and prescriptive ethics. From the viewpoint of normative ethics, ethical principles are derived from consistent and well-based standards of right and wrong that stipulate and inform human choices, decisions and actions in terms of rights, obligations, benefits of the common good, fairness, as well as other specific virtues. In light of this, ethics would include such standards as relate to reasonable obligations to refrain from rape, stealing, murder, assault, slander, and fraud; standards that also inspire virtues of honesty, compassion, and loyalty; as well as those relating to fundamental human rights.
From the viewpoint of prescriptive ethics, ethical principles relate to personal and communal ethical standards, associated with behaviour, feelings, laws, social habits norms and mores that could deviate from some more universal ethical standards, thus necessitating the need for constant review and overview of such standards to ensure that they remain congruent and in line with the progress and continuity of humanity. [32]
[33], defined computer ethics as "a field of study that examines ethical problems aggravated, transformed or created by computer technology". This definition is relevant to the course of this research because it captures the fact that some known ethical problems were, in effect, made worse by computers and computer technology, while some newer ones emerged with the advent of computing and information technology.
In line with this definition, [34] corroborated that the flexibility of virtual machines which enables users to create, copy, save / checkpoint, read and modify, share, migrate and roll back the execution state of entire machines by simply manipulating a file has generated new ethical concerns that have assumed a very recurring nature when modern businesses are faced with the prospect of cloud computing and virtualization.
The ethical concerns that lie very foundational to why most businesses, organizations and governments remain apprehensive of full virtualization of their IT infrastructure can be categorised into six parameters namely: co-tenancy, sharing, trust and privacy, hacks and breaches, boundary enforcement and restrictions, and integration and compatibility.

CO-TENANCY is the concern that is related to the type and calibre of individuals who own and operate the other virtual machines that share the same platform as well as the applications that run on them – who the co-tenants are. Usually, competitors within the same business endeavour or industry are not keen on getting on the same virtual platforms with each other for fears of breaches in their trade secrets by the competitors who are on the same platform. In reality, virtualization involves a lot of sharing. There are shared folders, hardware and resources, and these are the primary target during introspective attacks. Also, successful subversion and escalation of privileges could give competitors snapshots or entire clones of virtual machines with the content in them [34].

SHARING expresses the concern related to the amount / extent (type and depth) of information that virtual machines share with co-tenants and with the host. Considering that virtualization involves a lot of sharing of folders, hardware and resources, clients would want to know what kind of information from their virtual machines can be seen in the shared folder as well as the extent / depth of information. This concern is pertinent because spoofing the data contained in the shared folders or the virtual drive tables could offer very useful information about virtual machines and their content [14].

TRUST & PRIVACY (provider's level of interference with private virtual machines).
Trust issues express a concern over the extent of manipulation or interference that virtual service or cloud providers could have with virtual machines running on their infrastructure. Clients want to be convinced that service providers cannot directly access and manipulate their virtual machines as well as their content. This is primarily due to the fear that service providers may be able to sell unencrypted clones / snapshots of the content of their virtual machines to competitors.
Privacy concerns are most often used synonymously and interchangeably with trust. Privacy concerns however are more on the part of the clients, denoting the extent to which clients are able to keep the activities, data and operations of their virtual machines away from the prying eyes of

cloud service providers / administrators. [35], pointed out the fact that some activities of Cloud Service Providers (CSPs) may, in effect, go a long way to further complicate data privacy issues due to the extent of virtualization and cloud storage that are used to implement cloud services.

HACKS / BREACHES (security of virtual machines from intrusion and invasion)
Hacks / breaches constitute a concern over the level of security that is present to protect virtual machines and their operations running on a provider host from external sources (people who are neither providers nor co-tenants). This concern remains one of the top and most recurrent identified by various researchers and primarily arising from denial of service (DoS), distributed denial of service (DDoS), as well as infiltration attacks. An instance of this concern occurred in September 2014, when the iCloud server (a cloud storage facility owned and operated by Apple Inc.) experienced a major hack / breach that led to the leakage of private photographs belonging to top, popular Hollywood celebrities by the hackers, thereby putting the computing giant in bad light [36]. Also, a recent survey by the International Data Corporation (IDC) revealed that 87.5% of people surveyed at various levels (from IT executives to CEOs) opine that security remains the top-most challenge in every cloud service infrastructure [37] and [23] further pointed out security as the highest risk issue in virtualization and cloud computing with some IT executives holding the view that not all applications or data should be virtualized and/or put in the cloud.

BOUNDARY ENFORCEMENT / RESTRICTION. Every virtual machine residing on a host is entitled to a virtual space which is confined within virtual boundaries. Boundary Enforcement / Restriction expresses concern over the extent to which tenants' boundaries are carefully marked off and how much they are able to breach or push their virtual boundaries / limits as are clearly and technically mapped out in service level (SL) agreements. Clients would also raise concerns relating to what degree other tenants would be able to push or extend their virtual spaces and boundaries and begin to breach the space of other guests residing on the host. SL breaches are becoming a major concern to cloud providers as more research and technologies have gone into the task of trying to enforce this to the uttermost. The boundary issue becomes a bigger one when it is expanded in relation to property rights, contracts, and tenancy laws which in reality have industry and geo-political intricacies.

INTEGRATION & COMPATIBILITY. Integration is the concern about the ability of the host environment to receive the virtual guest (client) along with its already existing service and application requirements and programs. Compatibility is the concern over how able the host environment can cope with, and satisfy client requirements and program needs throughout the period of the tenancy (service level) agreement. Research by [11] revealed that some organizations are uncertain as to which virtualized applications, servers and environments would be able to satisfy completely the needs that were assuaged by their current IT framework which other raised the issue of compatibility with their already existing hardware resources.

Furthermore, and because virtualization is the core that powers cloud computing, it is no gainsaying that the security issues that bother on virtualization, to a large extent, also bother on cloud computing especially when viewed as a symbiotic relationship. Some of the most extensive studies represented by such sources as [14], [23], [27], and [34] highlighted in table 3, attempts a summary of the known common / major threats to virtualization, as well the solutions to these; grouped under the various ethical concerns highlighted in the above section.

Table 3: Summary of Ethical Concerns, Known Threats and Solutions [Source: authors]

ETHICAL CONCERNS	THREATS	SUGGESTED SOLUTIONS
CO-TENANCY	VM Introspection; Meta-data Copying; VM Information Leakage; Hijacking	Hypervisor Integrity and Hardening; Signature Verification; Monitoring, Auditing and Administration; Strong Passwords Enforcement for VMs; Encryption
TRUST & PRIVACY	VMM Subversion / Insertion; VM Cloning; Illegal VM Snapshots; Shared Folder Copying; Privilege Escalation; VM Information Leakage; VM Escape; Hijacking	VM Isolation; Hypervisor Integrity and Hardening; Shared Folder Encryption; Virtual Roles Separation
HACKS / BREACHES	VM Cloning; DoS and DDoS; VM Sniffing and Spoofing; Arbitrary Code Executions; Virtual Network Compromise; Hijacking	Hypervisor Security and Hardening; Monitors and filters on networks; Patches and Security Fixes; Intrusion Detection and Prevention Systems (IDSs and IPSs)
BOUNDARY ENFORCEMENT / RESTRICTION	VM Scaling; SLA Breaches; VM Masking; Malware corruption from other VMs and/or Host; VM Escape; VLAN Hopping; VM Promiscuity.	VM Isolation; SLA Enforcement; Resource Allocation; Virtual Boundary Separation; Anti-malware installations
INTEGRATION & COMPATIBILITY	Runtime Errors; Hardware and Software Incompatibilities; Corrupt Virtual Images	Sandboxing; Virtual Hardware and Software Verification; Disabling incompatible virtual hardware and software; Proper Backups and Regular Continuity Protection Measures;

VI. CONCLUSION

Virtualization has proven time and again to be one of the most promising technologies for business performance enhancement, sustainability and productivity. However, frequently recurring ethical concerns as well as the potential threats that are brought up due to the nature of flexibility which virtualization provides have kept begging for answers. A new course has been pointed out in this paper as clear ethical concerns of clients and subscribers regarding virtualization and cloud computing; cloud service providers are now better informed on how to directly address these ethical concerns by tapping from the solutions suggested herein. It must however be mentioned at this point that the technology of virtualization has undergone rapid, commendable improvements in various forms in recent times to provide answers and solutions to some of the challenges that previously plagued future prospects of this technology; and in line with these, this research has gone on to suggest a couple more.

VII. REFERENCES

[1] Davies, A. (2004, June). Computational intermediation and the evolution of computation as a commodity. *Applied Economics, 36*(11: 1131). doi:10.1080/0003684042000247334

[2] Christopher, S. (1959). Time Sharing in Large Fast Computers. *Proceedings of the International Conference on Information processing, UNESCO. 2.19*, pp. 336-341. UNESCO. Retrieved February 1, 2014

[3] IBM. (1970). *IBM System/360 Operating System: Conversational Remote Job Entry Concepts and Facilities.* International Business Machine (IBM). North Carolina, USA.: IBM Systems Reference Library. Retrieved February 1, 2014, from http://bitsavers.informatik.uni-stuttgart.de/pdf/ibm/360/rje/GC30-2012-0_CRJE_Concepts_and_Facilities_Jun70.pdf

[4] Nikiforakis, N., Joosen, W., & Johns, M. (2011). Abusing Locality in Shared Web Hosting. *Proceedings of the Fourth European Workshop on System Security: Article No. 2.* Salzburg, Austria:

Association for Computing Machinery. doi:10.1145/1972551.1972553

[5] Urgaonkar , B., Shenoy, P., & Roscoe, T. (2009, February). Resource overbooking and application profiling in a shared Internet hosting platform. *ACM Transactions on Internet Technology (TOIT): Article No. 1 , 9*(1).

[6] Bhattiprolu, S., Biederman, E. W., Hallyn, S., & Lezcano, D. (2008, July). Virtual Servers and Checkpoint/Restart in Mainstream. *ACM SIGOPS Operating Systems Review - Research and developments in the Linux kernel*, *42*(5), 104-113. doi:10.1145/1400097.1400109

[7] Buyya, R., & Bubendorfer, K. (2009). *Market-Oriented Grid and Utility Computing.* Wiley Publishing.

[8] Mell, P., & Grance, T. (September 2011). *The NIST Definition of Cloud Computing.* Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology, United States Department of Commerce. Gaithersburg, MD 20899-8930: National Institute of Standards and Technology. Retrieved January 28, 2014, from http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

[9] Ryan, S., & Jiangchuan, L. (2012). Understanding the Impact of Denial of Service Attacks on Virtual Machines. *Journal of the IEEE*.

[10] Tupakula, U., & Varadharajan, V. (2011). TVDSEC: Trusted Virtual Domain Security. *Institute of Electrical and Electronic Engineers (IEEE)*, 57-63.

[11] CDW Corporation. (2010, January 11). *CDW Server Virtualization Life Cycle Report (Medium and Large Businesses).* Retrieved from CDW Newsroom: http://webobjects.cdw.com/webobjects/media/pdf/Newsroom/CDW-Server-Virtualization-Life-Cycle-Report.pdf

[12] Strømmen-Bakhtiar, A., & Razavi, A. R. (2011). Cloud Computing Business Models. *Springer Computer Communications and Networks*, 43-60.

[13] Gartner AADI Summit. (2009). *Cloud Computing as Gartner Sees it.* Gartner's Application Architecture, Development & Integration Summit.

[14] Pearce, M., Zeadally, S., & Hunt, R. (2013, February). Virtualization: Issues, Security Threats, and Solutions. *Association for Computing Machinery (ACM) Computing Surveys*, Article 17: 1-39.

[15] Silberschatz, A., Galvin, P. B., & Gagne, G. (2013). *Operating System Concepts (Ninth Edition).* U.S.A.: Wiley Publishers.

[16] Chowdhury, K. N., & Boutaba, R. (2009, July). Network virtualization: state of the art and research challenges. *IEEE Communications Magazine*, *47*(7), 20-26.

[17] Soundararajan, V., & Anderson, J. M. (2010). The impact of management operations on the virtualized datacenter. *Proceedings of the 37th annual international symposium on Computer architecture* (pp. 326-337`). New York, NY, USA: ACM.

[18] Miller, K., & Pegah, M. (2007). Virtualization: virtually at the desktop. *Proceedings of the 35th annual ACM SIGUCCS fall conference* (pp. 255-260). New York, NY, USA: ACM.

[19] Bratus, S., Johnson, P. C., Ramaswamy, A., Smith, S. W., & Locasto, M. E. (2009). The cake is a lie: privilege rings as a policy resource. *Proceedings of the 1st ACM workshop on Virtual machine security* (pp. 33-38). Illinois, USA.: Association for Computing Machinery.

[20] Padala, P., Zhu, X., Wang, Z., Singhal, S., & Shin, K. G. (2007). *Performance Evaluation of Virtualization Technologies for Server Consolidation.* Hewlett Packard, Enterprise Systems and Software Laboratory. Palo Alto: Hewlett Packard. Retrieved February 2, 2014, from http://137.204.107.78/tirocinio/site/tirocini/Tirocinio-Zuluaga/Documents/virtualizzazione/Technologies%20for%20Server.pdf

[21] Popek, G. J., & Goldberg, R. P. (1974). Formal Requirements for Virtualizale Third Generation Architectures. *Proceedings of the fourth ACM symposium on Operating system principles*, 412-421.

[22] Sugerman, J., Venkitachalam, G., & Lim, B. - H. (2001). Virtualizing I/O Devices on VMware Workstation's Hosted Virtual Machine Monitor. *Proceedings of the General Track: 2002 USENIX Annual Technical Conference* (pp. 1-14). Berkeley, CA, USA.: USENIX Association.

[23] Nagaraju, K., & Sridaran, R. (2012, September). A Survey on Security Threats for Cloud Computing. *International Journal of Engineering Research & Technology (IJERT), Volume 1*(Issue 7), 1-10.

[24] Seshadri, A., Luk, M., Qu, N., & Perrig, A. (2007). SecVisor: A Tiny Hypervisor to provide lifetime Kernel code integrity for commodity OSes. *Proceedings of the 21st ACM SIGOPS Symposium on Operating Systems Principles* (pp. 335-350). Association for Computing Machinery.

[25] Franklin, J., Seshadri, A., Qu, N., Chaki, S., & Datta, A. (2008). *Attacking, Repairing, and Verifying SecVisor: A Retrospective on the Security of a Hypervisor.* Cylab Technical Report. CMU-Cylab-08-008.

[26] Hirano, M., Shinagawa, T., Eiraku, H., Hasegawa, S., Omote, K., Okuda, T., . . . Yamaguchi, S. (2009). A Two-step Execution Mechanism for Thin Secure Hypervisors. *Third International Conference on Emerging Security Information Systems and Technologies, IEEE.* (pp. 129-134). Institute of Electrical and Electronic Engineers.

[27] Pék, G., Butty´an, L., & Bencsáth, B. (2013, June). A survey of security issues in hardware virtualization. *ACM Computing Surveys (CSUR): Article No. 40, 45*(3). doi:10.1145/2480741.2480757

[28] Texas Reliability Entity. (2013). *Virtualization and Cloud Computing. "Security is a Process, not a Product".* Texas, USA.: Texas Reliability Entity.

[29] Kretschmer, T. (2010). *"WORK – Working Connected in Business and Society".* Bonn, Germany.: Deutsche Telekom AG.

[30] Griffith, R. (2007). Technology, Productivity and Public Policy. *Fiscal Studies, 28*(3), 273–291.

[31] Teece, D. J. (1986). *Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy.* Berkeley, CA 94720, U.S.A.: School of Business Administration, University ofCalifornia.

[32] Sembok, T. M. (November 2003). Ethics of Information Communication Technology (ICT). *Proceedings of the Regional Meeting on Ethics of Science and Technology* (pp. 239-325). Bangkok: UNESCO: Regional Unit for Social & Human Sciences in Asia and the Pacific (RUSHSAP).

[33] Maner, W. (1980). *Starter Kit in Computer Ethics.* Helvetia Press in collaboration with the National Information and Resource Center for Teaching Philosophy.

[34] Garfinkel, T., & Rosenblum, M. (2005). When virtual is harder than real: security challenges in virtual machine based computing environments. *Proceedings of the 10th conference on Hot Topics in Operating Systems* (pp. 20-20). Berkeley, CA, USA: USENIX Association.

[35] Winkler, V. J. (2011). *Securing the Cloud: Cloud Computer Security Techniques and Tactics.* Syngress Publishing.

[36] Duke, A. (2014, October 12). *CNN Entertainment (5 Things to know about the celebrity nude photo hacking scandal)*. Retrieved November 1, 2014, from CNN International: http://www.cnn.com/2014/09/02/showbiz/hacked-nude-photos-five-things/

[37] Lv, H., & Hu, Y. (2011). Analysis and Research about Cloud Computing Security Protect Policy. *Institute of Electrical and Electronic Engineers (IEEE)*, 214-216.

Analysis of Two-Gate Flip Flop versus Four-Gate Flip Flop

Professor Olawale J. Omotosho
School of Computing and Engineering Sciences
Babcock University, Ilishan-Remo, Ogun State
ojomotosho1@yahoo.com, oomotosho@babcockuni.ed.ng

Abstract – The design of SR-Flip Flops is principally governed by three states of transitions named RESTING, ACTIVE & FORBIDDEN states using a Two-Input device that is capable to flip flop or to be switched ON and OFF, such as Bistable Multivibrators. The allocation of these three transition states among the four possible combinations of the input variables give rise to 12 possible ways an SR-Flip Flop can be designed. Furthermore, the RESTING and FORBIDDEN transition states may be converted into useful states like the ACTIVE states in different ways, giving rise to different Flip Flops with different percentage of usefulness of the transition states. In the design of SR-Flip Flops, it is conventional to use only a particular type of logic gates to obtain the required hardware instead of using combination of different types of logic gates which leads to high cost in manufacturing and maintenance. Consequently, two NOR or NAND logic gates have been employed to build an SR-Flip Flop as published in many literature around the world. This has been found to be inadequate to satisfy the complete functions of an SR-Flip Flop. The proof of this discovery is the subject of this paper.

Key Words: Flip Flop, Bistable Multivibrators, Resting, Active & Forbidden.

I. INTRODUCTION

Conventionally, many authors have presented an SR-Flip Flop as a 2-Gate digital network comprising of either NOR or NAND gates without any convincing proof for its validity. However, design emanating from the properties of an SR Flip Flop produces a 4-Gate digital circuit. Therefore, there is a need to analyse the 2-Gate Configurations and compare them with the 4-Gate Configurations as to determine which of them is correct to represent an SR-Flip Flop. This is the purpose of this paper.

II. ANALYSIS OF TWO-NOR-GATE CONFIGURATION OF AN SR FLIP FLOP

The approach to be used in this paper is by posing an appropriate question such as:
Given the two configurations (Fig 2.1 & Fig 2.2a), identify the configuration that is probable to be used as an SR-Flip Flip stating the conditions considered for your choice. Can this choice of yours be used to satisfy fully the characteristics of a conventionally accepted SR-Flip Flop? If not, what modifications are necessary to meet the design requirements? If yes, how does it achieve the required characteristics?

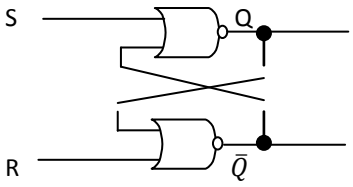


Fig 2.1 (NOR-Gate Config)

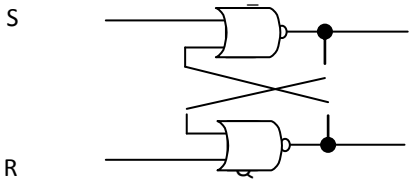


Fig 2.2a (NOR-Gate Config)

Solution:

From Fig 2.1,

$$Q_{n+1}(1) = \overline{S + \bar{Q}_n} = \bar{S}Q_n = 0 - 1 = 001,011 \equiv \langle 1 \rangle \dots\dots\dots\dots\dots\dots\dots\dots (a)$$
$$\bar{Q}_{n+1}(0) = \overline{R + Q_n} = \bar{R}\bar{Q}_n = -00 = 000,100 \equiv \langle 0 \rangle \dots\dots\dots\dots\dots\dots\dots\dots (b)$$

The Truth Table of an SR-Flip Flop (NOR Gate configuration) is given in Table 2.1

Table 2.1: Truth Table of SR Flip Flop (NOR Gate Configuration)

S/N	S	R	Q_n	Q_{n+1}	Logic Equation Corresponding to each transition	Remarks
0	0	0	0	0	$\bar{S}\bar{R}\bar{Q}_n = 000 = \langle 0 \rangle$	Equation (b) (Resting)
1	0	0	1	1	$\bar{S}\bar{R}Q_n = 001 = \langle 1 \rangle$	Equation (a) (Resting)
2	0	1	0	0	$\bar{S}R\bar{Q}_n = 010 = \langle 0 \rangle$	State not covered
3	0	1	1	0	$\bar{S}RQ_n = 011 = \langle 0 \rangle$	Equation (a)
4	1	0	0	1	$S\bar{R}\bar{Q}_n = 100 = \langle 1 \rangle$	Equation (b)
5	1	0	1	1	$S\bar{R}Q_n = 101 = \langle 1 \rangle$	State not covered
6	1	1	0	x	$SR\bar{Q}_n = 110 = \langle x \rangle$	Forbidden State
7	1	1	1	x	$SRQ_n = 111 = \langle x \rangle$	Forbidden State

NOTE: From this table, only two transitions/states (SRQ_n = 011 & 100) out of four Active states are accomplished by Figure 2.1 while there are two more transitions/states (SRQ_n = 010 & 101) that are not covered. Hence, Figure 2.1 cannot be a full representation of an SR Flip Flop.

Analytically, the same thing can also be proved as follows:
From the Truth Table of an SR-Flip Flop (NOR Gate configuration), we have
$SRQ_n = 110,111 \equiv \langle x \rangle$ …….. (c)

Plotting equations (a), (b) & (c) into a K-Map and a Truth Table for all the variables of 'S' and 'R' we have what is presented in Table 2.2.

Table 2.2: TRUTH TABLE for Fig 2.1

S	R	Q_n	Q_{n+1}		REMARKS
0	0	0	0	The Value of Q_{n+1} is determined from equation (b)	Resting States
0	0	1	1	The Value of Q_{n+1} is determined from equation (a)	
0	1	0	A	**This value of Q_{n+1} is indeterminate**	A = 0 or 1
0	1	1	1	The Value of Q_{n+1} is determined from equation (a)	The output follows 'R'
1	0	0	0	The Value of Q_{n+1} is determined from equation (b)	The output follows 'R'
1	0	1	B	**This value of Q_{n+1} is indeterminate**	B = 0 or 1
1	1	0	x	The Value of Q_{n+1} is determined from equation (c)	Forbidden State
1	1	1	x	The Value of Q_{n+1} is determined from equation (c)	

K-MAP

SR / Q_n	00	01	11	10
0	0	A	X	0
1	1	1	X	B

$Q_{n+1} = \bar{S}Q_n = \overline{S + \bar{Q}_n}$
A = B = 0

NOTES:

1. The Resting & Forbidden states are satisfied as a NOR-gate SR-Flip Flop.
2. The Active states partly follow 'R' suggesting that Fig 2.1 could be a NAND-gate SR-Flip Flop. That is, Q_{n+1} = R.
3. The Active states partly give 2 indeterminate outputs A & B. That is, the outputs can either follow 'S' or 'R' depending on the unspecified values of A & B. That is, Fig 2.1 could be either NAND-gate or NOR-gate SR-Flip Flop, depending on the values of 'A' and 'B'.

Hence, this configuration (Fig 2.1) does not satisfy the required conditions for a NOR-gate SR-Flip Flop.

Similarly, from Fig 2.2a,

$$Q_{n+1}(1) = \overline{R + \bar{Q}_n} = \bar{R}Q_n = -01 = 001, 101 \equiv \langle 1 \rangle \dots\dots\dots\dots\dots\dots\dots (a)$$

$$\bar{Q}_{n+1} = \overline{S + Q_n} = \bar{S}\bar{Q}_n = 0 - 0 = 000, 010 \equiv \langle 0 \rangle \dots\dots\dots\dots\dots\dots\dots (b)$$

From SR-Flip Flop (+ve Logic configuration), $SRQ_n = 110,111 \equiv \langle x \rangle$ …… (c)

Plotting equations (a), (b) & (c) into a K-Map and a Truth Table for all the variables of 'S' and 'R' we have what is presented in Table 2.3.

Table 2.3: TRUTH TABLE for Fig 2.2a

S	R	Q_n	Q_{n+1}		REMARKS
0	0	0	0	The Value of Q_{n+1} is determined from equation (b)	Resting States
0	0	1	1	The Value of Q_{n+1} is determined from equation (a)	
0	1	0	0	The Value of Q_{n+1} is determined from equation (b)	The output follows 'S'
0	1	1	A	**This value of Q_{n+1} is indeterminate**	A = 0 or 1
1	0	0	B	**This value of Q_{n+1} is indeterminate**	B = 0 or 1
1	0	1	1	The Value of Q_{n+1} is determined from equation (a)	The output follows 'S'
1	1	0	x	The Value of Q_{n+1} is determined from equation (c)	Forbidden State
1	1	1	x	The Value of Q_{n+1} is determined from equation (c)	

K-MAP

SR	00	01	11	10
0	0	0	x	B
1	1	A	x	1
Q_n				

$Q_{n+1} = \bar{R}Q_n = \overline{R + \bar{Q}_n}$

NOTES:

1. The Resting & Forbidden states are satisfied as a NOR-gate SR-Flip Flop.
2. The Active states partly follow 'S' suggesting that Fig 2.2a could be a NOR-gate SR-Flip Flop. That is, $Q_{n+1} = S$.
3. The Active states partly give 2 indeterminate outputs, A & B. That is, the outputs can either follow 'S' or 'R' depending on the unspecified values of A & B. That is, Fig 2.2a could be either NAND-gate or NOR-gate SR-Flip Flop, depending on the values of 'A' and 'B'.

Hence, this configuration (Fig 2.2a) does not fully satisfy the characteristics of an SR-Flip Flop because only half of its Active states will produce the required output while the other half may or may not produce the required output. However, this configuration is more acceptable than Fig 2.1 because half of its output follows 'S' as required for a NOR-gate SR-Flip Flop.

The modification required for Fig 2.2ato fully satisfy the characteristics of a NOR-gate SR-Flip Flop as presented in Table 2.1 is chosen from the four possible options presented on Table 2.4. The four options are obtained from the fact that the configuration produced two indeterminate states, thereby producing four different combinations (2^2 = 4) of these two states. That is, the combinations are AB = 00_2 (0_{10}), AB = 01_2 (1_{10}), AB = 10_2 (2_{10}) and AB = 11_2 (3_{10}).

Table 2.4: K-Maps for the various Options of Fig 2.2a

$Q_{n+1} = \bar{R}Q_n$ …. (0)

SR	00	01	11	10
0	0	0	x	B,0
1	1	A,0	x	1
Q_n				

AB = 00

$Q_{n+1} = S + \bar{R}Q_n$ ….(1)

AB = 01

SR	00	01	11	10
0	0	0	x	B,1
1	1	A,0	x	1
Q_n				

$Q_{n+1} = Q_n$ … (2)

SR	00	01	11	10
0	0	0	x	B,0
1	1	A,1	x	1
Q_n				

AB = 10

$Q_{n+1} = S + Q_n$ … (3)

AB = 11

SR	00	01	11	10
0	0	0	x	B,1
1	1	A,1	x	1
Q_n				

The required modification is only when AB = 01_2 (1_{10}) for an SR-Flip Flop that is designed with positive logic (NOR Gates Configuration). That is, all the Active states are required to follow 'S'

Similarly, the modification required for Fig 2.1 is chosen from the four possible options as presented in Table 2.5.

Table 2.5: K-Maps for the various Options of Fig 2.1

$Q_{n+1} = RQ_n$ … (0)

SR	00	01	11	10
0	x	A,0	0	0
1	x	1	1	B,0
Q_n				

AB = 00 = 10

$Q_{n+1} = \bar{S} + RQ_n$ ….(2)

AB

SR	00	01	11	10
0	x	A,1	0	0
1	x	1	1	B,0
Q_n				

$Q_{n+1} = Q_n$ … (1)

SR	00	01	11	10
0	x	A,0	0	0
1	x	1	1	B,1
Q_n				

AB = 01 = 11

$Q_{n+1} = \bar{S} + Q_n$ …(3)

AB

SR	00	01	11	10
0	x	A,1	0	0
1	x	1	1	B,1
Q_n				

The required modification is only when AB = 10_2 (2_{10}) for an SR-Flip Flop that is designed with negative logic (NAND Gates Configuration). That is, all the Active states are required to follow 'R'. In addition, the Resting and Forbidden states are to be exchanged. That is, SRQ_n (Forbidden states) = 000, 001 = x, and SRQ_n (Resting States) = 110 ≡ 0, 111 ≡ 1.

Alternative Solution

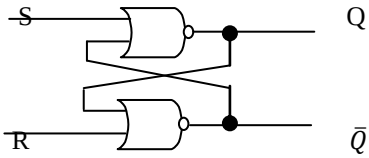


Fig 2.1 (NOR-Gate Config)

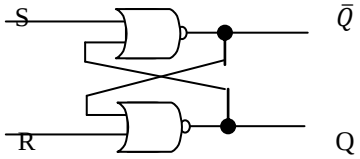


Fig 2.2a (NOR-Gate Config)√

From Fig 2.1,

$$Q_{n+1} = \overline{S + \bar{Q}_n} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (a)$$
$$\bar{Q}_{n+1} = \overline{R + Q_n} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (b)$$

If $Q_{n+1} = Q_n$, substituting equation (a) into equation (b) and vice-visa, we have

$$Q_{n+1}(1) = \overline{S + \overline{R + Q_n}} = \overline{S + \bar{R}\bar{Q}_n} \equiv \langle 1 \rangle \ldots\ldots\ldots\ldots\ldots\ldots\ldots (a1)$$
$$\bar{Q}_{n+1}(0) = \overline{R + \overline{S + \bar{Q}_n}} = \overline{R + \bar{S}Q_n} \equiv \langle 0 \rangle \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (b1)$$

From SR-Flip Flop (+ve Logic configuration), $SRQ_n = 110,111 \equiv \langle x \rangle$ …… (c)

Plotting equations (a1), (b1) & (c) into a K-Map and a Truth Table for all the variables of 'S' and 'R' we have what is presented in Table 2.6.

Examples: For S = 0, R = 1, Q_n = 0, we have from equation (a1), $Q_{n+1} = \bar{0} = 1$ and from equation (b1), $\bar{Q}_{n+1} = \bar{R} = \bar{1} = 0$. This procedure is continued for SRQ_n = 011, 100 & 101

Table 2.6: TRUTH TABLE for Fig 2.1

S	R	Q_n	Q_{n+1}	$\bar{Q}_{n+1}$	REMARKS
0	0	0	0	1	Resting State
0	0	1	1	0	
0	1	0	1	0	The output follows 'R'. That is, Q_{n+1} = R
0	1	1	1	0	The output follows 'R'. That is, Q_{n+1} = R
1	0	0	0	1	The output follows 'R'. That is, Q_{n+1} = R
1	0	1	0	1	The output follows 'R'. That is, Q_{n+1} = R

K-MAP

SR	00	01	11	10
Q_n 0	0	1	x	0
1	1	1	x	0

$Q_{n+1} = R + \bar{S}Q_n$

1	1	0	x	x	Forbidden State
1	1	1	x	x	

NOTES:

1. The Resting and Forbidden states are satisfied as a NOR-gate SR-Flip Flop.
2. The Active states fully follow 'R' as required for a NAND-gate SR-Flip Flop.This is only possible if it can be ascertained that at all times, the signals arriving at terminals 'S' and 'R' are received at different time and that the propagation times (τ) of both NOR gates are the same. That is, $Q_n = Q_{n+1}$ and $\tau_S = \tau_R$.

Hence, this configuration (Fig 2.1) does not fully satisfy the characteristics of a NAND-gate SR-Flip Flop because its Resting and Forbidden states are wrongly assigned.

From Fig 2.2a,

$$Q_{n+1} = \overline{R + \bar{Q}_n} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (a)$$
$$\bar{Q}_{n+1} = \overline{S + Q_n} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (b)$$

If $Q_{n+1} = Q_n$, substituting equation (a) into equation (b) and vice-visa, we have

$$Q_{n+1}(1) = \overline{R + \overline{S + Q_n}} = \overline{R + \bar{S}\bar{Q}_n} \equiv \langle 1 \rangle \ldots\ldots\ldots\ldots\ldots\ldots\ldots (a1)$$
$$\bar{Q}_{n+1}(0) = \overline{S + \overline{R + \bar{Q}_n}} = \overline{S + \bar{R}Q_n} \equiv \langle 0 \rangle \ldots\ldots\ldots\ldots\ldots\ldots\ldots (b1)$$

From SR-Flip Flop (+ve Logic configuration), $SRQ_n = 110,111 \equiv \langle x \rangle$ …… (c)

Plotting equations (a1), (b1) & (c) into a K-Map and a Truth Table for all the variables of 'S' and 'R' we have what is presented in Table 2.7.

Examples: For S = 0, R = 1, Q_n = 0, we have from equation (a1), $Q_{n+1} = \bar{1} = 0$ and from equation (b1), $\bar{Q}_{n+1} = \bar{0} = 1$. This procedure is continued for SRQ_n = 011, 100 & 101

Table 2.7: TRUTH TABLE for Fig 2.2a

S	R	Q_n	Q_{n+1}	$\bar{Q}_{n+1}$	REMARKS
0	0	0	0	0	Resting State
0	0	1	1	0	
0	1	0	0	1	The output follows 'S'. That is, Q_{n+1} = S
0	1	1	0	1	The output follows 'S'. That is, Q_{n+1} = S
1	0	0	1	0	The output follows 'S'. That is, Q_{n+1} = S
1	0	1	1	0	The output follows 'S'. That is, Q_{n+1} = S
1	1	0	x	x	Forbidden State
1	1	1	x	x	

K-MAP

SR	00	01	11	10
Q_n 0	0	0	x	1
1	1	0	x	1

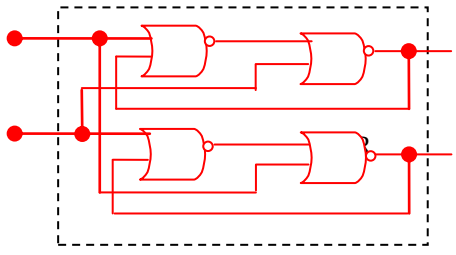
$Q_{n+1} = S + \bar{R}Q_n$

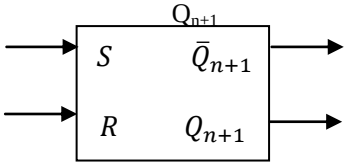
NOTES:

1. The Resting and Forbidden states are satisfied as a NOR-gate SR-Flip Flop
2. The Active states fully follow 'S' as required for a NOR-gate SR-Flip Flop.

Hence, this configuration (Fig 2.2a) does fully satisfy the characteristics of a NOR-gate SR-Flip Flop. Even though, this is only possible if it can be ascertained that at all times, the signals arriving at terminals 'S' and 'R' are received at different time and that the propagation times (τ) of both NOR gates are the same. That is, $Q_n = Q_{n+1}$ and $\tau_S = \tau_R$.

The logical circuit that satisfies fully the characteristics presented in Table 2.1 is shown in Fig 2.2b. The details of the synthesis/design are presented somewhere else.





Flip Flop Symbol (NOR)

Fig 2.2b: Four-NOR-gate SR-Flip Flop Configuration

III. ANALYSIS OF TWO-NAND-GATE CONFIGURATION OF AN SR FLIP FLOP

Similarly, the same approach of posing an appropriate question is employed here as follows:
Given the two configurations (Fig 3.1a& Fig 3.2), identify the configuration that is probable to be used as an SR-Flip Flip stating the conditions considered for your choice. Can this choice of yours be used to satisfy fully the characteristics of a conventionally accepted SR-Flip Flop? If not, what modifications are necessary to meet the design requirements? If yes, how does it achieve the required characteristics?

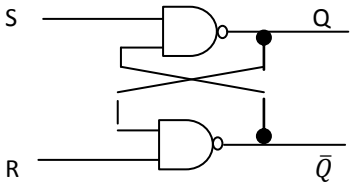


Fig 3.1a (NAND-GateConfig)√

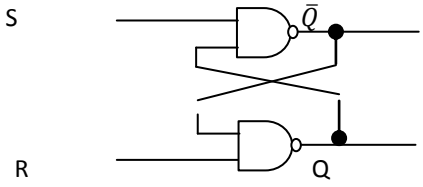


Fig 3.2 (NAND-GateConfig)

From Fig 3.1a,

$$Q_{n+1}(1) = \overline{S\bar{Q}_n} = \overline{1-0} = 011,001 \equiv \langle 1\rangle \dots\dots\dots\dots\dots\dots\dots\dots (a)$$
$$\bar{Q}_{n+1}(0) = \overline{RQ_n} = \overline{-11} = 100{,}000 \equiv \langle 0\rangle \dots\dots\dots\dots\dots\dots\dots\dots (b)$$

From SR-Flip Flop (NAND-gate configuration), $SRQ_n = 000{,}001 \equiv \langle x\rangle$ …….. (c)

The Truth Table of an SR-Flip Flop (NAND-gate configuration) is given in Table 3.1

Table 3.1: Truth Table of SR Flip Flop (NAND-gate Configuration)

S/N	S	R	Q_n	Q_{n+1}	Logic Equation Corresponding to each transition	Remarks
0	0	0	0	x	$\bar{S}\bar{R}\bar{Q}_n = 000 = \langle x\rangle$	Equation (c) (Forbidden)
1	0	0	1	x	$\bar{S}\bar{R}Q_n = 001 = \langle x\rangle$	Equation (c) (Forbidden)
2	0	1	0	0	$\bar{S}R\bar{Q}_n = 010 = \langle 0\rangle$	State not covered
3	0	1	1	0	$\bar{S}RQ_n = 011 = \langle 0\rangle$	Equation (a)
4	1	0	0	1	$S\bar{R}\bar{Q}_n = 100 = \langle 1\rangle$	Equation (b)
5	1	0	1	1	$S\bar{R}Q_n = 101 = \langle 1\rangle$	State not covered
6	1	1	0	0	$SR\bar{Q}_n = 110 = \langle 0\rangle$	Resting State
7	1	1	1	1	$SRQ_n = 111 = \langle 1\rangle$	Resting State

NOTE: From this table, only two transitions/states (SRQ_n = 011 & 100) out of four Active states are accomplished by Figure 3.1a while there are two more transitions/states (SRQ_n = 010 & 101) that are not covered. Hence, Figure 3.1a cannot be a full representation of an SR Flip Flop.

Plotting equations (a), (b) & (c) into a K-Map and a Truth Table for all the variables of 'S' and 'R' we have what is presented in Table 3.2.

Table 3.2: TRUTH TABLE for Fig 3.1a

S	R	Q_n	Q_{n+1}		REMARKS
0	0	0	0,x	The Value of Q_{n+1} is determined from equations (b) & (c)	Forbidden State, x. For this state, x is preferred.
0	0	1	1, x	The Value of Q_{n+1} is determined from equations (a) & (c)	Forbidden State, x. For this state, x is preferred.
0	1	0	A	**This value of Q_{n+1} is indeterminate**	A = 0 or 1
0	1	1	1	The Value of Q_{n+1} is determined from equation (a)	The output follows 'R'
1	0	0	0	The Value of Q_{n+1} is determined from equation (b)	The output follows 'R'
1	0	1	B	**This value of Q_{n+1} is indeterminate**	(B = 0 or 1)
1	1	0	C	**This value of Q_{n+1} is indeterminate**	Resting State, 0 (C = 0 or 1)

K-MAP

SR	00	01	11	10
0	0,x	A	C	0
1	1,x	1	D	B
Q_n				

$Q_{n+1} = \bar{S}Q_n = \overline{S + \bar{Q}_n}$
If A = B = C = D = 0

1	1	1	D	**This value of Q_{n+1} is indeterminate**	Resting State, 1(D = 0 or 1)

NOTES:

1. The Resting & Forbidden states do not satisfy any SR-Flip Flop because of the indeterminate states (SRQ_n = 010 [A], 101 [B], 110 [C], 111 [D]) and (SRQ_n = 000 [0,x], 001 [1,x]).
2. Only two Active states (SRQ_n = 011 [1], 100 [0]) follows 'R' as expected of a NAND-gate SR-FF. The other two Active states (SRQ_n = 010 [A], 101 [B]) are indeterminate.

Hence, this configuration (Fig 3.1a) does not satisfy the required conditions for either a NAND-gate or a NOR-gate SR-Flip Flop.

From Fig 3.2,

$$Q_{n+1}(1) = \overline{R\bar{Q}_n} = \overline{-10} = 101,001 \equiv \langle 1 \rangle \dots\dots\dots\dots\dots\dots\dots (a)$$

$$\bar{Q}_{n+1}(0) = \overline{SQ_n} = \overline{1-1} = 010,000 \equiv \langle 0 \rangle \dots\dots\dots\dots\dots\dots\dots\dots (b)$$

From SR-Flip Flop (+ve Logic configuration), $SRQ_n = 000,001 \equiv \langle x \rangle$ …… (c)

Plotting equations (a), (b) & (c) into a K-Map and a Truth Table for all the variables of 'S' and 'R' we have what is presented in Table 3.3.

Table 3.3: TRUTH TABLE for Fig 3.2

S	R	Q_n	Q_{n+1}		REMARKS
0	0	0	0,x	The Value of Q_{n+1} is determined from equations (b)& (c)	Forbidden State, x. For this state, x is preferred.
0	0	1	1, x	The Value of Q_{n+1} is determined from equations (a)& (c)	Forbidden State, x. For this state, x is preferred.
0	1	0	0	The Value of Q_{n+1} is determined from equation (b)	The output follows 'S'
0	1	1	A	**This value of Q_{n+1} is indeterminate**	A = 0 or 1
1	0	0	B	**This value of Q_{n+1} is indeterminate**	B = 0 or 1
1	0	1	1	The Value of Q_{n+1} is determined from equation (a)	The output follows 'S'
1	1	0	C	**This value of Q_{n+1} is indeterminate**	Resting State, 0 (C = 0 or 1)
1	1	1	D	**This value of Q_{n+1} is indeterminate**	Resting State, (D = 0 or 1)

K-MAP

SR	00	01	11	10
0	0,x	0	C	B
1	1,x	A	D	1
Q_n				

$Q_{n+1} = \bar{R}Q_n = \overline{R + \bar{Q}_n}$
If A = B = C = D = 0

NOTES:

1. The Resting states (SRQ_n = 110 [C], 111 [D]) failed to satisfy any SR-Flip Flop while the Forbidden states (SRQ_n = 000 [0,x], 001 [1,x]) are both forbidden and resting.
2. Only two Active states (SRQ_n = 010 [0], 101 [1]) follows 'S' as required for a **NOR**-gate SR-Flip Flop. The other two Active states (SRQ_n = 011 [A], 100 [B]) are indeterminate.

Comparing the analysis of both Fig 3.1a (Table 3.2)& Fig 3.2 (Table 3.3), Fig 3.1ais closer to a NAND-gate SR-Flip Flop than Fig 3.2 as indicated.

Additional modification required for Fig 3.1a(Table 3.2) to function fully as a NAND-gate SR-Flip Flop is chosen from the sixteen possible options presented on Table 3.4.1. The sixteen options are obtained from the fact that the configuration produced four indeterminate states (SRQ_n = 011, 100, 110 & 111 of Table 3.2), thereby producing sixteen different combinations (2^4 = 16) of these four states. Thus, the combinations are ABCD = 0000_2 (0_{10}), ABCD = 0001_2 (1_{10}), ABCD = 0010_2 (2_{10}), ABCD = 0011_2 (3_{10}), ABCD = 0100_2 (4_{10}), ABCD = 0101_2 (5_{10}), ABCD = 0110_2 (6_{10}), ABCD = 0111_2 (7_{10}), ABCD = 1000_2 (8_{10}), ABCD = 1001_2 (9_{10}), ABCD = 1010_2 (10_{10}), ABCD = 1011_2 (11_{10}), ABCD = 1100_2 (12_{10}), ABCD = 1101_2 (13_{10}) and ABCD = 1111_2 (15_{10}).

Table 3.4.1a Using Table 3.2: First Four Options (0 – 3)

$Q_{n+1} = \bar{S}Q_n$ …(0)

SR	00	01	11	10
0	x	A,0	C,0	0
1	x	1	D,0	B,0

K-MAP

ABCD = 0000_2 (0_{10})

ABCD = 0001_2 (1_{10}) →

$Q_{n+1} = RQ_n$ ….(1)

SR	00	01	11	10
0	x	A,0	C,0	0
1	x	1	D,1	B,0

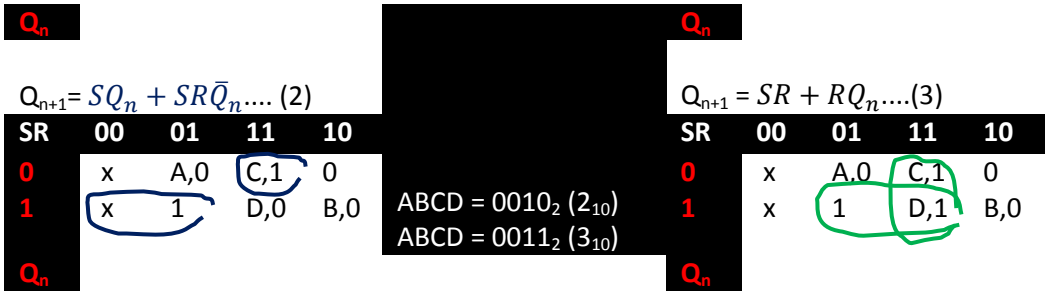


Table 3.4.1b Using Table 3.2: First Four Options (4 – 7)

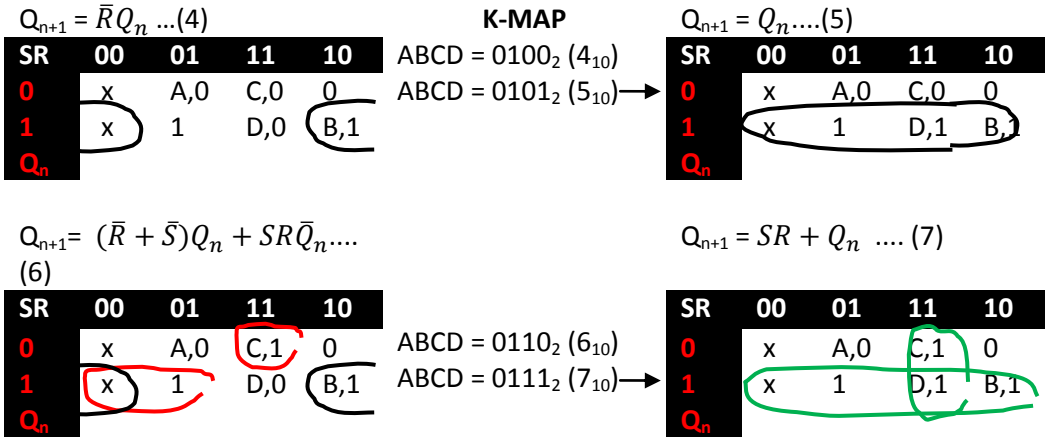
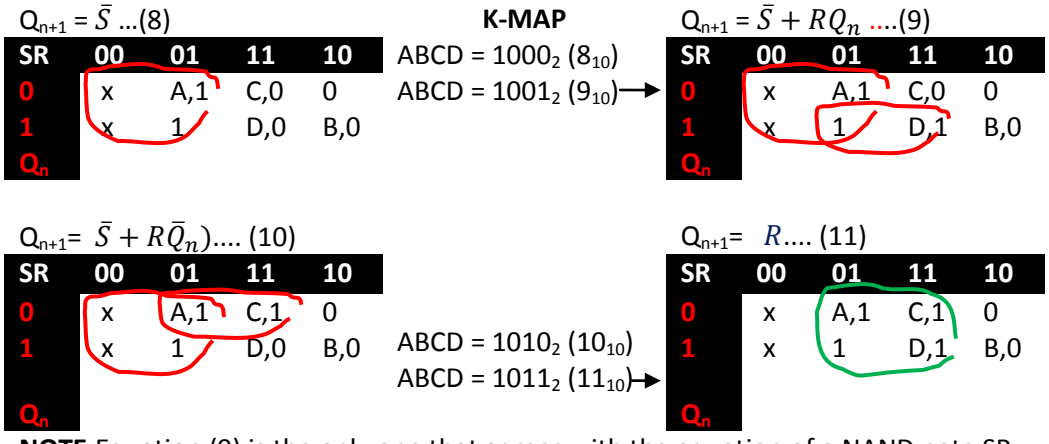


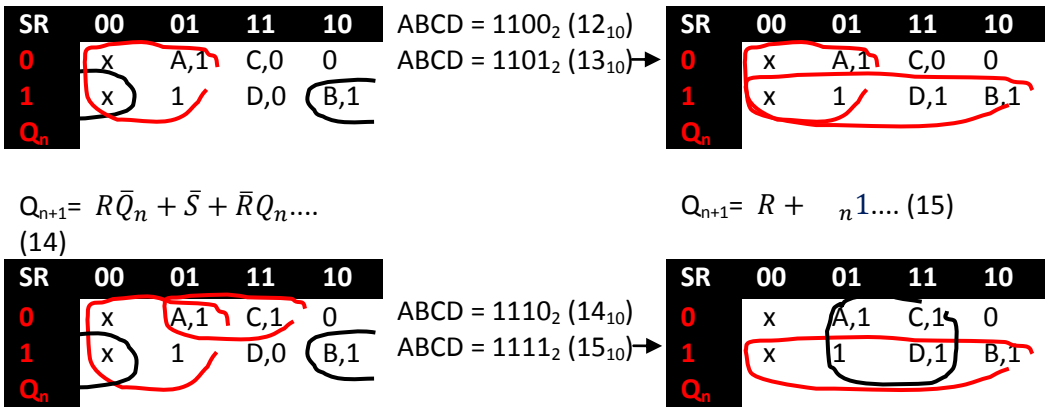
Table 3.4.1c Using Table 3.2: First Four Options (8 – 11)



NOTE: Equation (9) is the only one that agrees with the equation of a NAND-gate SR Flip Flop. Hence, the modification required is given by ABCD = 1001_2 (9_{10}).

Table 3.4.1d Using Table 3.2: First Four Options (12 – 15)

Qn+1 = $\bar{R}Q_n + \bar{S}$ ...(11) **K-MAP** Qn+1 = $\bar{S} + Q_n$ ....(12)



For Fig 3.2 to be modified to conform with the characteristics of a NAND-gate SR-Flip Flop presented in Table 3.1, the output of the two transitions/states (SRQ_n = 010, 101) must be complemented in addition to determining the appropriate values of the transitions/states A, B, C & D as it is done for Fig 3.1a. This is presented in Fig 3.4.2

Table 3.4.1a Using Table 3.3: First Four Options (0 – 3)

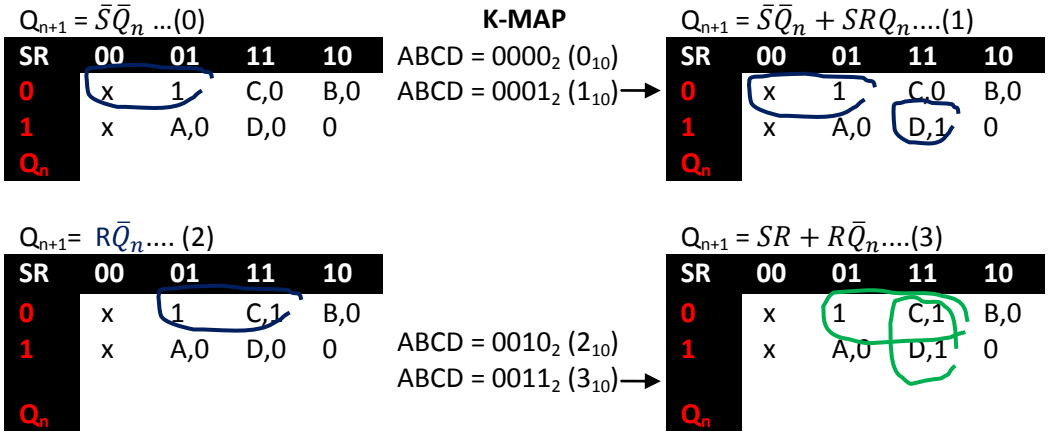


Table 3.4.2b Using Table 3.3: First Four Options (4 – 7)

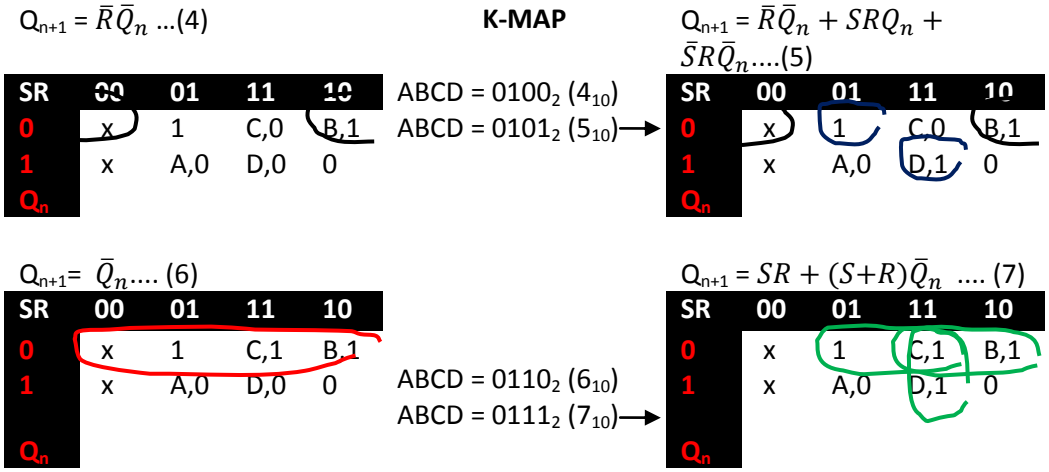
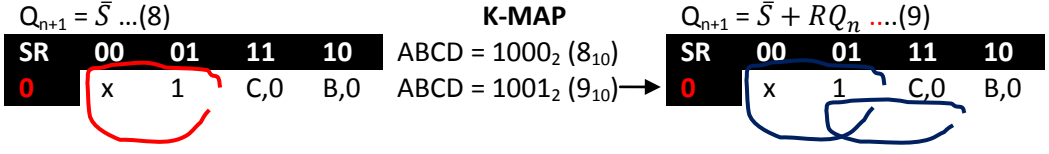


Table 3.4.2c Using Table 3.3: First Four Options (8 – 11)



1 / Q_n	x	A,1	D,0	0

1 / Q_n	x	A,1	D,1	0

$Q_{n+1}= \bar{S} + R\bar{Q}_n$.... (10)

SR	00	01	11	10
0	x	1	C,1	B,0
1	x	A,1	D,0	0
Q_n				

ABCD = 1010_2 (10_{10})
ABCD = 1011_2 (11_{10})

$Q_{n+1}= R$.... (11)

SR	00	01	11	10
0	x	1	C,1	B,0
1	x	A,1	D,1	0
Q_n				

NOTE: Equation (9) is the only one that agrees with the equation of a NAND-gate SR Flip Flop. Hence, the modification required is given by ABCD = 1001_2 (9_{10}).

Table 3.4.2d Using Table 3.3: First Four Options (12 – 15)

$Q_{n+1} = \bar{R}\bar{Q}_n + \bar{S}$ ...(12)

K-MAP

$Q_{n+1} = \bar{R}\bar{Q}_n + \bar{S} + RQ_n$ ....(13)

SR	00	01	11	10
0	x	1	C,0	B,1
1	x	A,1	D,0	0
Q_n				

ABCD = 1100_2 (12_{10})
ABCD = 1101_2 (13_{10})

SR	00	01	11	10
0	x	1	C,0	B,1
1	x	A,1	D,1	0
Q_n				

$Q_{n+1}= \bar{R} + \bar{Q}_n$.... (14)

SR	00	01	11	10
0	x	1	C,1	B,1
1	x	A,1	D,0	0
Q_n				

ABCD = 1110_2 (14_{10})
ABCD = 1111_2 (15_{10})

$Q_{n+1}= \bar{Q}_n + R$.... (15)

SR	00	01	11	10
0	x	1	C,1	B,1
1	x	A,1	D,1	0
Q_n				

Note that the required modifications of Fig 3.2 are too many. Hence, it is not engineering-wise to modify Fig 3.2 since the modification of Fig 3.1a gives a perfect design.

Alternative Solution

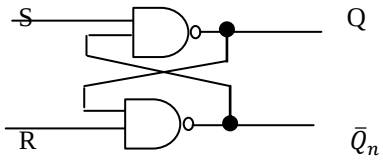


Fig 3.1a (NAND-Gate Config)

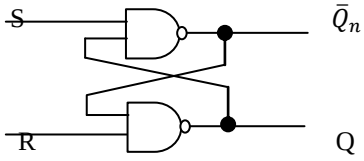


Fig 3.2 (NAND-Gate Config)

From Fig 3.1a,

$$Q_{n+1} = \overline{S\bar{Q}_n} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (a)$$
$$\bar{Q}_{n+1} = \overline{RQ_n} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (b)$$

If $Q_{n+1} = Q_n$, substituting equation (a) into equation (b) and vice-visa, we have

$$Q_{n+1}(1) = \overline{S\overline{RQ_n}} \equiv \langle 1\rangle \ldots\ldots\ldots\ldots\ldots\ldots\ldots (a1)$$
$$\bar{Q}_{n+1}(0) = \overline{R\overline{S\bar{Q}_n}} \equiv \langle 0\rangle \ldots\ldots\ldots\ldots\ldots\ldots\ldots (b1)$$

From SR-Flip Flop (-ve Logic configuration), SRQ_n = 00,001 $\equiv \langle x\rangle$ …… (c)

Plot equations (a), (b1) & (c) into a K-Map and Truth Table (Table 3.5) as follows:
Examples: For S = 0, R = 1, Q_n = 0, we have from equation (a1), $Q_{n+1} = \bar{S} = \bar{0} = 1$ and from equation (b1), $\bar{Q}_{n+1} = \bar{R} = \bar{1} = 0$. This procedure is continued for SRQ_n = 011, 100 & 101

Table 3.5: TRUTH TABLE for Fig 3.1a

S	R	Q_n	Q_{n+1}	$\bar{Q}_{n+1}$	REMARKS
0	0	0	x	x	Forbidden States

K-MAP

SR	00	01	11	10

0	0	1	x	x	
0	1	0	1	0	The output follows 'R'. That is, Q_{n+1} = R
0	1	1	1	0	The output follows 'R'. That is, Q_{n+1} = R
1	0	0	0	1	The output follows 'R'. That is, Q_{n+1} = R
1	0	1	0	1	The output follows 'R'. That is, Q_{n+1} = R
1	1	0	0	1	Resting States
1	1	1	1	0	

0	x	1	0	0
1	x	1	1	0
Q_n				

$Q_{n+1} = \bar{S} + RQ_n$

This is the same as equation (9) earlier obtained.

NOTES:

1. The Resting and Forbidden states satisfy the conditions required for a NAND-gate SR-FF
2. The Active states fully follow 'R' as required for a NAND-gate SR-Flip Flop.

Hence, this configuration (Fig 3.1a) fully satisfies the characteristics of a NAND-gate SR-Flip Flop because its Active states follow 'R' and its forbidden states are correctly assigned. This is only possible if it can be ascertained that, at all times, the signals arriving at terminals 'S' and 'R' are received at different times and that the propagation times (τ) of both NAND gates are the same. That is, $Q_n = Q_{n+1}$ and $\tau_S = \tau_R$.

From Fig 3.2,

$$Q_{n+1} = \overline{R\bar{Q}_n} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (a)$$
$$\bar{Q}_{n+1} = \overline{SQ_n} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (b)$$

If $Q_{n+1} = Q_n$, substituting equation (a) into equation (b) and vice-visa, we have

$$Q_{n+1}(1) = \overline{R\overline{SQ_n}} \equiv \langle 1 \rangle \ldots\ldots\ldots\ldots\ldots\ldots\ldots (a1)$$
$$\bar{Q}_{n+1}(0) = \overline{S\overline{RQ_n}} \equiv \langle 0 \rangle \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (b1)$$

From SR-Flip Flop (-ve Logic configuration), SRQ_n = 000, 001 ≡ ⟨x⟩ …… (c)

Plot equations (a1), (b1) & (c) into a K-Map and Truth Table (Table 3.6) as follows:
Examples: For S = 0, R = 1, Q_n = 0, we have from equation (a1), $Q_{n+1} = \bar{R} = \bar{1} = 0$ and from equation (b1), $\bar{Q}_{n+1} = \bar{0} = 1$. This procedure is continued for SRQ_n = 011, 100 & 101

Table 3.6: TRUTH TABLE for Fig 3.2

S	R	Q_n	Q_{n+1}	$\bar{Q}_{n+1}$	REMARKS
0	0	0	x	x	Forbidden States
0	0	1	x	x	
0	1	0	0	1	The output follows 'S'. That is, Q_{n+1} = S
0	1	1	0	1	The output follows 'S'. That is, Q_{n+1} = S
1	0	0	1	0	The output follows 'S'. That is, Q_{n+1} = S
1	0	1	1	0	The output follows 'S'. That is, Q_{n+1} = S
1	1	0	0	1	Resting States
1	1	1	1	0	

K-MAP

SR	00	01	11	10
0	x	0	0	1
1	x	0	1	1
Q_n				

$Q_{n+1} = \bar{R} + SQ_n$

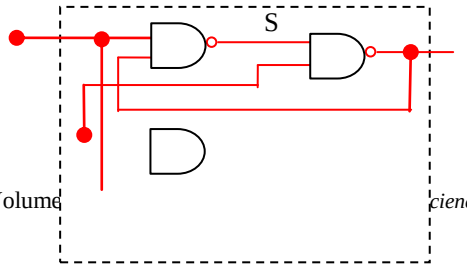
This is the same as equation (9) earlier obtained.

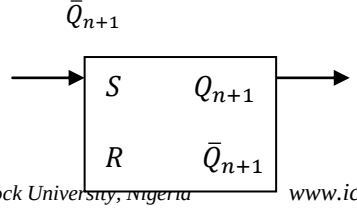
NOTES:

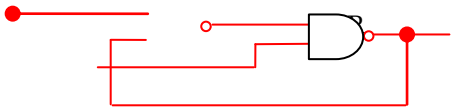
1. The Resting and Forbidden states satisfy the conditions required for a NAND-gate SR-FF
2. The Active states follow 'S' as if it were for a NOR-gate SR-Flip Flop.

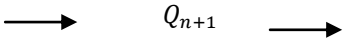
Hence, this configuration (Fig 3.2) fails to meet the requirements of a NAND-gate SR-Flip Flop.

The logical circuit that satisfies fully the characteristics presented in Table 3.1 is shown in Fig 3.1b. The details of the synthesis/design are presented somewhere else.









Flip Flop Symbol (NAND)

Fig 3.1b: Four-NAND-gate SR-Flip Flop Configuration

IV. CONCLUSION

From this analysis, it has been proved beyond all doubts that a two-gate configuration of an SR-Flip Flop is not sufficiently adequate to perform the full functions of an SR-Flip Flop. However, if it can be ascertained that, at all times, the signals arriving at terminals 'S' and 'R' are received at different times and that the propagation times (τ) of both gates are the same ($Q_n = Q_{n+1}$ and $\tau_S = \tau_R$), then the same result could be obtained as that of a 4-Gate network. But this is a condition that is practically impossible because, Q_n can never be at the same terminal with Q_{n+1} because of the inherent property of both signals.

VI REFERENCES

[1] O. J. Omotosho, (2012), "Fundamentals of Digital Systems", Franco-Ola publishers

[2] S. O. Ogunlere, (2011), "Digital Logic and Design", France-Ola publishers

[3] Ralph J. Smith; (1984), "Circuits, Devices and Systems", 4 th Edition

[4] W. H. Eccles and F. W. Jordan, (19 September 1919), "A trigger relay utilizing three- electrode thermionic vacuum tubes," The Electrician , vol. 83, page 298. Reprinted in: Radio Review , vol. 1, no. 3, pages 143– 146 (December 1919).

[5] Montgomery Phister (1958), "Logical Design of Digital Computers". http://books.google.com/books?

[6] K. G. Sharma, T. Sharma, B. P. Singh and M. Sharma in Dept of Electronics and Communication Engg. FET-MITS Lakshmangarh, Rajasthan, (INDIA).

[7] P. K. Meher, , " Extended Sequential Logic for Synchronous Circuit Optimization and its Applications" IEEE journal

[8] P. K. Meher, "Efficient implementation of accumulation in finite field over GF (2 m) and its applications", IEEE Transactions on Very Large Scale Integration Systems , accepted for future publication. [2] N. H. E. Weste and D. Harris, CMOS VLSI Design.

[9] B. P. Singh, N. Arora, K. Mehta, "Low Power Efficient D Flip Flop Circuit", International Symposium on Devices MEMS, Intelligent Systems & Communication (ISDMISC) 2011 Proceedings published by International Journal of Computer Applications® (IJCA)

Determination of DC Motor Load Balance in Control Loop Using Modeling and Simulation Techniques

Kadiri K. O.
Department of Electrical/Electronics Engineering
Federal Polytechnic Offa, Kwara State, Nigeria

Alabi O. A.
Computer Engineering Department
Federal Polytechnic Offa, Kwara State. Nigeria

Nicolas Goga
University of Groningen, Netherlands

Abstract- This paper describes a simpler way to control the motor load balance using modeling and simulation techniques. The performance of the system is simulated. The speed is regulated by PI controller. Simulink is utilized with MATLAB to get a reliable and flexible modeling schema. Current gain controller was calculated with some approximations in the conventional design process which is replaced by proposed model order reduction method, in order to highlight the effectiveness of the load balance method used. The technique proposed suppresses torque oscillations. This drive has high accuracy, robust operation from near zero to optimal level.

Keywords: Cascade control, Matlab Simulink, DC motor modeling, Controller

I. INTRODUCTION

The use of DC motor varies with concurrent technology today. DC motors used in many fields or various applications such as steel rolling mills, electric trains and robotic manipulators require current and speed controllers to perform tasks. Major problems in applying a conventional control and balance algorithm in a controller design are the effects of nonlinearity in a DC motor. The nonlinear characteristics such as friction and saturation could degrade the performance of conventional controllers of load mechanism. Many advanced model-based load control methods such as variable structure in load control and model reference adaptive have been developed to reduce these effects. However, the performance of these techniques depends on the accuracy of system models and parameters. In this paper, current controller of two quadrant DC motor drive is emphasized. The linear operation of DC motor drive was taken into account in the design stage of current load balance controller. In conventional design methods, some of simplification processes are considered to design the load
balance system with controller parameter values but whereas in this proposed method, the model order reduction technique was introduced for the controller parameter design values [1].

DC motors load factor cannot be exempted in developing of open and close loop system since it is regarded as the essential machine in the most control systems such as electronics systems, electrical systems in homes, industrial machines, vehicles, trains and process control. It is well represented in the mathematical model as a crucial aspect for a control system design [1], [2]. For a DC motor control system, there are many models and techniques to represent the machine behavior with a good accuracy on design prototype.

However, the parameters of the model are essential because the mathematical model cannot give correct behavior and characteristics without correct parameters in the model [3, 4]. In this study, the design of Generic Model controller and their effect to industrial application of DC motor load balance at steps included characteristic, structure and the mathematical model, and simulation of stability response for speed control of DC motor control system are considered. The design and simulation on visibility study was to demonstrate the basic theoretical to predict pratical results of the system used by Generic Model Control to achieve better response by less noise, stability, steadiness and less overshoot.

DC machines in load system are characterized by their versatility in diffrent dimension by means of various combinations of shunt, series and separately-excited field windings alignment.This can be designed to display a wide variety of volts amperes or speed torque characteristics for both dynamic and steady-state operation in a closed loop system. More so, because of the simplicity with which it can be control systems of DC machines, that are frequently used in many applications requiring a wide range of motor speeds and a precise output motor control to generate uniform output.

In this research paper, the DC motor modeling was synthesized and determinant of load effect depends on steady electrical and mechanical performances. The DC motor is driven by applied voltage respect to

the source. Figure 1 focus the equivalent circuit of DC motor with separate excitation. The characteristic equations of the DC motor are represented as [5],[6]

II. OVERVIEW OF RESEARCH DESIGN

The present invention solves the above problems by providing separately excited direct current motors with a simplified control system for maintaining a load balance between the drive motors. The field windings of the two motors are each divided into separate halves with the one-half of the field winding of one motor being coupled in series with one-half of the field winding of the other motor and the field of a separate exciter is coupled between the armature circuits of the two motors to sense any unbalance in the current flowing between the two armatures since any unbalance of the armature current flow is directly related to an unbalance in the torque load on the two motors. As a consequence, the resulting current flow in the field will vary the current flow in the armature of the separate exciter which is coupled between the common junction of the series-connected half-fields of the two drive motors so the current flow from the armature of the separate exciter through the drive motor fields will be in the direction to rebalance the current flow through the armature of the two drive motors.

The separate exciter functions to cumulatively compound the motors during motoring operation and differentially compound them during the braking operation with the possibility of serious accidents if the invention also includes means in the armature circuit of the separate exciter for reversing the current flow through the motor fields. Thus, when the direction of rotation of the traction motors is reversed, the direction of current flow through the separate exciter can also be reversed so that the separate exciter can balance the torque loads under both forward or reverse motoring and braking conditions. The reversal of the current flow through the armature of the separate exciter is conventionally accomplished by relay-operated switches or a switch that is operated in conjunction with the reversing controls on the vehicle [6].

Design view on load Imbalance measurements

A load balance is measured with current drawn by the DC motor considering the three phase induction motor, the balance load was achieved through the current that they draw on each phase should be about the same (less than 10%, as measured below). The imbalance cause can deteriorate the stator insulation as a result of voltage imbalance measured less 3% of the resultant output. In the design model, current and voltage calculation is accomplished by taking side of the drive. The internal capacitor bank value is set to determine balance load of the DC because the charges on the internal capacitor determine the workforce on the drives from current and voltage draw from the three phases [7]. The design can suffer from voltage sag or low bit current from any of the phases and the use of internal bank capacitor is employed to create drive balance load.

In the research, load balance was calculated using percentage imbalance system as follow;

Percent imbalance = maximum deviation from average / average of three phases X 100%

In the experiment, the following measurement values were obtained on the DC motor;

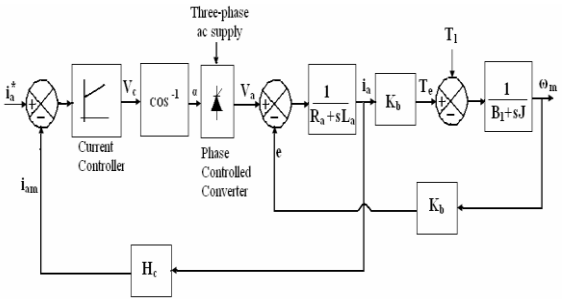
Taking measured: Phase A = 449 A; Phase B = 470 A; Phase C = 462 A respectively.

We determined the average as = (449 + 470 + 462) / 3 = 460 A

Max Deviation = 460 – 449 = 11 A

Imbalance = (11 / 460) x 100 % = 2.4 %

Figure 1: *DC motor and current-control loop*



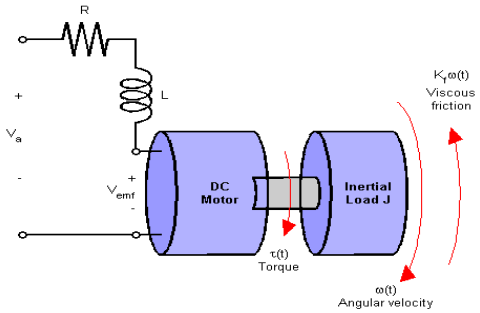
The Design of DC Motor Load profiling

The relationship between the torque and current was justified since the determinant of DC drives depend on electrical energy (current) into mechanical energy (torque) to generate electro-magnetic energy as shown in Fig 1. The load of DC motor depends on torque and current therefore voltage is in state and current is used to compensate DC motor load balance. The research evaluates the impact of a load on the motor drive system, by measuring the current it draws over some period of time. Since current draw typically varies over time as the load varies. Current imbalance was determined to ensure DC health status [8].

Motor shaft voltages and bearing currents.

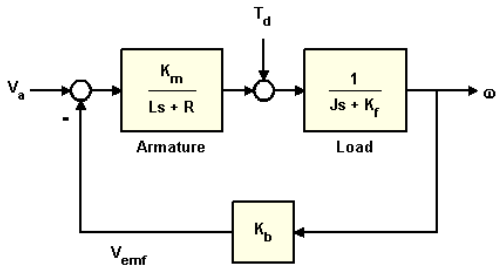
The current produce in DC motor is a typical sine wave (analogue Signal), this was used to measure voltage (between 1 to 2volt) rotor-to-frame (ground), using stranded wire or a carbon brush. The Pause Wave Modulation (PWM) waveform can cause breakdown voltages of 8 to 15V to occur between the shaft (most specifically, the bearing) and the frame. This damages bearings, causing pitting and scarring. Therefore, many solutions have been proposed, but the most recommended one is the shaft grounding of a device.

Fig 2: The armature current controller DC motor



In armature-controlled DC motor in real time, it is the applied voltage (emf) that controls the angular velocity (w) of the shaft. This paper shows two major techniques for minimising the sensitivity of angular velocity (w) to load variations changes in the torque opposed by the motor load in the feedback system [8],[9].

Fig 3: Reducing Techniques with load sensitivity



The figure (3) can be reduced to second order, to facilitate a simple controller synthesis, by judiciously selecting a close system above, a simplified model of the DC motor is represented. The torque Td models load disturbances. The strategy is to minimize the speed variations induced by such disturbances in the system.

Taking physical constants as: R = 5.0 Ohms L = 0.15Henrys and Km = 0.1; Kb = 0.1 refers to torque and back e.m.f constants

Kf = 0.2; % Nms

J = 0.02; % kg.m^2/s^2

In the first stage design and implementation of a state-space model of the DC motor with two inputs (V_a, T_d) and one output (w):h_1 = tf(Km,[L R]) for armature function

The precisely state-space formed is to minimize the model order is computed first. The plot of the angular velocity response is to a step change in voltage, (Va): stepplot(dcm(1))

A motor speed control system has the actual speed amplitude approaches to the desired speed r ,

Show that Open-loop: Gc(s) = k, G(s) = k1/(Ts+1)

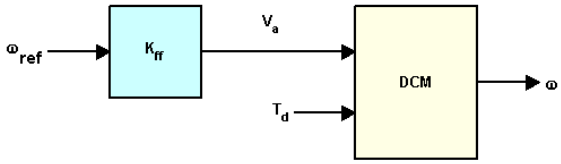
Assume, K_1 = 1, T = 12

- and closed-loop with a proportional control term k, H(s) = 1

taking, K_1 = 10, T = 12.

The characteristic of frequency response on closed system time domain was configured using indication feed foward control signal[5],[9].

Fig 4: Feed forward Control Design



Feedforward Control

The feed Feedforward is an open loop system with Voltage control Va. Taking DC motor gain Kff is defined and set to the reciprocal of the DCM gain from Va to w.Kff = 1/ dcgain (dcm (1)) Where;

Kff = 10.1000

Using eveluation method, the face of load disturbances is evaluated using feedforwards design and the simulation result shown the response command w ref=1 with a disturbance Td = -0.1Nm between t=5 and t=10 seconds:

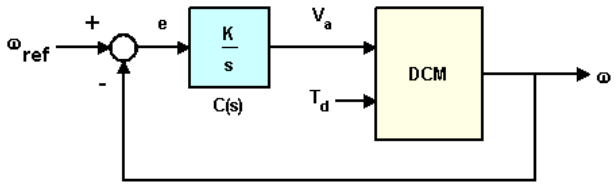
Taking Time {t} = 0:0.1:15;

Td = -0.1 * (t>5 & t<10); % load disturbance

Observation: Once discovered, the amplitude feed forward control handles load disturbances poorly in the closed loop system.

The Next layout shows the feedback control structure as represented in Fig 5 below.

Fig 5: Feedback Control Structure



Feedback Control

According to the findings, the enforced zero steady-state error, uses integral control of the form C(s) = K/s where K is to be determined. To determine the gain K, The experiment in the simulation show the use root locus technique applied to the open-loop 1/s * transfer(Va->w):h = rlocusplot(tf(1,[1 0]) * dcm(1));

Root Locus Design

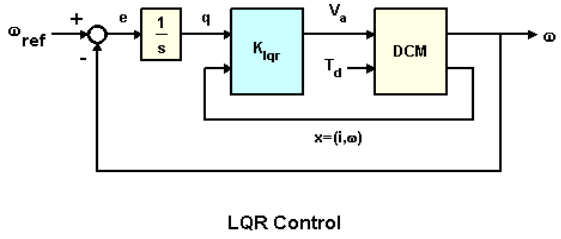
The main idea of root locus design is to obtain the closed-loop response from the open-loop root locus plot for the DC motor response. By adding zeroes and poles to the original system, the root locus is modifiable using a new closed-loop response. Foremost the root-locus for the system itself imposed with a unit circle. Using m-file editor , the following

commands can be executed and rerun. These give root-locus plot as shown above[6],[10].
The syntax is as follow;

rlocus(numd,dend)
title('Root Locus of Original System')
zgrid(0,0)

or
input the follow to derive DC motor controlcommand line
*h = rlocusplot(tf(1,[1 0]) * dcm(1));*
setoptions(h,'FreqUnits','rad/s');
set(gca,'Xlim',[-15 5],'Ylim',[-15 15]);
The results were equally obtained on the curves GUI which was read to determine the gain values and related DC motor focus and control measures. The precuite choice is K = 5. Note that the GUI Design with Guards Tool offers an integrated GUI to perform such designs in the M. File editor. By justifying and correlating new design with the initial feed forward design on varrying load and determine load balance and gain control.

Fig 6 Linear Quadratic Regulator (LQR) Control System



LQR Control

Using the LQR control system in Fig 6 above and in addition to the integral of error denoted, the LQR scheme also uses the state vector x=(i,w) to synthesize the driving voltage mechanism.

Va. The resulting voltage is of the form

Va = K1 * w + K2 * w/s + K3 * i where i is the armature current. By derivation for better disturbance rejection. The cost function is used to penalises the large integral error such as cost function where the optimal LQR gain for actual cost function is computed as follows: dc_aug = [1 ; tf(1,[1 0])] * dcm(1); % add output w/s to DC motor model
K_lqr = lqry (dc_aug,[1 0;0 20],0.01);
A simulation is done on closed loop model in purpose to denote
P = augstate(dcm); % inputs:Va,Td outputs:w,x
One technique for compensator design is to work with Bode diagrams of the open-loop response (loop shaping) [10],[11]. Using Bode diagrams, extrapolating design to gain and phase margin specifications, in view of making adjustment of the bandwidth and add notch filters for disturbance rejection in the system [12].

A case of DC motor system show how to create a compensator using Bode diagram design techniques. the transfer function of the DC motor is can be represented as follow;

Transfer function:
2.5
---
s^3 + 16 s + 50.01

For this design, the design criteria is categorized as follow:

I. Rise time of less than 0.1 second
II. Steady-state error of less than 7%
III. Overshoot of less than 15%
IV. Gain margin greater than 40 dB
V. Phase margin greater than 50 degree

III. DESIGN PROTOTYPE

The DC motor parameters and ratings of a speed controlled DC motor drive maintaining the field flux constant are as follows 220V, 8.3A, 1470 rpm, Ra=4Ω, J=0.0607kg-m2, La=0.072H, Bt=0.0869N-m/rad/sec, Kb=1.26V/rad/sec. The converter is supplied from 220V, 3-PhaseAC at 50Hz. The converter is linear, and its maximum control input voltage is ±12V. The tacho-generator has the transfer function

$$G_w(s) = \frac{0.065}{V_{cm}} = \frac{1.35 \times 220}{12} = 24.75$$

V_{dc} (max) = 247.5V
The rated DC voltage required is 220V, which corresponds to a control voltage of 7.09V. The transfer function of the converter is,

$$Gr(s) = \frac{24.75}{(1 + 0.00247s)}$$

Current transducer gain: The maximum safe control voltage is 7.09V, this is related to the maximum current error [11]:
Imax = 20A

$$Hc = \frac{7.09}{I_{max}} = 0.355V/A$$

Motor transfer function;

$$K_1 = \frac{B_1}{K^2_b + R_a B_t} = 0.0449$$

T_1 = 0.1077 sec and T2 = 0.0208sec

$$Tm = \frac{J}{B_t} = 0.7sec$$

The subsystem of the design transfer functions are

$$\frac{Ia(S)}{Va(s)} = \frac{K_1 (1 + sTm)}{(1 + sT_1)(1+sT_2)} = \frac{0.0449 (1 + 0.7s)}{(1+0.0208s)(1+0.1077s)}$$

$$\frac{\omega_m(s)}{Ia(s)} = \frac{(1+sT_m)}{(1+sT_m)(1+sT_2)} = \frac{0.0499\ (1+0.7s)}{(1+0.0208s)(1+0.1077s)}$$

Design of current controller

$T_2 = 0.0208$sec; Tc = 0.03sec

By applying the proposed model order reduction technique, with the value of c = 0.707, the value of K is achieved as,

K = 357.192

IV. DISCUSSION

The genesis of this research was to determine load factor on DC and the contemporary of the discovery focus on drive, load and motor itself. Basic motor measurement could target imbalance to determine motor health status. Then other simple drive measurement was done, it helps to proceed systematically and check for causes of over or under voltage trips. DC profile management on load to find the cause for intermittent overload trips. Finally, in this finding, premature motor bearing or winding failures, as well as control signal interference can work well when a Fluke Scope Meter test tool is used to make additional motor measurements. We evaluate using modeling and simulation techniques to arrived on most significant signal impact for all the possible sources of problems, but the result has certainly identified a lot of the most common ones feedback (closed-loop) control that can be used to stabilise systems model, speed up the transient response of the DC motor mode, enhance the steady-state characteristics, provide and source for disturbance rejection state, and decrease the sensitivity to parameter variations using plot.

Analysis of DC based simulation motor without load (TL = 0 Nm) depends on the state-space model of DC motor that the load torque is taken as TL = 0.0 (without load), and the set point value of the motor speed is set at ωset = 350 rad/sec. LQR controller calculates the target value of armature current which in turn secures by adjusting the armature voltage and change resistor and Inductor for gain control. The dynamic responses of armature voltage, armature current, motor speed and motor torque with proposed control system are shown in Fig 2 and Fig 3 respectively.

Analysis based of simulation of DC motor with load (Td = -0.1 Nm): In the state-space model of DC motor, the load torque is taken as TL = -0.1 Nm in the same time the set point value of motor speed is setting on disturbance limit between 5-10sec. The dynamic responses of armature current, armature voltage, motor speed and torque response with control system as shown in Fig. 7.

Fig 7:DC Motor Frequncy respose to close loop

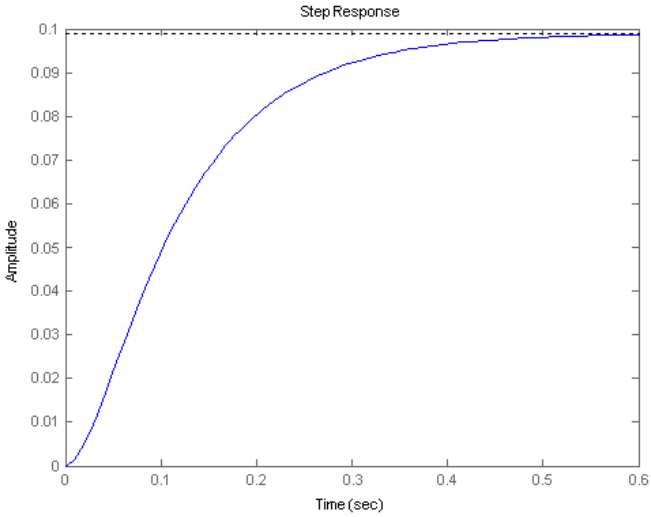
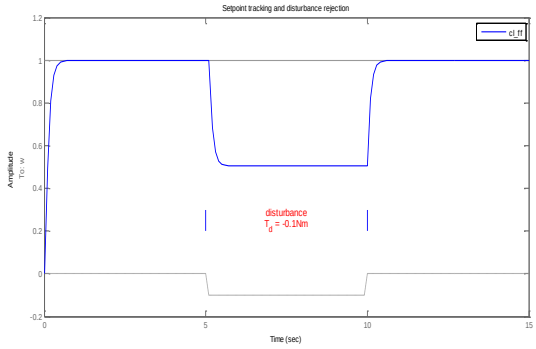


Fig 8: DC Motor Stepping and Disturbance Rejection



In figure 8 above, the original constraint setup and the derivation of the above equations indicate proportionality distance in amplitude steady fall and sharp raises as a result of load imbalances.

Fig 9: *Root Locus* Design Method for the *DC Motor system*

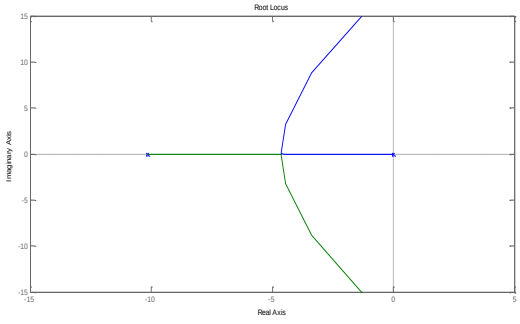
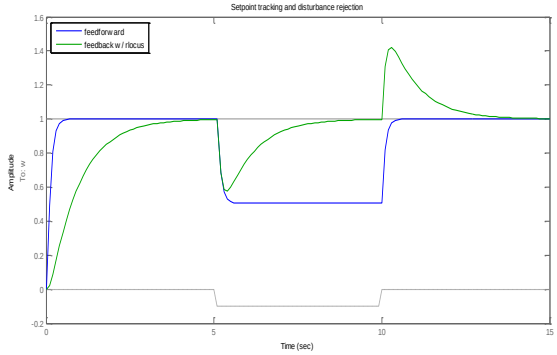
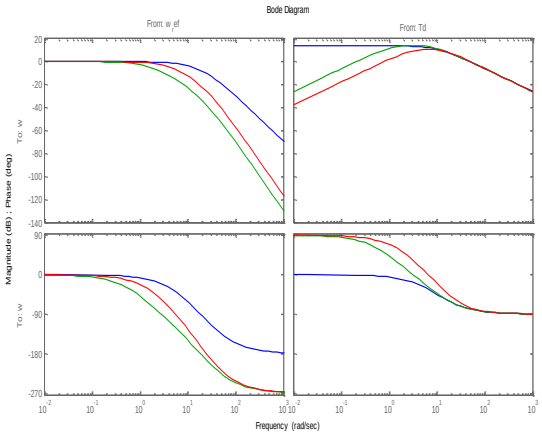


Fig 10: Feed back Fucos on the effect of load tracking distance and rejection



The result in fig 10 showing tracking balance as a result of load effect. The system balance and focus on load tracking. The plotted is to compare the closed-loop Bode Finally three designs simulation could be compared following case test simulation Bode plot

Fig 11: Comparison of Designs using multi plot



V. CONCLUSION

The simulation results show the load balance using root locus and Bode plot to evaluate transfer functions of the DC motor. High performance was achieved by Simulation results indicate that the determinant controller has a better performance in transient, load balance, error rate, feedback control and steady state response.

The design of the gain constant of current controller varying with different load is of paramount importance in meeting the dynamic specifications of the motor drive efficiency. This paper helps to target load control effect in the same model using order reduction technique. The validity of the proposed method was illustrated through the design example. The proposed model order reduction method is a simple mathematical expressing and producing the stable reduced order system if the given system is stable load. The optimal value of current controller gain to suppress the oscillations at the output was reduced by trial and error technique but it may be selected optimally by applying the neural network concepts or genetic algorithm system.

References

[1] Santana, J., J.L. Naredo, F. Sandoval, I. Grout and O.J. Argueta, 2002. Simulation and Construction of Speed Control for a DC Series Motor. Mechatronics, 12: 1145-1156.

[2] Abdulrahman, A.A. Emhemed and Rosbi Bin Mamat, 2012. Modelling and Simulation for Industrial DC Motor Using Intelligent Control. Procedia Engineering, 41: 420-425.

[3] M.F. Hutton and B. Friedland, "Routh approximations for reducing order linear time invariant systems", IEEE Trans. on Auto. Control, Vol.20, pp. 329-337, 1975.

[4] R.K. Appiah, "Linear model reduction using Hurwitz polynomial approximation", Int. J. Control, Vol.28, No.3, pp. 477-488, 1978.

[5] Y. Shamash, "Truncation method of reduction: a viable alternative", Electronics Letters, Vol.17, pp. 97-99, 1981.

[6] P.O. Gutman, C.F. Mannerfelt and P. Molander, "Contributions to the model reduction problem", IEEE Trans. on Auto. Control, Vol.27,
pp.454-455, 1982.

[7] Saab, S.S. and R.A. Kaed-Bey, 2001. Parameter Identification of a DC Motor: An Experimental Approach. IEEE International Conf. on Elec. Circuit and Systems. (ICECS), 4: 981-984.

[8] Ali Bekir Yildiz, 2012. Electrical equivalent circuit based modeling and analysis of direct current motors. Electrical Power and Energy Systems, 43: 1043-1047.

[9]
*www.researchgate.net/...**disturbancerejection**/.../50463520 c09c984383.p...*
by Z Gao - Cited by 106 - Related articles

[10] http://instruct.uwo.ca/engin-sc/391b/CTM/examples/motor2/digital.html

[11] http://www.mathworks.com/help/control/examples/dc-motor-control.html

[12] IJCSIS) International Journal of Computer Science and Information Security,
Vol. 7, No. 1, 2010 Design of Current Controller for Two Quadrant DC Motor Drive by Using Model Order Reduction Technique K.Ramesh *Et-al*

The Cooperative Prefix Design Approach

A Worked Sample-BUHR Mobile Solution

Okoro U.R and Kuyoro S.O.
Computer Science Department
Babcock University
okororaymond@ymail.com

Makinde A.
Human Resource Unit
Babcock University
makinde_abiola@yahoo.com

Otuza C. and Egwuonwu T.K.
Babcock Business School
Babcock University
tkegwuonwu@yahoo.co.uk

***Abstract*— The field of Design Science is experiencing changes from a user-centered design process in which construct references hovers just around the two giant design terms-functionality and usability to that of participatory design approach where user's tacit knowledge, pre and post experiences becomes full part of the design process. The collaborative fashion used to develop the Babcock University Human Resource (BUHR) mobile solution shrouds the "for" and "with" the cooperative and end-user mindset, giving a new dimension and approach to design process and user-device interaction experience.**

***Keywords*—User-Center Design, Participatory Design, Cooperative Design, Artefact**

I. INTRODUCTION

Human Computer Interaction and Design (HCI/HCID) addresses two basic terms in machine designs: functionality and usability. HCI emphasizes that all devices should be consistent, with a repertoire of gestures that feel natural and communicate their states and capabilities to the user while minimizing all cognitive friction. HCID aims at keeping users and their skills, needs and overall profiles central when developing innovative systems. It also creates awareness and appreciation of the crucial implications of good interfaces in terms of overall system performance and user satisfaction. [1][2] presented that every design should produce an agreeable meet between the user, the machine and the required services in order to achieve a certain performance both in quality and optimality. People are cynical about the methods and goals of consumerism. They are beginning to use their influence to get what they want, when they want it and how they want it.

This work presents a community approach towards building a product that satisfies individual and cooperate needs. A mobile solution that focused on the cooperative and end-user device conflict resolution is developed using Babcock University Human Resource (BUHR) as case study. The development tools are C#, Java for Android and IOS-SDK platform.

The remaining part of this work is arranged as follows Section 2 summarizes the global smart phones users demographics, Section 3 presents the existing design approach in HCI, Section 4 proposes a hybrid of user-centered design and participatory design as the new approach in this work, Section 5 presents a mobile solution that showcase the outcome of the proposed approach and Section 6 gives conclusion and recommendations.

II. GLOBAL SMART PHONES USERS

A. Smart Phone Defined

Not all mobile phones are smart phones. A smart phone is simply a Personal Digital Assistant (PDA) integrated with mobile phone features. Any mobile phone that can makes calls, have internet access, a touchscreen interface and still incorporates the functions of a computer and an operating system which can allow users to install and run apps is a smart phone. In 2012, out of 5 billion mobile phones users in the world, 1.08 billion are smartphone See Figure 1 [3]. In 2014 it is expected that the number of mobile users will total 4.55 billion 1.75billion will be smart phones users.

B. Nigerian Factor

Nigeria is the second biggest smartphone user in Africa trailing South Africa. In a report [4][6][7], it was presented that only 25% of Nigerians use smart phones as against the 105 million mobile telephone subscribers in Nigeria (that's 59% of Africa's mobile phone). 59% of Nigerian phone users were casual users while 16% were sophisticated users. According to the report released by [5][6][7] over ten (10) million smart phones worth $1 billion (N167bn) dollars were sold in Nigeria in 2013. Within two years a leap from 4 million smart phone users to 10 million users suggest that at least 30 million smart phones will be sold by 2015 in Nigeria. See Figures 2 and 3.

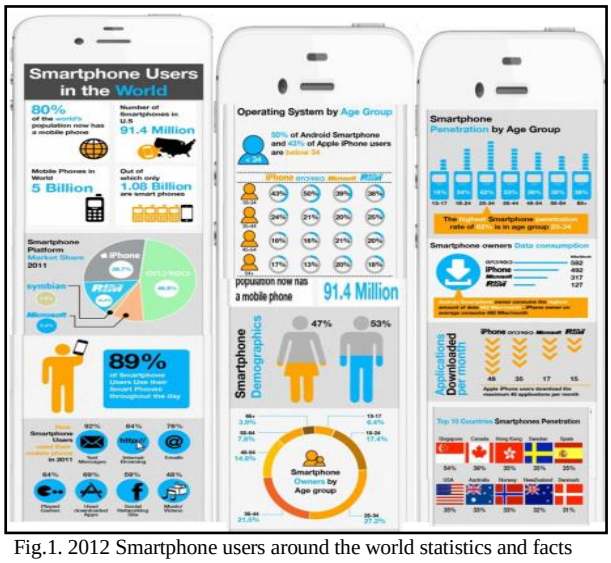


Fig.1. 2012 Smartphone users around the world statistics and facts
Source: [3]

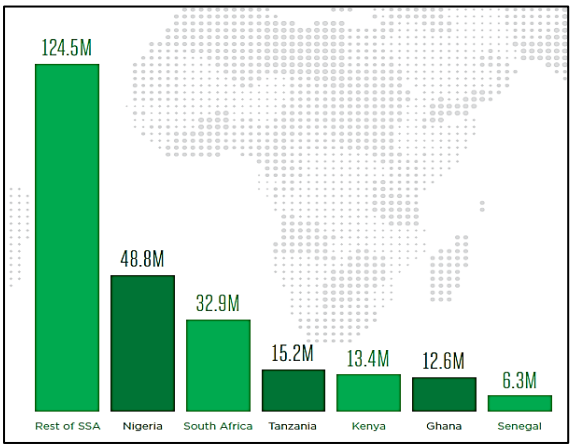


Fig. 2. Unique Subscribers in millions Q2 2013 [7].
Source: GSMA Intelligence

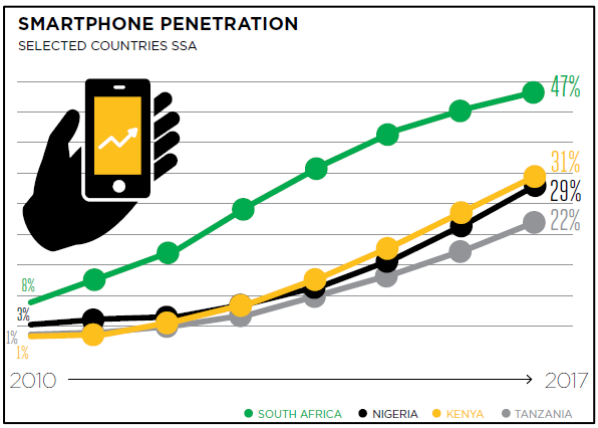


Fig. 3. Smart Phone Penetration Selected Countries [7].
Source: Strategy Analytics

"The subscriber penetration in Nigeria is currently under 30%, but with the figure in South Africa already at 65%. In absolute terms, the Nigerian market is expected to add 27 million new subscribers over the next five years, compared to around 7 million in South Africa."[7]

III. EXISITING DESIGN APPROACH

A. *User-Centered Design*

In User Centered Design (UCD) the artefact being constructed is the focus. All references hovers around the two giant HCI terms of functionality and usability. There is usually a researcher or social scientist to interface between the users and designer. The distinct but independent roles of the researcher and the designer do not make the user an important part of the team. The researcher collects primary data or uses secondary sources to learn about the needs of the user. The researcher interprets this information, often in the form of design criteria. The designer interprets these criteria, typically through concept sketches or scenarios. [8]

B. *Participatory Design*

Participatory design (PD) is seen as a new applied research in HCI design principles characterized by user involvement. [9]. Participatory Design as a branch of HCI which focuses on preserving tacit knowledge; unfortunately many designers ignore the aspect of user's tacit knowledge and hence construct artefacts based on the theory of cognition. PD was approached by [10] as a new dimension to computer systems design in which the people destined to use the system play a critical role in designing it, following some underlying HCI principle, practices and standards. User's tacit knowledge and experiences becomes a full part of the design process.

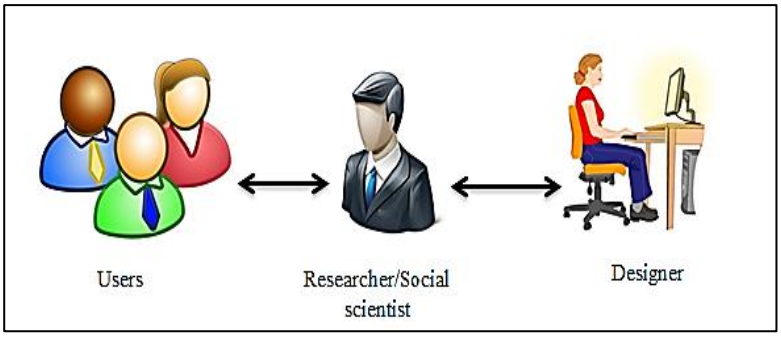


Fig. 4. An Independent Lateral Relationship between teams in UCD [9][10]

IV. HYBRID OF USER-CENTERED DESIGN AND PARTICIPATORY DESIGN APPROACH

A. *Cooperative Prefix Design (CPD) Approach- An Extension of UCD and PD*

A prefix design approach uses the strength of both UCD and PD in designing by:

1. Giving users a voice in the design process, thus increasing the probability of a usable design
2. Enabling technical and non-technical participants to participate equally
3. Providing an opportunity for developers to meet, work with and understand their users
4. Providing a forum for identifying issues
5. Providing an opportunity to get or enhance user buy-in
6. Using techniques which can be easily learned and applied in future activities.

In a Cooperative Prefix design approach more experts in different fields are assembled to work towards a superordinate goal via cooperate conflict resolution strategies [11].

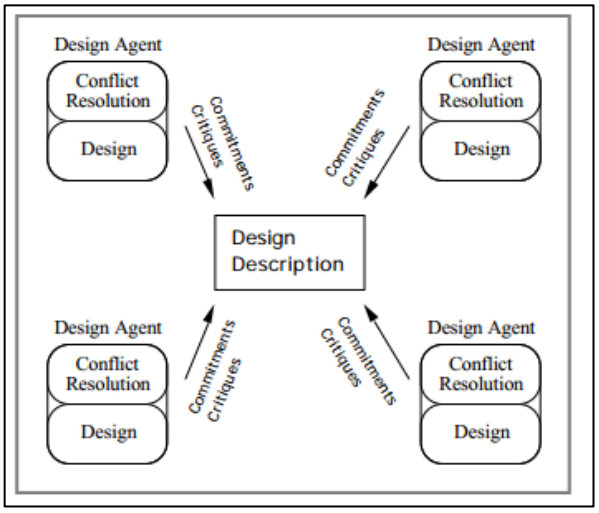


Fig. 5 Components Cooperative Design System, [11]

With the help of a social scientist/ contextual researcher, a distributed community design interest is fully "built-in". All experts' resolution is now being harnessed with the knowledge of both *naive* and *parametric* device users before the artefact is instantiated.

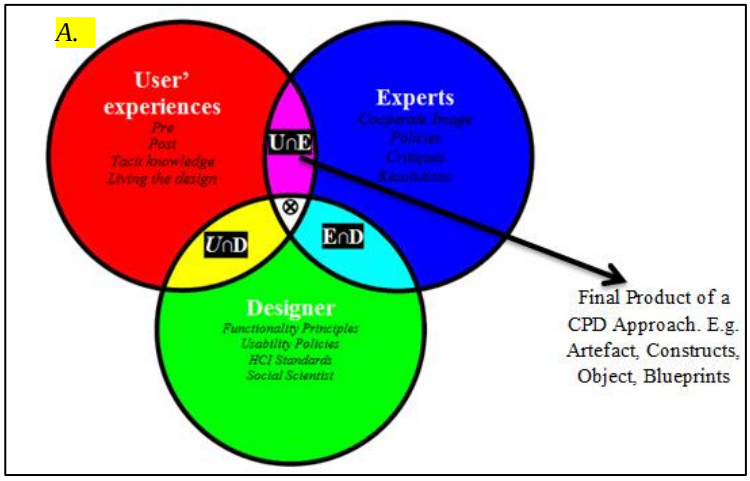


Fig. 6 Cooperative Prefix Design (CPD) Approach Model

B. *Cooperative Prefix Design Approach Explained*

From the model in figure 6, **U ∩ D** represents the fit between the User's device knowledge, their '*dream-about'* and the voiced ideas of the researcher or designer. **U∩ E** represents the participatory resolutions made between the cooperate body, cooperatives experts and the end users. In a cooperate environment such as Babcock University conflicts are usually resolved at a BU Workers forum where Administration gives each Staff and Faculty a direct participation opportunity in the design and analysis of any project. **D∩ E** represents the resolutions made between the social scientists and or Artefact designer, Cooperate image, initiative and interest. Later in this paper, the BUHR mobile app is a sample developed to further expose the philosophy of extended participatory design approach by adopted in Babcock University.

Now, the fit between the resolutions as given by **U∩ D, U∩ E and D∩ E** is the area **⊗ = U∩ E∩ D.** At this point it is important to know that the user's (cooperate and end-users) contextual experience forms two-third (2/3) of the components of the object description and instantiation.

V. Babcock University Human Resource Mobile Application Using the Template

A. *A Worked Sample*

Babcock University Human Resource Department outside upholding Babcock's cooperative interest has a responsibility of providing and sustaining aptitudes of Staff and Faculty in their various units and work terrain. The Information Technology benchmark for a 21century worker in BU is to make Babcock's business part of each worker's personal life and that we will call it 'carry-me-along'. Using the phenomenal growth of mobile internet over the fixed internet, most BU workers own and make use of various impressive mobile devices such as Iphone, Ipads, Android Devices and other smart phones; BUHR differentiates itself among competitors by capitalizing on sustainable technological landmarks. An engaging mobile app was developed to extend the features and functionality of the former BUHR website used by Staff and Faculty. A cooperative prefix design approach model has given the Administration (Cooperate/Canned Users), Staff and Faculty (end users) an equal opportunity to participate in the design of this mobile app. From the prefix model in Figure 6, we have divided the App users into the following categories:

B. *Casual Users*

50.2% members of staff are under this category. These users indicated interest in accessing the categorized information in the app from time to time. (see Figure 7)

C. *Sophisticated Users*

49.8% Faculty members of Babcock University are sophisticated users. These users are very specific and familiar with customized apps and so are more advanced users that the casual users. (See figure 7)

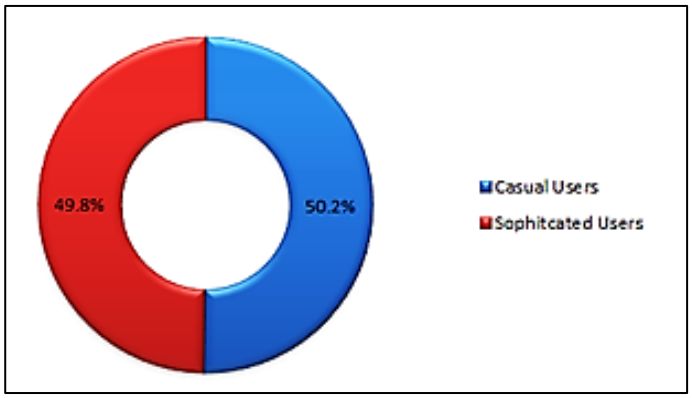


Fig. 7 User Categories in Babcock (2014)

D. Features of the Application

B. The needs and queries from users to Babcock University

C. Human Resource (BUHR) unit formed the mechanics of menu

D. categories and construct approach of the BUHR app. The

E. categories include: About US, Other info, Transfer

F. Promotion, Dates and Events, Vacancies, Duties, Dress Code,

G. Internet and Phone Use, Leaves, Contact Us, Suggestion Box,

H. Map, and Phone Book.

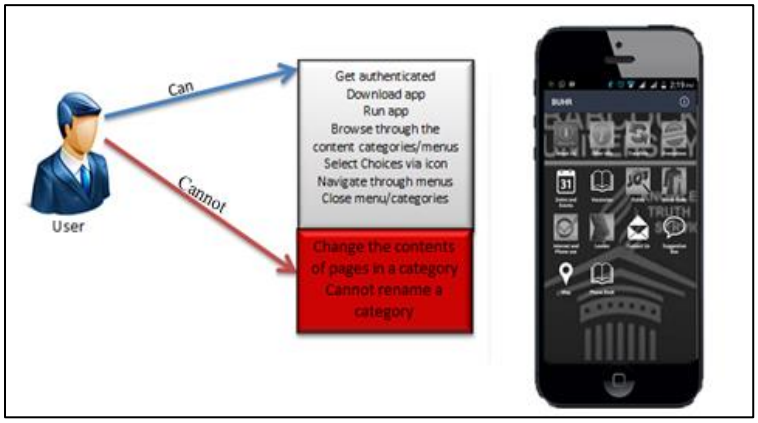


Figure 8: User case diagram and a CPD worked example for menu (2014)

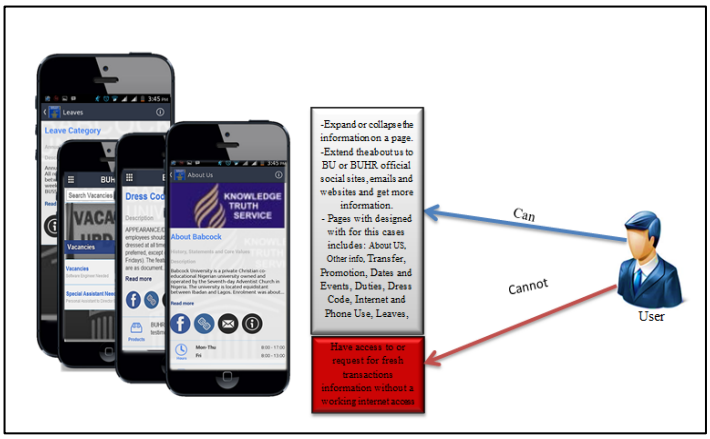


Figure 9: User case diagram and a CPD worked example for pages (2014)

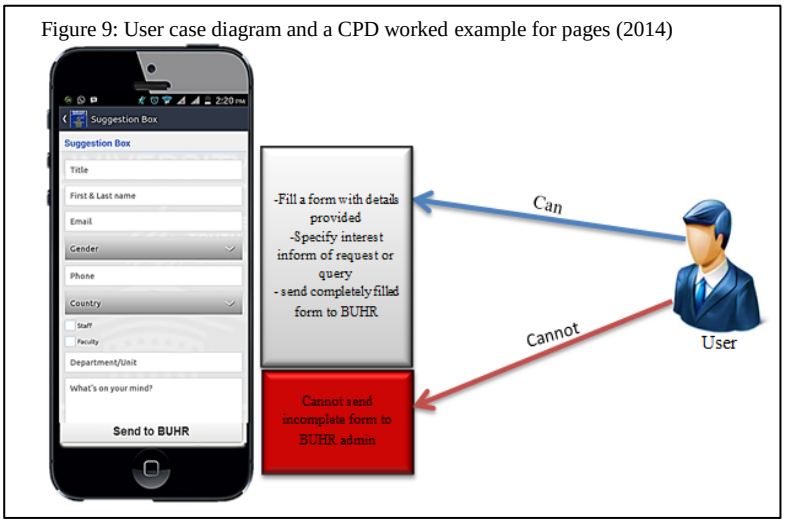


Figure 10: User case diagram and a CPD worked example for BUHR Form (2014)

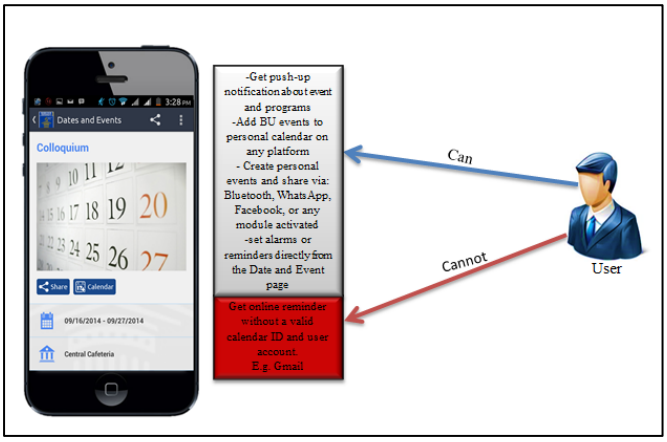


Figure 11: User case diagram CPD worked example for BUHR Dates and Events (2014)

I.

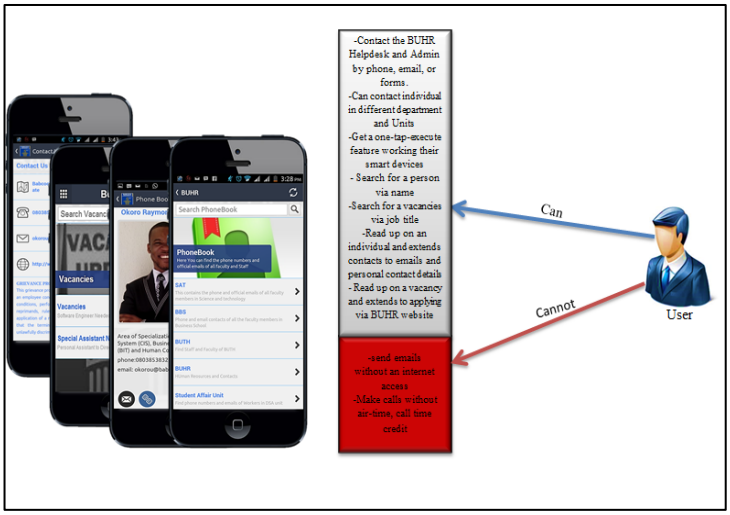


Figure 12: User case diagram and a CPD worked example for Contact, Phone and vacancy Pages (2014)

VI. CONCLUSION AND RECOMMENDATION

The world of Human Machine Interaction has evolved passed beyond placing our attention just functionality and usability. A Shift of User-Centered Design and Participatory design advocates that user post and pre experience represent full part of the design inception. A social scientist, or design experts can intermediate between Users, Cooperate Users/ Administration and the Designers resolving conflicts as per concepts and needs of the local community. For instantiation purposes, the CPD Model applied to design this worked sample (BUHR Mobile app) can be extended to an entire organization workflow and process design projects.

ACKNOWLEDGMENT

Special thanks to the management of Babcock University Human Resources Unit for arranging conflict resolutions sessions used in the CPD Approach for this example. We also appreciate Prof. Idowu Sunday for his valuable insight on the mobile app design parameters.

REFERENCES

[1] D. Te'eni, "Designs that fit: an overview of fit conceptualization in HCI", in P. Zhang and D. Galletta (eds), *Human-Computer Interaction and Management Information Systems: Foundations*, M.E. Sharpe, Armonk. 2006

[2] D. Te'eni, J. Carey and P. Zhang, "Human Computer Interaction: Developing Effective Organizational Information Systems", John Wiley & Sons, Hoboken, 2007

[3] Smartphone users around the world statistics and facts 2012-2014. Available at www.go-gulf.com. Retrieved 14/10/2014

[4] K. Macllwainem, "A report on Nigerian Mobile Subscribers" TNS Africa and Middle East Team Report 2014". Available at www.punchng.com/business/technology25-0f-nigerian-mobile-subscribers-use-smartphones-tns, 15/10/2014

[5] Moverick 2014STATS: IS NIGERIA BECOMING A SMARTPHONE MARKET?" available at mobility.ng/2014-is-nigeria-becoming-a-smartphone-market/. June 4, 2014

[6] Reports on Smart Phone Users (2014) available at http:www.mobimatter.com/comscore-reports November-2011-u-s-mobile-subdciber-market-share/ retrieved 14/10/2014

[7] GSMA, Sub-Saharan Africa Mobile Economy 2013. pp 18,25 . 2013 Report available at www.gsmaintelligence.com,info@gsmaintelligence.com. Retrieved 15/10/2014

[8] B. Elizabeth and N. Sanders, "From User-Centered to Participatory Design Approaches SonicRim. In Design and the Social Sciences". J.Frascara (Ed.), Taylor & Francis Books Limited. 2012

[9] R. R. Johnson, "User-centered technology: A rhetorical theory for computers and other mundane artifacts". New York, NY: SUNY Press. 1998.

[10] D. Schuler D and N Aki, "Participatory Design, Principles and Practices"; library of congress cataloging in publication data. ISBN 0-8058-0951-1, ISBN 0-8058-0952-x. Lawrence Erlbaum Associates, Inc, 365 Broadaway Hillsade, New Jersey 07642,1993

[11] K. Mark, "Supporting Conflict Resolution in Cooperative Design Systems . IEEE Transactions on Systems, Man and Cybernetics. Special Issue on Distributed Artificial Intelligence. Volume 21, Number 6, December 1991. (Advanced Research Lab Hitachi Ltd Hatoyama, Saitama 350-03, Japan

Challenges of Computing K-Means Convergence Criteria in MapReduce Framework

Osofisan Adenike
Department of Computer Science
University of Ibadan
Ibadan, Nigeria

Akande Oyebola
Department of Computer Science
University of Ibadan
Ilishan-Remo, Nigeria

Abstract- Past implementations of k-means algorithm on traditional MapReduce framework used heuristic to determine the terminating point of k-means algorithm iteration. Whereas heuristic is not good enough for an empirical study. Therefore, suitable convergence criteria such as minimum sum of square error (MSSE) which aims to minimise sum of square error of both intra and inter cluster should be used. MSSE is an important convergence criteria for a k-means algorithm because it uses only data for its computation and thus reduces noise, improves accuracy, and reliability. However, computation of MSSE as a convergence criteria for a k-means algorithm has been so difficult to express on a MapReduce framework. In this work we discover the reasons for the difficulty and proposed that a solution can be provided by using a new complex data structure for the key-value pair of MapReduce framework and introduction of a new optimization function in the reduce phase.

***Keywords*—k-means, MapReduce, minimum sum of square styling.**

I.INTRODUCTION

Clustering is an unsupervised learning of a hidden data concept which is generally divided into two types: hierarchical, and partition-based (Kaur Supreet, and Kaur Usvir , 2013). K.Jain, 2010, dated the first data-clustering article to 1955. Generally, the objective of clustering is to develop an automatic algorithm that will discover natural groupings in an unlabeled data (MacQueen, 1967). Partition-based clustering technique is based on iterative relocation of data points between clusters (Ayramo and Tommi, 2006). It aims to find global optimality by exhaustive enumerations of all partitions. K-means and k-mediods are examples of heuristic to achieve this global optimality (Jiawei Han and Micheline Kamber, 2006). But instead these heuristic methods are only sure of local optimal and can only work with small and moderate size data. So many research papers have concluded that for a scalable clustering algorithm, the algorithm should work in parallel (Wooyoung Kim, 2009) and (C.L. Philip Chen, and C.-Y. Zhang, 2014).

There are also researches which extended the algorithms of these heuristics to handle very large data sets e.g. (Madjid Khalilian et al., 2010) and (Silvia Nittel et al., 2004).Other researches made use of parallelization of data (Kilian Stoffel and Abdelkader Belkoniene, 1999) and (Xiaobo Li and Zhixi Fang, 1989) and (Sanpawat Kantabutra and Alva L. Couch, 2000) and (Jitendra Kumara et al., 2011). This present work focuses on k-means algorithm, and use of parallelization technique to handle big data.

Three computations are carried out in a k-means algorithm. The aim of these computation is to ascertain that data in the same group called cluster are similar to each other and dissimilar to data in another group. The similarity is computed using a distance measure called Euclidean distance. K-means starts by choosing a k number of initial points that serve as initial centroids. The first one is the distance computation. It computes the Euclidean distance of each data to the selected initial centroids, and group each data with the centroid nearest to it. Secondly, it computes new means or new centroid for each resulting clusters. Thirdly, convergence of each group is then computed using minimum sum of square error (MSSE). MSSE is the objective function which k-means algorithm seeks to minimise. It is best used as convergence criteria for a k-means algorithm because it uses only data for its computation therefore reduces noise, improve accuracy, and reliability. Thus, MSSE can computationally determine the terminating point during iteration of a k-means algorithm.

Distance computation in a k-means algorithm becomes complex as the number of objects increases. Thus, divide and conquer principle is applied to very large data. This implies that k-means algorithm works separately and simultaneously with different division of data and then combine the results of each division in order to arrive at final solution. Different parallel frameworks by researchers such as (Sebastian Celis and David R. Musicant, 2002) and (Philipp Haller and Heather Miller, 2011) and (Aleksandar Prokopec et al., 2011) etc. have

been developed, one of such is MapReduce framework. MapReduce framework was developed by Google developers. It is capable of not only to parallelise data but also computation.

Apart from the distance computation which can be complex with very large data set, MSSE computation can also be very complex. The computation of MSSE is found to be an NP-Hard problem when *k i.e. the number of cluster is >=2* and the data is one dimensional. The same thing also holds when *k=2* with multidimensional (Daniel Aloise et al., 2012). For one dimensional data time complexity is $O(n^3)$, for more than one dimensional data time complexity is $O(n^{sk+1})$ time. This is very time consuming.

Nevertheless, computation of MSSE as a convergence criteria for a k-means algorithm is difficult to express in a MapReduce algorithm (Böse Joos-hendrik et al., 2010) and (Jimmy Lin and ChrisDyer, 2010). In this work we discover the reasons for the difficulty in expressing MSSE in a MapReduce framework, and proposed that a solution can be provided by using a new complex data structure for the key-value pair of MapReduce framework and introduction of a new optimization function in the reduce phase.

The following sub-section discusses the related works, section two explains the MapReduce framework and implementation of k-means on it. Section three discusses challenges of computing MSSE on traditional MapReduce framework and the proposed solution.

RELATED WORKS

Early researches on the use of parallel framework to handle analysis of data is only based on parallelising the data in a shared-memory architecture. In addition, these framework used restrictions of restrictive programming model to parallelise computation automatically. These algorithm are processor-intensive (Jimmy Lin and ChrisDyer, 2010) which only have rudimentary support for dealing with large amount of data. Thus, MapReduce programming model, a data-intensive framework, was developed by Google researchers (Dean Jeffrey and Ghemawat Sanjay, 2004) and (Dean Jeffrey and Ghemawat Sanjay, 2008). Weizhong Zhao et al.,2009, published the first known paper where k-means algorithm was implemented on MapReduce programming model. They successfully used MapReduce to parallelize the data and to parallelize distance computation. The success of Weizhong Zhao et al., 2009, was based on the possibility that k-means distance computation between an object 'a' and the center is irrelevant to the distance computation between another object 'b' and the cluster centre. This implies that the distance measures of different objects to the cluster can be computed in parallel. This easily solves the computational challenge of k-means, which involves a total of *nk* (*n* is the number of objects while *k* is the number of clusters) distance for each cluster to compute the distance between each object and the centre of that cluster. Thus, Weizhong Zhao et al., 2009 implemented k-means algorithm on a MapReduce framework. The evaluation shows that the parallel k-means algorithm on MapReduce framework can handle very-large dataset with nearly linear-speedup. In terms of scale-up, it is found that the run-time of PKmeans (Parallel Kmeans) remains constant as the original system when *m-larger system* processed *m-larger jobs.* The size-up also increases as the size of dataset increases.

Though this was successfully implemented but the convergence criteria using MSSE was not implemented. The authors used heuristic to determine when to terminate the iterative clustering process. Nevertheless, heuristic is difficult to validate. However, MSSE is the key criterion for convergence in order to achieve empirical result (Daniel Aloise et al., 2009).

Böse Joos-hendrik et al., 2010, developed a shared-memory online MapReduce model as opposed to the traditional shared-nothing batch-oriented MapReduce framework. The aim is to reduce time-to-solution of a MapReduce framework. To achieve the aim, the new model is able to monitor progress and estimate convergence of preliminary results. This new model works well for an algorithm whether it is iterative or non-iterative. Progress is based on the relative progress of existing reducers. Progress of a reducer is measured by the fraction of *values* (*value1….valuen*) which correspond to key *k* that the reducer handling key *k* has processed so far. The total number or input which corresponds to key *k* is measured by monitoring the size of all input data processed by all mappers, and for each *k* how many pairs are emitted by mappers per input byte. The convergence curve is used to estimate the expected changes of the results during processing. This is done by plotting the metric dif (dif means difference) on the *y-axis* against the processing progress (*x-axis*). Metric dif() which is >=0 is the algorithm specific difference between two intermediate results and it is stored as *signatures* sig(*R*). For a k-means algorithm the authors expressed metric dif() as the distance between two intermediate results in an auxiliary cluster (AC). The distance is measured in terms of MSSE. Note that MSSE computation is based on two intermediate results while in traditional k-means algorithm the MSSE is based on sum of distance between the new centroid and each data point in a cluster. Furthermore, their work is not on the whole dataset, can only estimate within the cluster convergence, cannot perform inter-cluster dissimilarity among the clusters.

MAPREDEDUCE

Google Incorporated, developed MapReduce programming model as a solution to the complexity of code to deal with parallelization of computation, distribution of data, and handling of failure in a distributed network. This relief the Programmers from writing complex codes and thus concentrate on providing simple and straightforward computations. Dean Jeffrey and Ghemawat Sanjay, 2004, presented MapReduce as the simple and powerful interface for automatic parallelization and distribution of large-scale computations, and show how its implementation achieves high performance using large clusters of commodity machines.

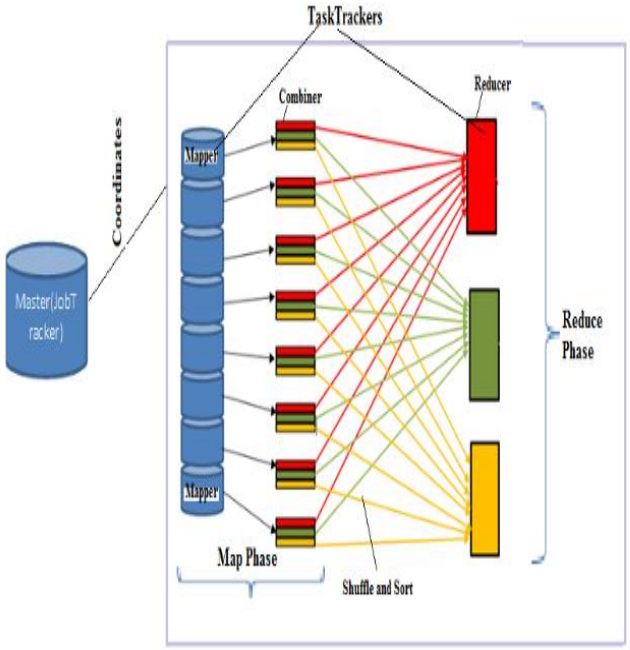


Figure 1: Architectural Framework of MapReduce

MapReduce is a library consisting of two functions, Map and Reduce. The map function takes in a set of input key/value pairs and the reduce function produces a set of output key/value pairs. Algorithm on MapReduce can be optimized with the use of partitioners and combiners. Combiner can be described as local aggregate function which acts like a mini-reducer, while partitioners is used for allocating unique intermediate results from the combiners to different reducers.

Figure 1 above is the architectural framework of MapReduce. The MapReduce framework uses one node as the master called the JobTracker and the other nodes are the slaves called the TaskTrackers. All these nodes are inexpensive commodity machines. The master/JobTracker assigns map and reduce tasks to different TaskTrackers, keep the state (idle, in-progress, or completed) of each map and reduce tasks, and informs the reduce tasks of the location of intermediate data on the mapper local disk. It also keeps location information of the input files (or blocks), i.e. the worker machine containing each input file and those ones containing the replica of each input file. Note that map and reduce tasks are user specified code to solve the problem at hand. In addition, multiple map tasks or reduce tasks can be assigned to different TaskTrackers.

MapReduce framework works on a distributed file system called Google File System (GFS), (Ghemawat, S., Gobioff, H., and Leung, S.-T, 2003). GFS is a Google proprietary distributed file system, another file system was developed by Yahoo!, and this is called Hadoop distributed file system (HDFS) (Tom White, 2009). The distributed filesystem (DFS) is the storage system, while MapReduce is the analysis system (White, 2009). A DFS is a filesystem that manages storage across a network of machines (White, 2009). Job submitted by the client is broken into blocks of not less than 128MB (this is called an input split). The DFS is divided into NameNode and DataNode. The NameNode determines the DataNode that will store each input split. Thus the DFS allows the client to write into the local disk of such DataNode.

MAPREDUCE AND K-MEANS

Weizhong Zhao, Huifang Ma, and Qing He, 2009, used the map function for distance computation of each data points in a block to the centroids and then determine the closest centroid to each data point. The map function takes in as input <key, value> pair which represent a record from the sequence of file stored in an HDFS, and an array of the initial centres or centroids. The key (*key*) is the offset (in bytes) of the record to the start point of the data file. The output is the intermediate results that consists of index of the closest centre point (*key'*), and sample information. The intermediate data is stored in the local disk of the host. The combiner function will fetch the intermediate data of the same map task. The input of the combiner is the intermediate data. The combiner then computes the partial sum of the value of instances attached to the same centre point, and the count of such instances. This information is then transferred to the reducer where the aggregate sum of closest points to each center is computed from the partial sum computed in the all the combiners. After this, the mean is computed and this forms the new centers, which will be the centers for the next iterative step. Note that MapReduce is a batch-oriented framework, this implies that until the map function finishes, the combine function will not start, and until the combiner finishes the reduce function will not start. Furthermore, the computation of distance is done in

parallel, the iterative steps are serial, and the architecture is a shared-nothing architecture (no sharing of memory or data).

Challenges of Computing K-means MSSE on MapReduce Framework

It is difficult for traditional MapReduce framework to compute K-means MSSE due to the following reasons: firstly, each map function only handles part of data therefore the intermediate results passed to the combiner is a partial intermediate results.

Also, each combiner acts in isolation and does not have access to output from other mappers. Secondly, the combiner aggregates the intermediate results before passing it to the reducer. Thus the reducer could only find the new centroids since it does not have access to intermediate results.

In this research, we find out that for the reducer to be able to compute the MSSE it needs to have access to the intermediate results and be optimized with another function which can compute the MSSE. The MSSE will generate the objective function, and each objective function from different reducers will be passed to an online collector which will aggregate them to produce the global objective function.

Proposed Solution

One of the way to improve the efficiency of an algorithm in a MapReduce framework is to use complex data structure. Another one is the use of optimizers in extending and improving the efficiency of an algorithm (Jimmy Lin and ChrisDyer, 2010). Thus this work makes use of complex data structure and efficient use of optimizers to explicitly express the computation of MSSE and to overcome the limitation in (Weizhong Zhao, Huifang Ma, and Qing He, 2009).

Use of new complex data structure different from the ones used by (Weizhong Zhao et al., 2009) will give the reducer the opportunity of access to instances belonging to each cluster. Thus, the reducer will be able to compute the MSSE. Table **I** shows the input and output based on the data structure used by (Weizhong Zhao et al., 2009). Table **II** shows the input and output based on the newly proposed complex data structure of this current research.

Table I: *Complex Data Structure Used By Weizhong Zhao et al., 2009.*

Function	Input	Output
Map task	<key, value>	<key', string of values'>
Combiner (optimizer)	<key', string of values'>	<key', string consisting sum and count'>
Reduce task	<key', string consisting sum and count'>	<key', value''>

Table II: *Newly Proposed Complex Data Structure*

Function	Input	Output
Map task	<key, value>	<key', string of values'>
Combiner (optimizer)	<key', string of values'>	<key', string consisting of values', sum, and count'>
Reduce task	<key', string consisting of values', sum, and count'>	<key', string of value consisting of mean and values'>
MSSE function	<key', string of value' consisting of mean and values'>	<key', string of new mean and objective function>

CONCLUSION AND FUTURE WORK

The practical experiment is ongoing using simulated MapReduce framework as a single-threaded Java application. The second stage is to experiment simulated MapReduce framework using multithreaded Java application, and the last stage will be the deployment and implementation of the proposed algorithm into MapReduce framework on Hadoop distributed file system. Each experiment will be tested using Nigeria Population Census data of 2005.

REFERENCES

[1] Böse Joos-hendrik and Andrzejak Artur and Högqvist Mikael. (2010). Beyond Online Aggregation: Parallel and Incremental Data Mining with Online MapReduce. Proceedings of the 2010 Workshop on Massive Data Analytics on the Cloud.

[2] C.L. Philip Chen, C.-Y. Zhang, Data-intensive applications, challenges, techniques and technologies: A survey on Big Data, Inform. Sci. (2014), http://dx.doi.org/10.1016/j.ins.2014.01.015

[3] Daniel Aloise, Amit Deshpande, Pierre Hansen and Preyas Popat. (2009). NP-hardness of Euclidean sum-of-squares clustering. Machine Learning, 245-248.

[4] Dean Jeffrey and Ghemawat Sanjay. (2004). MapReduce: simplified data processing on large

clusters. Proceedings Of The 6th Conference On Symposium On Operating Systems Design And Implementation (pp. 137-149). San Fransico: USENIX Association.

[5] Dean Jeffrey and Ghemawat Sanjay. (2008). MapReduce: simplified data processing on large clusters. ACM Communication, 107-113.

[6] Ghemawat, S., Gobioff, H., and Leung, S.-T. (2003). The Google file system. 19th Symposium on Operating Systems Principles (pp. 29-43).

[7] Jiawei Han and Micheline Kamber. (2006). Data Mining: Concepts and Techniques. San Francisco: Morgan Kaufmann.

[8] Jimmy Lin and ChrisDyer. (2010). Data-Intensive Text Processing with MapReduce: Morgan & Claypool.

[9] Jitendra Kumara, Richard T. Millsa, Forrest M. Hoffmana, William W. Hargroveb. (2011). Parallel k-Means Clustering for Quantitative Ecoregion Delineation Using Large Data Sets. International Conference on Computational Science, ICCS 2011 (pp. 1602–1611). Singapore: Procedia Computer Science.

[10] K.Jain, A. (2010). Data Clustering: 50 years beyond K-means. Elsevier Science, 651-666.

[11] Kaur Supreet, and Kaur Usvir . (2013). A Survey on Various Clustering Techniques with K-means Clustering Algorithm in Detail. International Journal of Computer Science and Mobile Computing , 155-159.

[12] Kilian Stoffel and Abdelkader Belkoniene. (1999). Parallel k/h-Means Clustering for Large Data Sets. Euro-Par '99: Proceedings of the 5th International Euro-Par Conference on Parallel Processing (pp. 1451–1454). London: Springer-Verlag.

[13] MacQueen. (1967). Some Methods for Classification and Analysis of MultiVariate Observations. Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability (pp. 281-297). Berkeley: Berkeley: University of California Press.

[14] Madjid Khalilian, Norwati Mustapha, MD Nasir Suliman, MD Ali Mamat. (2010). A Novel K-Means Based Clustering Algorithm for High Dimensional Data Sets. Proceedings of the International MultiConference of Engineers and Computer Scientists (pp. 503-507). Hong Kong: IAENG.

[15] Pierre Hansen and Daniel Aloise. (2010). A survey on exact methods for minimum sum-of-squares clustering.

[16] Sanpawat Kantabutra and Alva L. Couch. (2000). Parallel K-means Clustering Algorithm on NOWS. Technical Journal, 243-248.

[17] Silvia Nittel, Kelvin T. Leung, and Amy Braverman. (2004). Scaling Clustering Algorithms for Massive Data Sets using Data Streams. ICDE '03 Proceedings of the 19th International Conference on Data Engineering (pp. 830-841). Washington: IEEE Computer Society.

[18] Tom White. (2009). Hadoop: The Definitive Guide. ??: O'reilly and Yahoo! Press.

[19] Weizhong Zhao, Huifang Ma, and Qing He. (2009). Parallel K-Means Clustering Based on MapReduce. CloudCom 2009 (pp. 674-670). Heidelberg: Springer-Verlag.

[20] White, T. (2009). Hadoop: The Definitive Guide. United State: O'Reilly Press.

[21] Wooyoung Kim, Parallel clustering algorithms: survey, in: Parallel Algorithms, Spring, 2009. http://www.cs.gsu.edu/~wkim/index_files/SurveyParallelClustering.pdfs as at September 30, 2014

[22] Xiao, H. (2010). Towards Parallel and Distributed Computing in Large-Scale Data Mining: A Survey. ???, 1-30.

[23] Xiaobo Li and Zhixi Fang. (1989). Parallel clustering algorithms. Parallel Computing (pp. 275-290). Elsevier.

Model-based Development of Digital-Library Information System

Akintola K.G.
Computer Science Department,
University of Houston-Victoria,
Texas 77901
United States of America.
Akintola2087@yahoo.com

Ojokoh B.A.
Computer Science Department,
Federal University of Technology
Akure
bolanleojokoh@yahoo.com

Boyinbode O.K
Computer Science Department,
University of Cape Town,
Rondesbosch 7701
South Africa.
okboyinbode@yahoo.com

Abstract- In this paper we present the report of development of a web-based library information system using OOHDM methodology. We investigate web development efforts using an engineering approach against the hacking practices of the old time technique of web development. A digital library is chosen as a case study for this project. A digital library is a library in which collections are stored in digital formats (as opposed to print, microform, or other media) and accessible by computers. The digital content may be stored locally, or accessed remotely via computer networks. A digital library is a type of information retrieval system [Wikipedia] system.
We find that engineering web application solves some of the problems found in hacking practices of web development. There is an improvement in areas such as usability, scalability, security, maintenability, and repeatability of web applications. The organization of this report is as follows: In Section 1, we present the motivation for our research. Section 2 describes the OOHDM methodology. Section 3 presents system development. In section 4 we present system implementation and finally section five summarizes our findings and outline areas for further research. This development is according to postulations of Schwabe and Rossi [1998].

I. MOTIVATION FOR THE RESEARCH

In the roundtable discussions on the topic "Do old tried and true principles of software engineering make any sense on completely new playing field of Internet?" Various software leaders gave reasons for why Web should be engineered and some believe that the hacking practice that is operational from the beginning should be stuck to while others seek a new approach that reflects the nature of web applications. One of the software leaders Tom Giib in (pressman 1998) who supported the notion of engineering the web applications made this statement: "Internet based systems should be engineered. The top few critical and qualitative requirements should be specified and a suitable architecture derived from this as well as from cost and other constraints. Not doing so substantially increases the projects risk of failure because of misunderstanding of what the project is all about".
Web development right from the onset and up till the present moment has been dominated by the Hacking practices. Very many popular portals started from primitive simple sites developed by amateurs (Ginige 2001). Due to the limited knowledge of engineering principles by the developers, hacking was the adopted option. However to justify their practice some reasons were given as responsible for this approach among which are according to (pressman 1998), web based systems must be conceived, planned and produced in record time. Also that web based systems have short lifetimes.
However, these sites developed by ad-hoc methods have proved to be difficult to maintain, to scale and ultimately fail. The cause of these failures are lack of vision, short sighted goals , flawed design and development process, and poor management of development efforts – not technology (Ginige etal 2001).
So to better manage web based systems design and development in a systematic and repeatable manner we need a process that outlines the various phases, steps, and activities of web-based systems development (Ginige etal 2001).
Web engineering is a discipline concerned with establishment and use of sound scientific, engineering and management principles and discipline and systematic approaches to the successfully development, deployment and maintenance of high quality web-based systems and applications (Murugesan etal 2001).
Web based systems have moved from ordinary web presence systems to full-fledged and complex application systems such as e-commerce which requires an engineering approach and a sound methodology in order to successfully build the systems. Not only this, a sound process helps developers address the complexities of web-based systems , minimizes the development risks and deals with the likelihood change and also delivers the web applications within the time providing feedback for management as the project continues (Ginige etal 2001).
Also in Ginige etal (2005) a well engineered system is supposed to be: functionally complete and correct, usable, Robust, reliable , Maintainable ,secure , perform satisfactorily, scalable, portable, reusable, interoperable, universally accessible, and well documented. So we really need to engineer the web to meet with all these requirements.
Building complex Web applications such as e-Library applications is a time consuming task. We need to understand the underlying domain. We must carefully design the applications' architecture and user interface, if we want them to be usable. Similar web based applications possesses the same requirements. So there is the need for engineering web based systems so that when developing similar systems in future we can reuse the design and thereby economize time of system development.
Also nowadays, greater emphases are now on performance, correctness, and availability of web based systems so the development process will assume greater significance.

In this paper we present a report of a web-based library information system which was designed and developed using the web engineering technology called the object oriented

hypermedia technology (OOHDM).

II. THE OOHDM

According to Schwabe and Rossi in(John Karat 2008) **OOHDM** (**Object Oriented Hypermedia Design Methodology**) is a method for the development of web applications. It was one of the first methods to postulate the separation of concerns that defines its various models - requirements, conceptual, navigation, abstract interface and implementation. OOHDM, and its successor, SHDM (Semantic Hypermedia Design Method, which uses semantic models) are supported by an open source, freely available environment, HyperDE.(wikipedia). It is a model-based approach to develop Web applications. It allows the designer to specify a Web application, seen as an instance of a hypermedia model, through the use of several specialized meta-models. Each model focuses on different aspects of the application.Once these models have been specified for a given application, it is possible to generate run-time code that implements the application.

In OOHDM a Web application is built in a five-step process supporting an incremental or prototype process model. Each step focuses on a particular design concern, and an appropriate model is built. Classification and generalization/specialization are used throughout the process to enhance abstraction power and reuse opportunities. We next summarize the five activities.

The OOHDM design approach

- Clearly separates conceptual from navigational aspects, and interface layout
- Navigation objects are OO views on conceptual objects
- Structure the navigation space into sets called navigational contexts
- Two schemas: Navigational Class Schema and Context Schema

i Requirements Gathering

The first step is to gather the stakeholder requirements. To achieve this, it is necessary to first identify the actors (stakeholders) and the tasks they must perform. Next, scenarios are collected (or drafted) for each task and type of actor. The scenarios are then collected to form use cases, which are represented using User Interaction Diagrams (UIDs). The UIDs are validated with the actors, and redesigned if necessary.

ii Conceptual Design

In this step a conceptual model of the application domain is built using well-known object-oriented modeling principles. There is no concern for the types of users and tasks, only for the application domain semantics. A conceptual schema is built out of subsystems, classes, and relationships. OOHDM uses UML with slight extensions, to express the conceptual design.

iii Navigational Design

In OOHDM, an application is seen as a navigational view over the conceptual model. This reflects a major innovation of OOHDM [also adopted by other methods such as UWE Schwabe and Rossi in John Karat (2008) which recognizes that the objects (items) the user navigates are not the conceptual objects, but other kinds of objects that are "built" from one or more conceptual objects, to suit the users and tasks that must be supported. In other words, for each user profile we can define a different navigational structure that reflects the objects and relationships in the conceptual schema according to the tasks this kind of user must perform. The navigational class structure of a Web application is defined by a schema containing navigational classes. In OOHDM there is a set of predefined basic types of navigational classes: nodes, links, anchors and access structures.

iv Abstract Interface Design

The abstract interface model is built by defining perceptible objects—also called widgets—that contain information (e.g., a picture, a city map, etc.) in terms of interface classes. Interfaces are defined as recursive aggregations of primitives classes (such as exhibitors or capturers) or of other interface classes. Interface objects map to navigational objects, providing them with a perceptible appearance, or to input values. Interface behavior is defined by specifying how to handle external and user-generated events and how communication takes place between interface and navigational objects.

v Implementation

Implementation maps interface and navigation objects to run-time objects and may involve elaborate architectures, e.g., client–server, in which applications are clients to a shared database server containing the conceptual objects. A number of DVD-ROM–based applications, as well as Web–sites, have been developed using OOHDM, employing various technologies such as Java (J2EE), .NET (aspx), Windows (asp), Lua (CGILua), ColdFusion, and Ruby (Ruby on Rails.

III. SYSTEM ANALYSIS AND DESIGN

3.1 Requirements Gathering:

In OOHDM we build a different navigational model for each user profile; in this application we clearly have at least two different user profiles: the reader, who is looking for a book to read or borrow or just for information about the book, and the administrator, who maintains the books database. We will mostly discuss the application for the reader profile. Once we have identified the actors, we must identify the tasks the user will accomplish using the application, in order to obtain usage scenarios. Clearly, there are many tasks to be supported in our application scenario.

Some of the typical tasks for the reader user profile are

- Find a book given its title.
- Find a book given an author's name.
- Find information about an author.
- Find books of a particular author.
- Find recently released books.
- Choose books to borrow given one of the above criteria

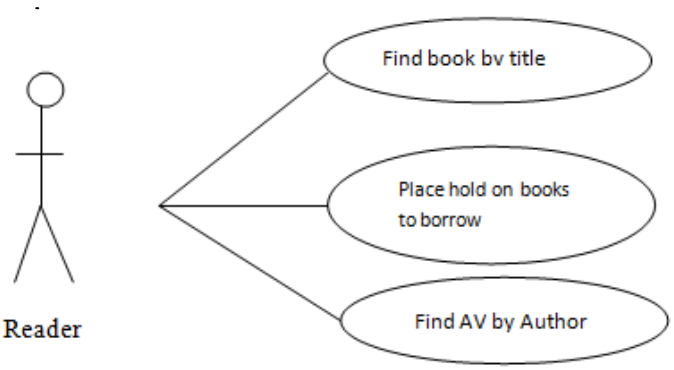


Fig 1 Typical use cases in Library system.

J. Use Case Specification

We next describe the usage scenarios. A scenario represents the set of subtasks the user has to perform to complete a task. Scenarios are specified textually using the point of view of the final user, in our case a reader.

As an example of the first task ("Find a book given its title"), a possible scenario would be:

"I enter the book title or part of it, and I see a list of matching books titles. For each book matching the title, I get some information such as the book cover, the year the book was published, the list of authors, ISBN. I can get additional information such as publishers, edition, pages, summary, pubcode, pubyear, email, publisher's name, address. For some books, I can also see the table of contents. "After reading the information, I can decide to place hold on it or to quit."

After collecting several such scenarios, a generalization is captured in **a use case**, defined next. We use the following heuristics:

1. Identify those scenarios related with the task at hand.
2. For each scenario, identify information items that are exchanged by the user and the application during their interaction.
3. For each scenario, identify which data items are associated among themselves, they typically appear together in the text of the use case.
4. For each scenario, identify those data items organized as sets. Usually, the use case text refers to them explicitly as sets.
5. The sequences of actions appearing in scenarios should also appear in the use case.
6. All operations on data items that appear on scenarios should be included in the use case.

After defining the data involved in the interaction, the sequence of actions, and the operations, we can specify the use case. A use case will be constructed based on the sequence of actions, detailed with the information about the data items and operations involved. Use cases can also be enriched with information from other use cases or provided by the designer.

The resulting use case for the previous scenario is the following:

Use case: "Find a book given its title."

1. The user enters the book title (or part of it).
2. The application returns a list of books matching the data entered or the information about the book (if only one book matches see step 4). For each book, the title, author, author, and cover image are shown.
3. In case the user wants to see more information on the book, he selects it.
4. The system returns detailed information for the book: ISBN, title, cover image , authors etc. If the user wants to place hold on it, he can place-hold to include it in the hold cart to reserve later for borrowing later (use case: reserve a book given its title). If he wants, he can see the table of contents.
5. If the user wants to know information about an author of the book, he can select the author and the application will return his name, date of birth, a photograph, and a list of books he has authored. The specification of other use cases follows a similar process.

User Interaction Diagrams

- Specify the interaction described in a use case
- Are used to support the communication between the designer and users and to validate the use cases
- Serve as input for obtaining the class diagram through the application of guidelines
- Can be incorporated in the requirements and analysis workflows of the Unified Process
- Are used also as input for synthesizing the Navigation Design

Use cases are described using a graphical notation called a User Interaction Diagram (UID), which captures the flow of information and helps detail the information items and choices made by the user. The specification of UIDs from use cases can be done following the guidelines described below. For illustration purposes, we detail the process of building the UID for the use case "Find a book given its title," as described above.

1. Initially, the use case is analyzed to identify the information exchanged between the user and the application. Information provided by the user and information returned by the system are tagged accordingly. Next, the same information is identified and made evident in the use case.
2. Items that are exchanged during the interaction are shown in UID states. Information provided by the user and that provided by the system are always in separate states. Information that is produced from computations should be in separate states from the information used as input to this computation. The ordering of states depends on the dependencies between data provided by the user and returned by the application. In Figure 2, we show the first draft of the UID where parts of the use case are transcribed; information exchanged is shown in italics.
3. After identifying the data items exchanged, they must be clearly indicated in the UID. Data entered by the user (for example, the movie title) are specified using a rectangle; if it is mandatory, the border is a solid line; if it is optional, the border is a dashed line, as shown in Figure 2 Ellipsis (…) in front of a label indicates a list (e.g., …book indicates a list of books). The notation book *(Title, Author(Name), CoverImage)*

is called a *structure*. A shaded ellipsis represents a separate UID.

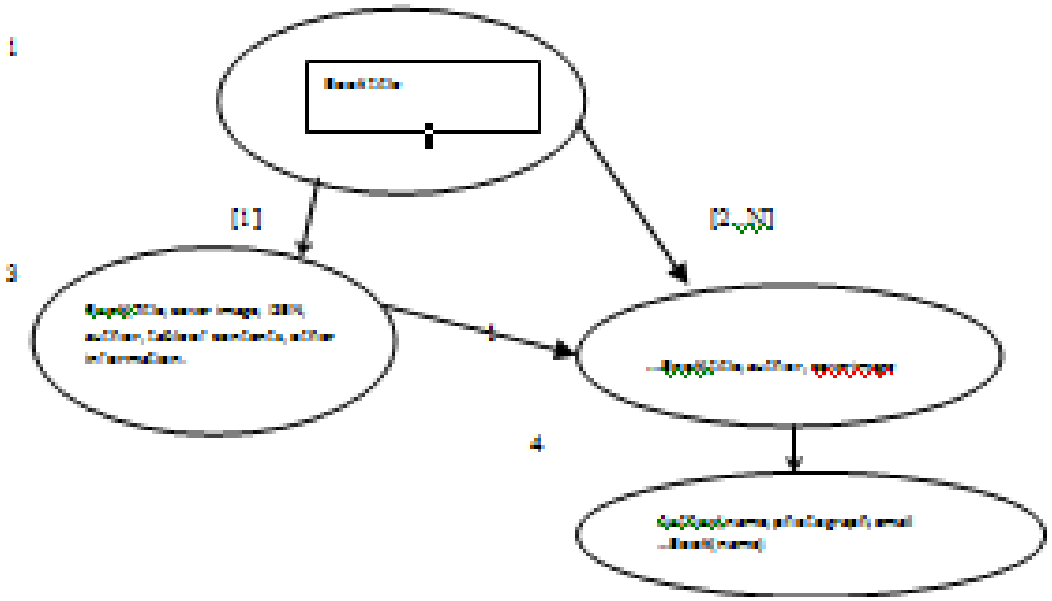


Fig. 2. Defining the UID.

4. Transitions between interaction states must be indicated using arrows. Multiple paths as indicated in the use cases might arise as shown in Figure 4. Labels between brackets indicate conditions (e.g., [2..N] indicates more than one result); a label indicating cardinality represents a choice. (In the example, "1" indicates only one option may be chosen. For any choice, the source of the arrow is the list from which the option is selected or the whole state if it is not ambiguous.)

5. Finally, operations executed by the user are represented using a line with a bullet connected to the specific information item to which it is applied, as shown in Figure 4. The name of the operation appears in parentheses.

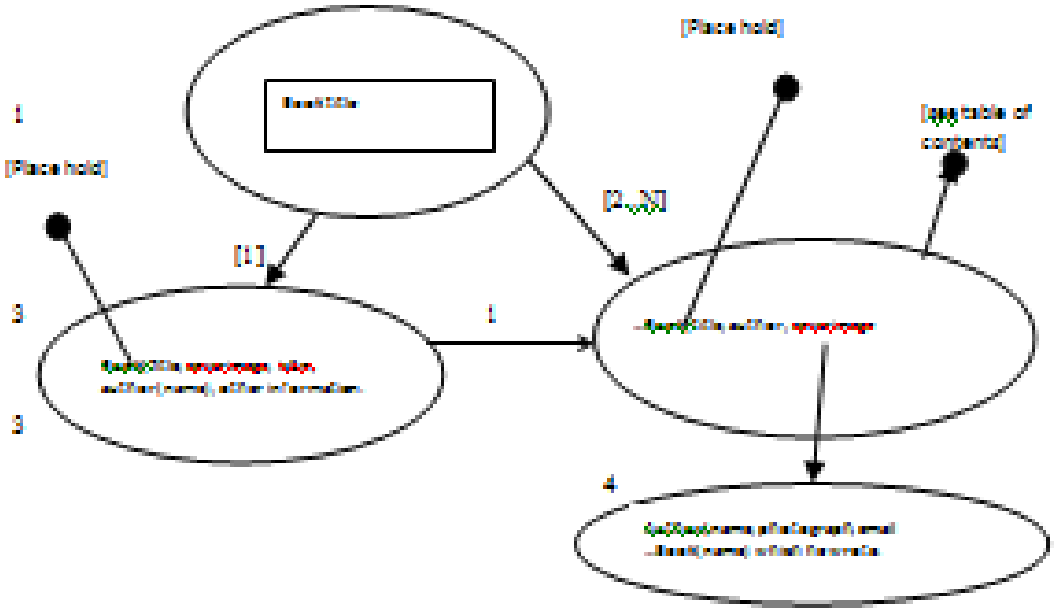


Fig 3. Refining Interactions states in UIDs.

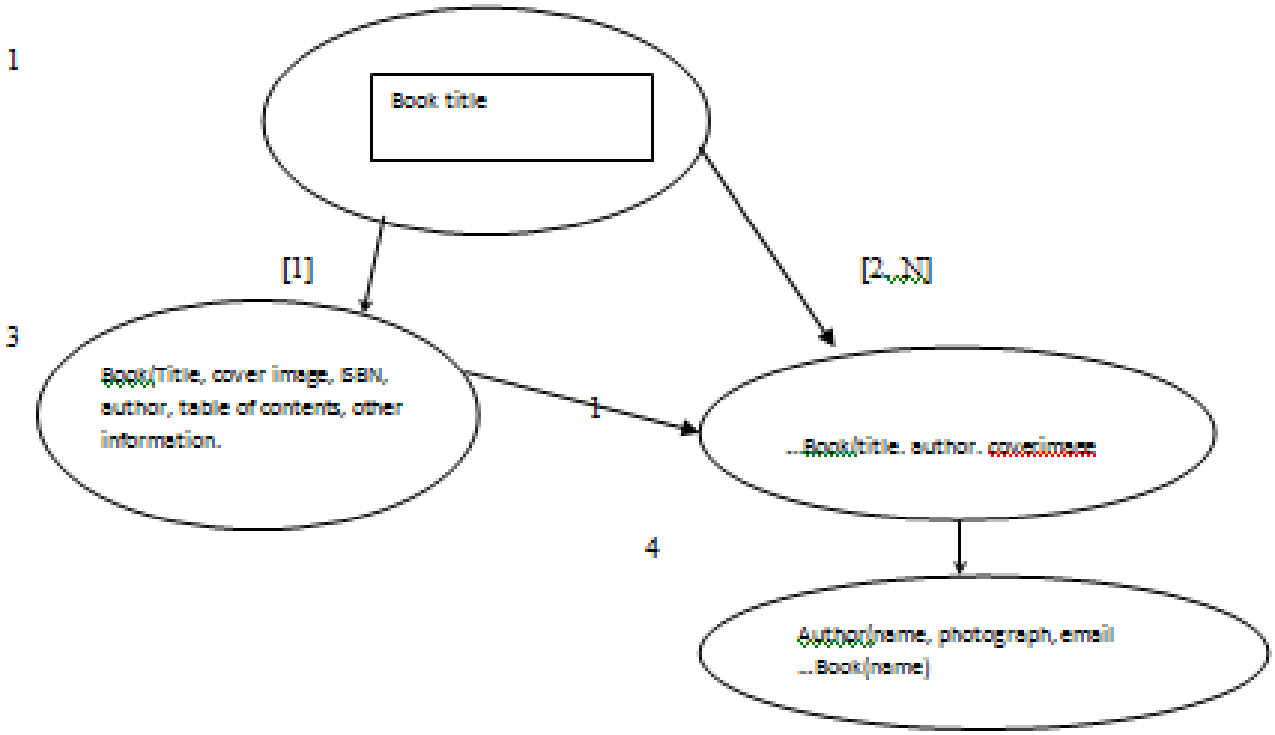


Fig 4 Transitions between interactions states in UIDs.

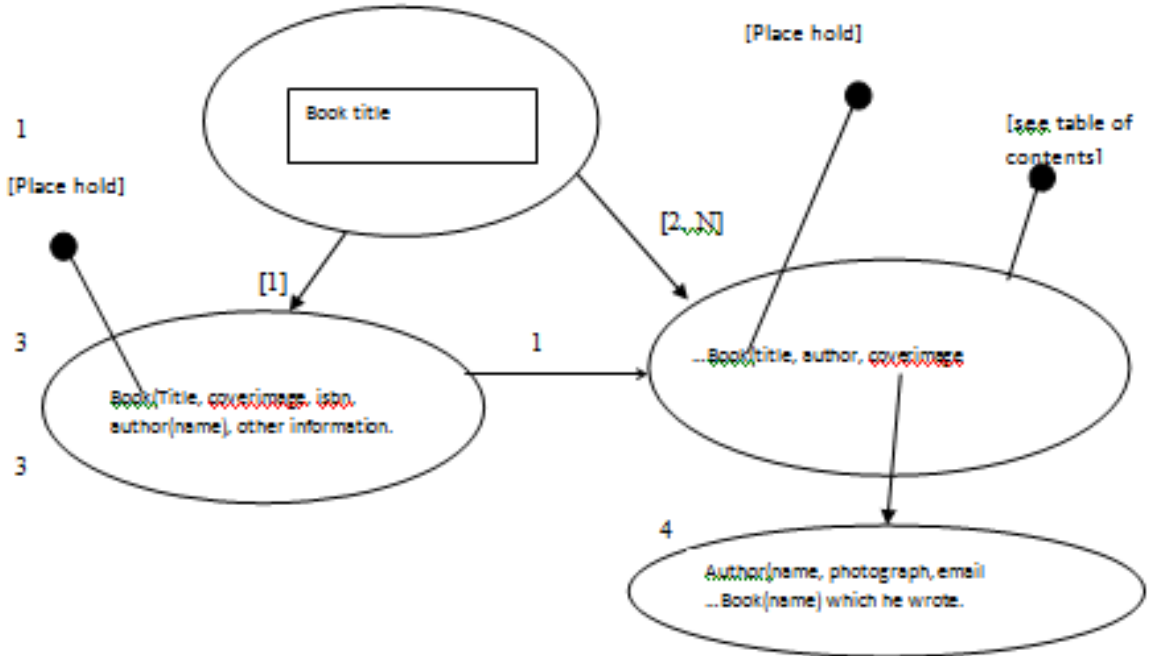


Fig 5. A complete specification of the UID for the use case "Find a book by title including the placehold)

Conceptual modeling

The conceptual model in OOHDM comprises a set of classes (their attributes and behaviors) and their relationships using UML. To develop a conceptual model, the information gathered from use cases and UIDs can help to identify core information classes that can later be refined. Fig 6 is a conceptual model derived from the UID "Find book given its title." using Figure 5.

After several refinements have to be made by hand, incorporating the designer understands of the domain. Among other concepts, the designer must identify

• Generalization and specialization hierarchies—for example, Book and AV can be recognized as subclasses of collection.
• Association classes—for example, Author writes book.
• Hidden classes—for example, reservation.
• Redundant classes.
• The parity of relations.

This conceptual model might need further improvements as the application evolves, since these classes are "just" the ones we derive from the requirements gathering activity. After analyzing the complete set of UIDs and performing needed adjustments, we can obtain the conceptual model of Figure 7. Notice that we have included a generalization class collection, abstracting both books and AVs.

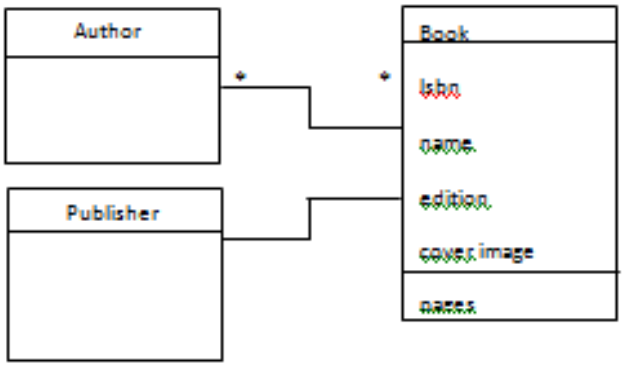


Fig. 6 Initial conceptual model

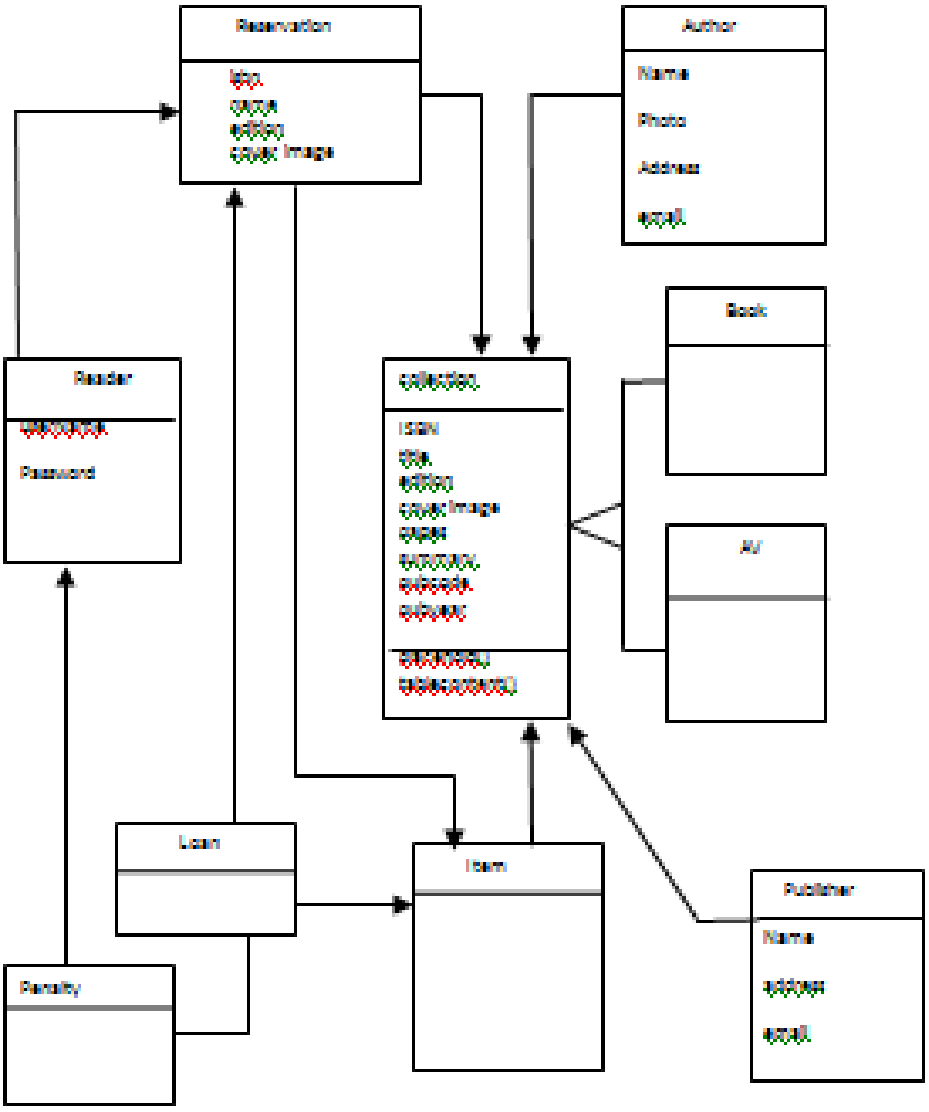


Fig.7 Conceptual Model for the Library website.

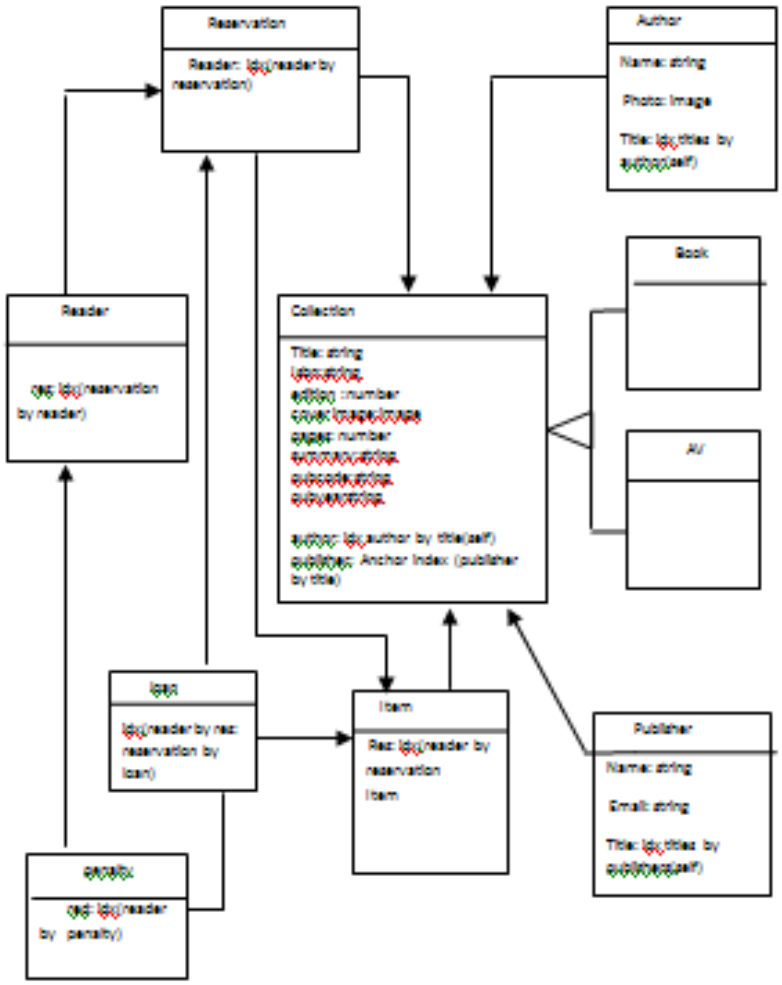


Fig. 8 Navigational class Model for the Library website.

IV. NAVIGATIONAL DESIGN

Navigational Schema

For each set of user profiles, we define a different navigational class schema and context schema. The navigational schema contains the nodes and links of the application. Nodes contain perceivable information (attributes) and anchors for links. Anchors are objects that allow triggering links. Links, meanwhile, are the hypermedia realization of conceptual relationships.

Nodes and Anchors

Nodes are derived from conceptual classes by selecting those classes we want that the user to perceive; attributes are defined in an opportunistic way according to usage needs. Sometimes it is necessary to combine attributes from different objects to describe a node. In the example we may want nodes representing books to contain an attribute with the names of all the authors that wrote the book, eventually using the names as anchors to each Author's page. As shown in the conceptual model of Figure 7, the name of the author is an attribute of Class Author and should not be included in Class book. As an example we define the Node class Book, including as one of its attributes the name of the publisher and an author for the link that connects both nodes. We say that the conceptual class book is the subject of Node class book. In OOHDM we defer the definition of how objects will be perceived until the interface design activity.

Nodes may also possess attributes that are used to trigger operations in their object counterparts in the conceptual model.

```
NODE Book [FROM Book :m]
publisher: String [SELECT Name] [FROM publisher:P WHERE P publishes
m
.... (other attributes "preserved" from the conceptual classBook}
publishedBy: Anchor [PublishedBy, Publishers in Alphabetical order]
```

Links

Links connect navigational objects. The result of traversing a link is expressed either by defining the navigational semantics procedurally as a result of the link's behavior or by using an object-oriented state-transition machine similar to Schwabe and Rossi in (John Karat 2008). The syntax for defining Link classes also allows one to express queries on relationships as shown in the example below.

LINKPublishedBy
SOURCE: Book: M
TARGET: Publisher:P
WHERE S.P publishes S.M
END

Navigational Contexts

OOHDM structures the navigational space into sets, called navigational contexts, represented in a context schema. Each navigational context is a set of nodes, and it is described by specifying its elements, indicating its internal navigational structure (e.g., if it can be accessed sequentially) and associated indexes. Generally speaking, contexts are defined by properties of its elements, which may be based on their attributes or on their relations, or both. Navigational contexts usually induce associated access structures called indexes, which are collections of links pointing to each of the context's elements.

Another way to understand contexts is that they provide an abstraction mechanism that allows us to specify the navigation opportunities available to sets of objects all at the same time, without having to do so for each individual element within the context. In this respect, contexts play a role with respect to navigation that is analogous to the role classes play with respect to object structure and behavior—they allow us to specify navigation properties that are common to all its elements without requiring individual specification. Schwabe and Rossi in (John Karat 2008).

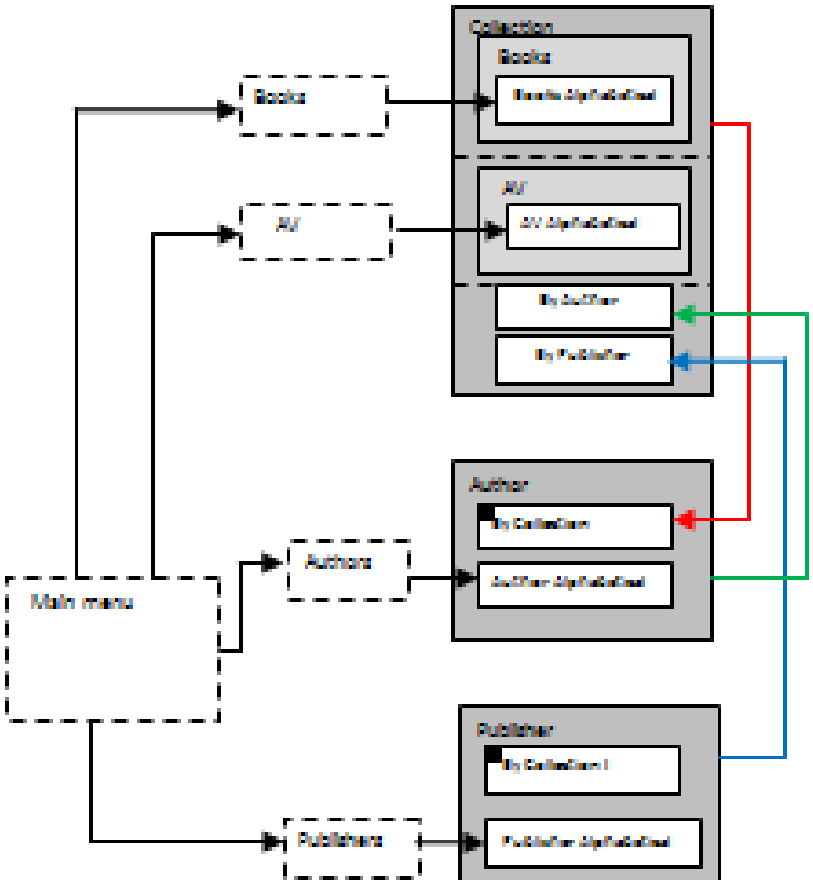


Fig. 9 Partial Navigation context diagram for the Library Users.

Figure 9 contains a portion of the navigational context diagram for our example; we will use it to explain the notation. The details of each context and access structure are described by a context (respectively, access structure) specification card as shown in Figures 10 and 11.

Context: Collection by Author
Type: static
Parameters: a: Author
Elements: c: Collection WHERE a writes b
In context Classes:
Ordering: by name, ascending
Internal navigation: by index (Collection **by Author**)
Operations:
Users: Client **Permissions**: read
Comments:

Figure 10 The context specification card.

Access structure : Collection by Author
Type: simple
Parameters: a: Author
Elements: b: Book WHERE a Writes b
Attributes Target
title: b.name.....Ctx Collection by Author (self)
Author
cover: b.coverimage
"See Table content"..... table_content()
placehold"..... placehold()
Ordering: by name, Ascending
Users: Client **Permission**: read
Comments:

Figure 11. The access structure specification card.

Consider the access structure specification for Book by Author. Since this is an index induced by a context (Book by Author), it will contain one entry for each element in the context. Each entry has four attributes—the (book) title, which is an anchor to the book by Author context (for the movie corresponding to this entry); the author's who writes this book; an activation that allows one to see the table of contents of the book; and an activation that allows one to placehold the book. The former two attributes are calls to methods which will have to be mapped to active interface elements that can trigger the associated operations when activated by the user.

V. INTERFACE DESIGN

The role of the interface is to make the navigation objects and application functionality perceptible to the user, which is the goal of the interface design. From the application business logic's point of view, all that is needed regarding the interface is the definition of the information exchange between the application and the user, including activation of functionalities. The most abstract level is called the abstract interface and focuses on the various types of functionality that can be played by interface elements with respect to the information exchange between the user and the application. The vocabulary used to define the abstract interface is established by an abstract widget ontology **s**hown in Figure 12, which specifies that an abstract interface widget can be any of the following:

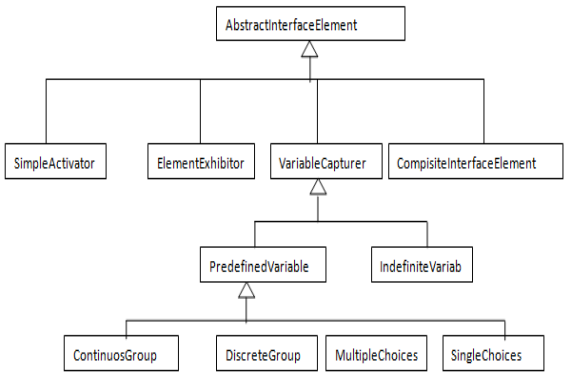


Fig 12 Abstract widget ontology. (Schwabe and Rossi in John Karat (2008)

VI. THE IMPLEMENTATION

We use HTML/CSS/PHP?MySQL. PHP ad MySQL, offers possibilities of a dynamic interactive web applications for the online library. The implementation uses a PhP program for server scripting language, HTML for client side and cascading for formatting the webpages, My SQL for database server. Web browser (like Internet Explorer 4.0, Netscape 4.7, Mozzilla, Opera) are supported by the aplication. The web server used is an Apache web server (1.3.26 software version) and database used is the MySQL database server (3.23.52 software version). MySQl and PHP, was picked for this projects. As a data based management system, MySQL can store very complex information. As a server side scripting language, PHP can perform complicated manipulation of data. MySQL and PHP as a pair have several advantages: They are open source codes,

- They are good for building dynamic Web sites;
- They are easy to use.
- They are fast. So MySQL, and PHP were designed with speed as a major goal.
- They communicate well with one another. PHP has built-in features for communication with MySQL;

VII. FINDINGS AND CONCLUSION

We have been able to present a model driven web based application development in this project using OOHDM approach. We find that adequate planning before modeling and development yields a better solution. Not only that, Information that should be included in the system are readily available after adequate analysis of the requirements. The structure and usability of the system is enhanced from the beginning stage to the final product by segmenting development to different stages. In the project we showed how to build the conceptual model, the Navigational model, and the Interface model. Finally we implemented the model using CSS, HTML, PhP, and MySQL. This approach was found better because it is suitable to both Waterfall and Rapid application development. Not only this, It engenders division of concerns and division of labour which leads to efficiency, timely production of applications. When people do the things they know better, the end result is always good. Having design concerns separated and developed individually was found to lead to an effective and efficient web development. In conclusion, web engineering is the best way to go when developing web applications.

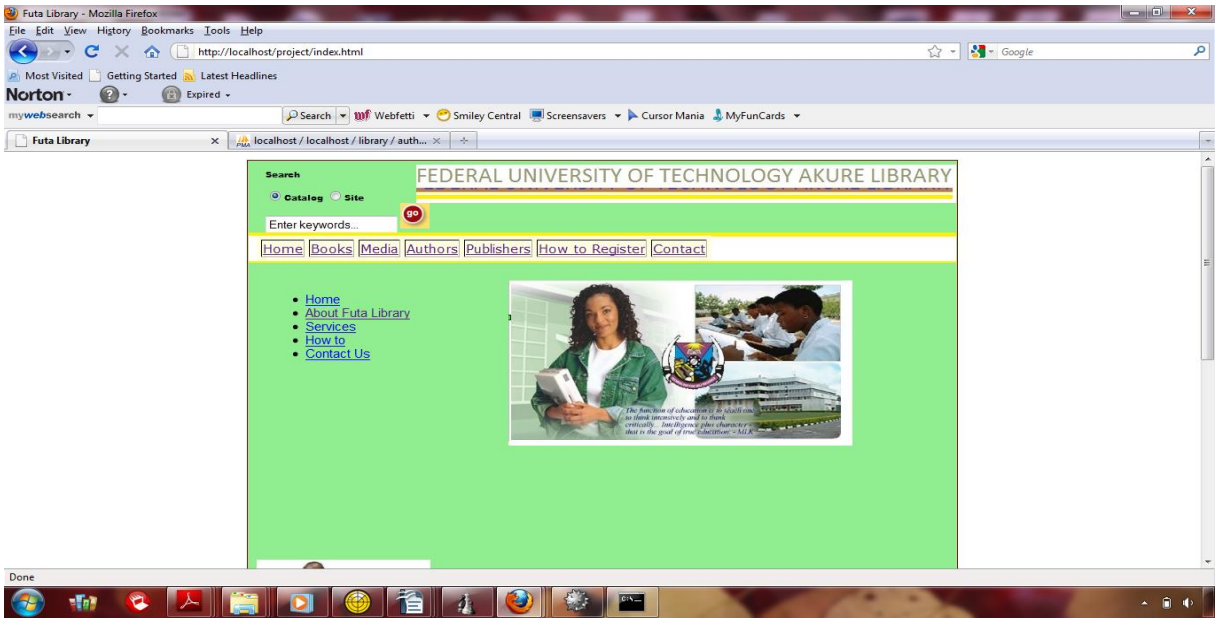
REFERENCES

[1] Athula Ginige and San Murugesan (2001) The essence of Web Engineering "Managing the Diversity and Complexity of Web Application Development IEEE Multimedia.
[2] Bebo White (2002) WEB ENGINEERING Journal of Web Engineering, Vol. 1, No.1 (2002) 003-017 © Rinton Press.
[3] Ceri et al., 2002 Designing Data-intensive Web Applications, Morgan Kaufmann, San Francsisco.
[4] GustavoRossi etal (2008) Web Engineering: Modelling and Implementing Web Applications © Springer-Verlag London Limited 2008.

[5] John Karat (2008) Human-Computer Interaction Series Springer-Verlag London Limited 2008.
[6] Koch and Kraus, 2002 The Expressive power of UML-based web engineering Second International Workshop on Web-oriented Software Technology (IWWOST02) Malaga Spain June.
[7] Murugesan , Yogesh Deshpande, Steve Hansan and Athulua Ginige "Web Engineering: Anew Discipline for Development of Web-Based Systems, Web Engineering LNCS 2016 < Springer pp 3-13(2001)
[8] Pressman (2008) Can Internet -Based Applications Be Engineered? IEE software September/October 1998
[9] San Murugesan Athula Ginige, (2005) Web Engineering: Introduction and Perspectives idea group inc.
[10] Schwabe and Rossi, 1998 An Object oriented Approach to web - Based Application Design, Theory and practice of Object Systems 4(4) Wiley and sons, New York.
[11] Wikipedia http://en.wikipedia.org/wiki/Digital_library.

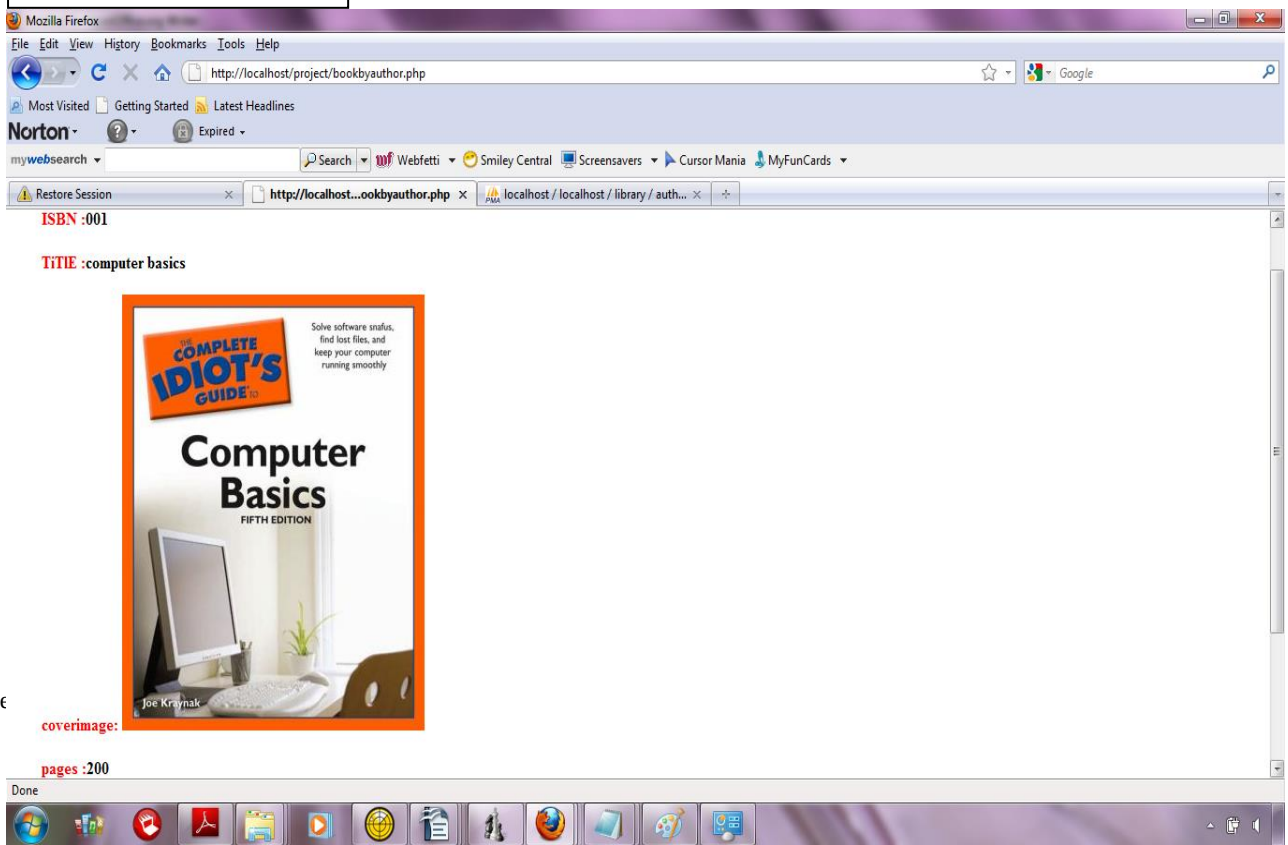
The home page of the application is as shown below.

Home Page



The page given by the context book by authors

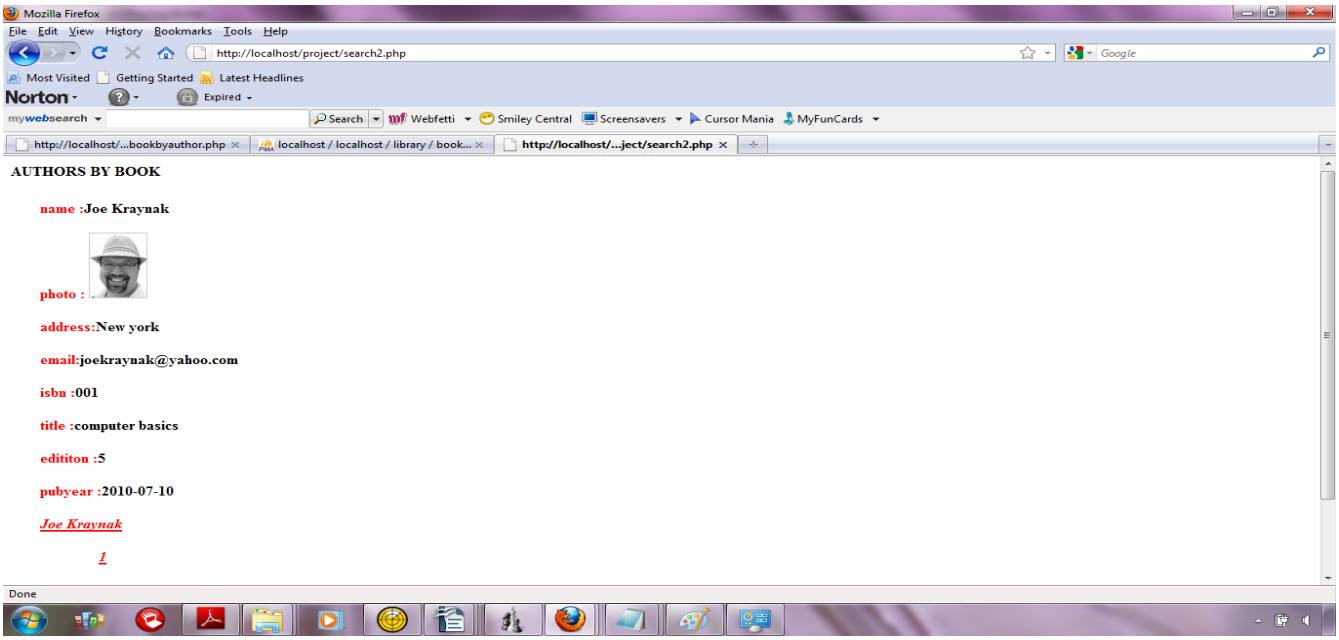
Books by Author



Volume

The Page given the context Authors by book

Authors by Book



The Use and Limitations of Mathematical and Statistical Software Packages in Tertiary Institutions in Zambia

Dayakar Sidhabattula
Department of Mathematics
Rusangu University, Zambia
dayakar.africa@gmail.com

Goga Nicolae
Department of Computer Sciences,
Technical University of Constructions,
Bucharest, Romania
n.goga@rug.nl

Goga Maria
Department of Educational Sciences,
Technical University of Constructions,
Bucharest, Romania
maria.goga@yahoo.com

***Abstract*- The usage of computers in academia has become an essential and integral part of the teaching learning process, and Zambia is no exception. This research paper investigates the usage of mathematical and statistical packages in educational institutes in Zambia at tertiary level. The result of the study shows that students are familiar with these packages, but availability is the concern. Lack of updated knowledge, non-participation in workshops and poor infrastructure are impediments to the low level usage. The mixed method approach was used in this study. A prototype tool was used to examine the ability of the students.**

I. INTRODUCTION

Computer knowledge is changing and developing. Hence, it is imperative to keep up to date with latest software packages. It is penetrating our daily lives in every moment and aspect. Education is not escaping from this penetration. Software is used more and more in educational programs, in all levels and domains. While for interactive programs this is quite easy, for mathematics it becomes more difficult. In this paper we concentrate on the use of software for mathematics in higher education.

Another important remark is that while in other continents there are many studies that discuss the impact which software has in education, there are few studies of this type in Africa. Major studies were done by the International Mathematical Union in 2009. According to the International Mathematical Union (IMU) 2009 Africa is a continent of contrasts. The African continent has more than 50 countries and they exhibit huge variability in size, population level, wealth and culture. These variations are also seen in educational development, specifically in mathematics and the usage of mathematics software. In this paper we address this identified gap by reporting a research done at the University of Zambia regarding the use of mathematics software in higher education.

This paper is organized as follows. Next section presents a relevant literature review. Section 3 presents the methodology used in this research. Section 4 describes the results while Section 5 draws the conclusions.

II. LITERATURE REVIEW

John K. Hunter (2011) from the University of California suggests that Mathematical software development is an uncommon area. This provides a good reason for the present study to determine the use and limitations of Mathematical and Statistical Software programs. Hunter tries to suggest that it is possible to create working user interfaces rapidly with little risk if one sticks to standard Graphical User Interface (GUI) components such as buttons, sliders, and data entry fields. He views mathematical software development as usually much harder than modern user interface software development, taking longer per line of code, involving much more debugging and being less predictable. UH Kortenkamp (2008) suggests that software packages have some significant impact on mathematics education. He farther says that the role of mathematics education is to device rules that lead to design principles for creating the artifact. For him computer graphics and mice became standard output and input devices. O'Leary (2000) underscores the need of mathematics as he said that the nineties really introduced new ways to do mathematics for everybody. Further he describes that extrapolating from the past ten year cycle of software innovation would lead to the expectation of a new quantum leap for the millennium. This quantum leap is not just faster software caused by new hardware.

There is evidence in literature that hardware will become better and better as Jean-Claude Trichet in International Symposium New Technologies and Monetary Policy November 2001 notes:

A. Application of the Mathematics Software:

"Mathematical Reasoning in Software Engineering Education" by Peter B. Henderson (2002) of Butler University suggests that Software is abstract and one can find many notations, symbols, abstractions and precision. He further suggests that mathematics is requisite for modeling software systems and many application domains (engineering, science, economics, etc.) are mathematically based. He adds that mathematical reasoning about software systems is essential.

Henderson (2002) quotes Devlin as saying:

> *What role does mathematics play in software engineering? Consider the following two statements. "Software practitioners do not need or use mathematics." "Software practitioners do need to think logically and precisely." An apparent contradiction since the reasoning underlying software engineering and mathematics are similar. Perhaps when a software practitioner says, "I don't use mathematics" this really means "I don't explicitly use mathematics." Practicing engineers don't explicitly use calculus on a daily basis, but they do implicitly use mathematical reasoning all the time.*

He suggests that Mathematical Software is applicable on Groups, Algorithms, Programming (GAP) - a system for computational discrete algebra, with particular emphasis on Computational Group Theory. Santa Marta (2002) School of Mathematics The Colombia University says that GAP provides a programming language, a library of thousands of functions implementing algebraic algorithms written in the GAP language as well as large data libraries of algebraic objects. GAP is used in research and teaching for studying groups and their representations, rings, vector spaces, algebras, combinatorial structures, and more.

While Marta is focusing on software in general, Santiago R. Simanca (2002) is specifically focusing on Maple:

> *Maple is a comprehensive computer system for advanced symbolic mathematics. It includes facilities for interactive algebra, calculus, discrete mathematics, graphics, numerical computation and many other areas of mathematics. It also provides a unique environment for rapid development of mathematical programs using its vast library of built-in functions and operations.*

K. Application of the Statistics Software

In System Software Solutions (S-CUBE, 2008) it says that the GAUSS Mathematical and Statistical System is a fast matrix programming language widely used by scientists, engineers, statisticians, biometricians, econometricians, and financial analysts. The GAUSS system is ideally designed for computational intensive tasks. R is 'GNUS', a freely available language and environment for statistical computing and graphics which provides a wide variety of statistical and graphical techniques: linear and nonlinear modeling, statistical tests, time series analysis, classification, clustering, etc.

The SAS (Statistical Analysis System) is a powerful programming language and a collection of ready-to-use programs called procedures. It has data access, data management, data presentation, and data analysis capabilities. Its roots are in academia, particularly statistical analysis, but SAS has broadened its applications to include dozens of tools from data entry products to a scalable performance data server product. The power of the SAS System comes from the integration of all of the SAS applications so they work together. Additionally, the SAS System is portable; you can easily move SAS applications across platforms. Finally, SAS has flexible user interfaces; this means if you would rather not learn the SAS programming language you can use the SAS user interface products for a menu-driven, task-oriented, and point-and-click environment. STATA is a complete, integrated statistical package for data analysis, data management, and graphics.

Sober (2002), the C++ Software developer-CICS Transaction Gateway(CTG-IBM) Software Corporation, say that the SAS covers a wide range of statistical techniques. STATA includes a variety of routines to analyze complex data and is a general purpose statistical package with good graphics capabilities and a graphic editor. Among the highlights of STATA are that it is relatively easy to learn for beginners. A fast and complete matrix programming language is an integral part of STATA.

L. Educational Achievements Made in the Area of Mathematics in Africa

Major studies were done by the International Mathematical Union in 2009. According to the International Mathematical Union (IMU) 2009 Africa is a continent of contrasts. The African continent has more than 50 countries and they exhibit huge variability in size, population level, wealth and culture. These variations are also seen in educational development, specifically in mathematics.

According to the IMU report the number of PhD holders in mathematics in South Africa, as of 2007, was 502; the number in nearby Tanzania – a country of comparable area and population was only 38. Similarly, of the population aged 10 to 18 years in

Mozambique and Burkina Faso, fewer than 10% attend secondary school, while in South Africa and Swaziland more than 50% of that age group attends secondary school. Another way to summarize the disparities in mathematical development is to look at the continent as a whole. We might take five different sections of "the mathematical geography."

1. North Africa is a relatively advanced region, partly ' due to governments' broad commitment to research and education at all levels and sustained support from nearby southern Europe.
2. Francophone Africa, reflecting its colonial past and continued support from the world-wide Francophone, is relatively advanced in secondary and tertiary mathematics, but its strength in pure mathematics is not matched by activities in applied fields. This limits the ability of students to see mathematics-related career opportunities beyond teaching.
3. The Anglophone countries of East Africa are generally weaker in pure mathematics, with most emphasis directed toward applied fields.
4. Some of the smaller countries in Southern Africa offer exciting models of collaboration, primarily Tanzania, Botswana, and Zimbabwe, as well as the other countries of the Southern African Development Community (SADC).
5. Some South African schools, notably private schools and formerly all white universities have achieved relative strength in mathematics.

One of the participants in the IMU study, John Templiton, F (2009), found that while African countries differ from one another in many features, they are broadly similar in the issues that concern international and national conditions that hinder mathematical development. From this report, it seems clear that these conditions are virtually the same throughout the continent.

Another participant in that study, Dr. Bregman (2009), says that teaching weaknesses begin in the primary schools, where mandatory education laws have caused enrolments to soar. Secondary school education suffers from a lack of teachers with mathematics training and low student participation. He estimates that only about 30% of the secondary school age group in sub-Saharan Africa attend junior secondary school and around 15-20% of the group attend senior secondary school.

Mahlol (2006), one of the contributors, says that overall the story of mathematical development in Africa is discouraging. Based on the outstanding achievements of some individuals and institutions, it is clear that no African country lacks talented potential mathematicians, but without a stronger educational structure at all levels, few of them are able to reach their potential.

Charles (2010), a professor in Human-Computer Interaction, describes the state of mathematics in Tanzania as poor, especially at the primary and secondary levels. He cites about statistics of the National Examination Council of Tanzania (NECTA) showing that in the 2007 national examinations, only 17.4% of candidates passed mathematics. He also attributed this phenomenon to the insufficient number of competent mathematics teachers and partly to a tumultuous history since the country- was a German colony. In another study related to use and limitations of mathematical software conducted in South Africa, Moila (2006) investigated the use of educational technology in mathematics teaching and learning in Phusela Secondary School in Limpopo Province, South Africa. This was a case study, carried out in one week. The method used in the investigation was mixed method approach, and one case was investigated and evaluated. The study focused on the relationship between uses of Integration of Information and Communication Technology (ICT). It was found out that software tools are not used in mathematics teaching and learning in Phusela School.

The contributing factors established in that study were:

- The absence of guidelines (policy) on the use of ICT tools in the whole school.
- Inadequate training of educators on ICT tools usage in teaching and learning, and
- Lack of relevant ICT resources for rural schools.

M. Studies Done at Tertiary Level

According to Shafika (2007) ICT tools and Policy development in South Africa, "Since 1994 the South African education landscape has undergone major transformation in Education system". This study further adds, "ICTs in education laid the basis for the e-Education and White Paper adoption in 2004"

According to Cerniewicz, Ravjee, et al (2006) in the Survey of ICT and Education in Africa South Africa's 24 state-funded tertiary institutions and private institutions in varying degrees, have policies, strategies, or programs in place related to ICTs. They also pointed out that the use of technology is defined more by the nature of the institution than by policies, referencing the University of Stellenbosch's E-Campus Strategy incorporating all e-learning business and the University of Pretoria's Telematic Learning and Education Innovation Strategic Plan 2002-2005.

According to the African Leadership in ICT 2011 "E-education is aimed at connecting learners and teachers to information, ideas and to one another via effective combinations of pedagogy and technology in support of educational reform".

The African Virtual Open Initiatives and Resources (AVOIR) Project, initiated by the University of the Western Cape (UWC), is a collaborative effort among several African higher education institutions contributed to the development of Africa through Free and Open Source Software (FOSS) development activities. It has taken an existing e-learning platform application, Knowledge Environment for Web-based Learning (KEWL) that was developed at UWC, and rebuilt it to run on a FOSS platform. Women's Net services include technology planning especially in creating awareness of free and open source software solutions, training on the strategic use of technology to meet the needs of women. Women's Net also has a dedicated technology training centre with 18 computers networked on an open source platform.

These studies that were done at the University level suggest that African Universities are exposing their students to software development activities, but they have not focused on the use and limitations of the Mathematical and Statistical Software programs, which therefore justifies the purpose of the present study.

III. METHODOLOGY

The mixed method approach was used in this study. In the department of Mathematics in the University of Zambia there are twenty five students in first year, eighteen students in second year, twelve students in third year and six students in fourth year. Among them there are eight female and fifty two male students bringing the total enrolment to sixty for the academic year 2012-2013. While the researcher intended to do purposive sampling of the sixty students, at the time of appointment only forty two (seven female and thirty five male students) availed themselves, hence the data was collected from forty two students and three lecturers (two male and one female). Prototype tool was used to observe the ability of students in usage of software.

IV. ANALYSIS

This section presents the analyses of data and discussion of the findings and their interpretations. This is presented by using descriptive statistics such as frequencies, percentages, and totals. The Statistical Package of Social Sciences (SPSS) was used to determine the frequencies and percentages of the data. The data obtained from the questionnaires were presented in text, tabular, and graphic forms. The responses from the personal interviews were presented in tabular and text forms. The analyses done were for the forty two students and three lecturers of the department of Mathematics in the University of Zambia. The data analysis consists of part A and part B. Part A deals with the data analysis pertaining to questionnaires and interviews. Part B deals with source code. The researcher developed and distributed forty two (42) questionnaires for the students and three for the lecturers. Eight interviews with students and three with lecturers were conducted.

The number of students is forty two (93.3%) and number of lecturers three (6.7%). With regards to the usage of mathematical and statistical programs in the University of Zambia the study shows that only six (13.3%) respondents agree with the usage of Matlab, whereas three (6.7%) respondents agree with the usage of Scilab. It also shows that 28 (62.2%) respondents agree the use of Excel program. When it comes to Encarta only 22 (48.9%) agree to use this program. In case of SPSS, there are 18 (40%) respondents who agree to use this program. By this it appears that most respondents use EXCEL and SPSS as compared to other software. However four of the respondents who responded to all the questions did not indicate the software program they used.

When respondents were asked about frequency of the usage of various mathematical and statistical programs, out of forty five respondents thirty four (75.5%) said that they use these programs very often. Although there were six students who disagreed and five students who strongly disagreed, hundred percent of the lecturers said that they use mathematical and statistical packages very often. It is interesting to note this, the aspect of lecturers active participation is very encouraging, which inspires the students to use also the software. This study also revealed that about half of the respondents (55%) said that they do not have technical support. Which means that probably there was a lack of manpower or appropriate mechanisms to handle technical difficulties while using mathematical and statistical packages. Forty five of them expressed that they were able to use technical support to solve their day to day technical difficulties.

The researcher further investigated about the helpfulness of the mathematical and statistical packages in their studies: 86.7% of the students agreed that these packages helped them in their studies whereas fewer of them (13.3%) disagreed. For further clarification the responses were cross tabulated between gender and category. Thirty male students and six female students agreed positively however five male and one female student disagreed whereas all three lecturers expressed that these mathematical and statistical packages help them very much in their teaching/learning process.

Mathematical and statistical packages help not only in calculations but they also help in many other aspects of the teaching and learning process like grading and preparing presentations etc. When the

respondents were asked about this issue 22.2% of them strongly agreed, 46.7% of them agreed whereas 22.2% disagreed, 6.7% strongly disagreed which indicates that majority of them using mathematical and statistical packages more than one purposes.

Some respondents looked at mathematical and statistical packages as not very useful as they believe these packages make students lazy. Interestingly this group is very small in number (17.8%) whereas majority (77.7%) opposed this view. Two of them did not respond.

Majority of the respondents seem to be traditional as they expressed that it is better to solve the mathematical and statistical problems manually instead of using software programs. The reasons for their specific choices may vary. 42.3% of them agreed that it is better to calculate by hand whereas 57.8% disagreed. This indicates that there should be some awareness program to educate people in eradicating some of the misconceptions that they may have in using mathematical and statistical programs. This data also indicates that students are having misconceptions. Hence it is suggested that teachers need to take some initiative in educating the students to lead them in right direction as it is very important to use statistical packages due to time, speed and accuracy considerations.

V. CONCLUSIONS

The study found that lecturers are able to be updated with the latest mathematical and statistical packages, which is a positive sign for the academic growth of the student body, as long as the student fraternity is able to get updated as well. There is a crucial need to organize special workshops and seminars so that students will be accepted to be better.

This paper upholds that there should be equal opportunities for male and female students. However the researcher observed that the male students are more active in using these software packages than female students. This matter alarms any researcher so it can be corrected by providing incentives and scholarships, regular seminars and workshops for female students which can motivate them to have active participation in using these software packages. Since Zambia has little over 50% percentage of women population ratio, there may be a direct impact on the younger generation when women are empowered with these packages.

This study also found that students and teachers have an optimistic attitude towards these packages, since 99% of the people whom the researcher interviewed agreed with the active role of these packages in real life situations. The study revealed that students are in favor of these software packages although the availability of mathematical and statistical software packages is a concern. This phenomenon suggest that there needs to be ample resources to meet the needs of the students with regards to availability of mathematical and statistical packages. It also found that lecturers are ahead of students in using software packages. Since the lecturers updated themselves in recent developments of software packages they also ought to train the students to research at the advanced level in using mathematical and statistical software packages. Male students use more available software packages than their female counterparts. Equal opportunities should be ensured without any gender discrimination. It is suggested that there be some incentives and special provisions for female students to encourage them to be active in learning and using mathematical and statistical packages. Prototype tools, as the ones worked out by the author of this thesis, can be used easily by lecturers and students. Most people are aware of such packages and appreciate using them in teaching and learning mathematics as they minimize difficulty and enhance their conceptual understanding.

Based on the study's findings it is concluded that mathematical and statistical packages contribute tremendously in the teaching and learning process of mathematics. There are inadequate facilities with regards to infrastructure and availability of this software. This has become a hindrance in using statistical and mathematical packages regardless of gender. Students and teachers have almost similar opinions about using mathematical and statistical packages.

Based on these conclusions recommendations were made to the tertiary institutions to enhance the use of mathematical and statistical packages. Institutions must provide hands-on experience to lecturers and students to familiarize with statistical and mathematical packages. It is recommended that tertiary institutions should increase the infrastructure to meet the needs of students and lecturers. Institutions should take further steps to motivate students and lecturers to use mathematical and statistical packages. Institutions should take initiatives in making various software packages available. Institutions and academic departments are advised to conduct seminars and workshops regularly to update students and lecturers with advanced techniques and software packages. Providing incentives to lecturers is essential to encourage them to use these packages. Extending lab hours and instruction is another way of encouraging students to use mathematical and statistical packages. Continued guidance and feedback will help students to use these packages effectively. Industry -University partnerships will provide ample experience and opportunities for students.

Summarizing, the researcher came to the following conclusions by doing this study

1. There is overwhelming acceptance of these software packages by students and lecturers.
2. There is a real need of workshop and seminars to familiarize there software packages.

3. There is a genuine concern for the scarcity of these software packages.
4. Some programs which are available in the market are not within the reach of the students parching ability.
5. The researcher is optimistic that these software packages are for the overall benefit of students as well as teachers.

REFERENCES

[1] Lori A. Clarke, (2003) 25th Int. Conference on Software Engineering (ICSE) *Extended version in IEEE Transactions on Software Engineering*. p12-14 Retrieved from http://www.icse-conferences.org/2003/home/ICSE2003

[2] Alant, E. & Mophosho, M.(2003) Early Childhood Intervention: Web-based Training.p5-15 Retrieved from http://web.up.ac.za/sitefiles/file/46/9742/Annual Bridges.org(2003); Digital Divide Assessment of the city of Cape Town. p8. Retrieved from http://www.che.ac.za/documents/d000127/11-ICTs_HE_Landscape_Jul2006.

[3] C. Abts, A.W. Brown, S. Chulani, B.K. Clark, E. Horowitz, R. Madachy, D. Reifer, and B. Steece, (2000) *Software Cost Estimation with COCOMO II*, Prentice Hall.p23-25 Retrieved from http://csse.usc.edu/csse/about/people/faculties/BarryBoehm.html

[4] Capers Jones (1998): *Sizing Up Software*, Scientific American. p 4-16. Retrieved from http://insights.cermacademy.com/2012

[5] C Dreyer (2013) *Learning and Teaching through information and Communication Technologies*. Pretoria: Department of Education. p5-10 Retrieved from L http://web.up.ac.za/default.aspipk

[6] Cerniewicz, L., N. Ravjee and N. Mlitwa. N.(2013) Higher Education Monitor: Information and

a. Communication Technologies and South African Higher Education. p16-23 Retrieved from http//www.infodev.org/higher education

[7] David Joyner U. S. Naval Academy, Annapolis. p 12. Retrieved from http://cs.nyu.edu/artg/systemquality/Fall2002/readings

[8] Devlin, K., "*The Real Reason Why Software Engineers Need Math*," Communications of the ACM, Oct. 2001 (44:10) p. 21-22. Retrieved from http://www.cs.geneseo.edu/~baldwin/math-thinking/MathInSE

[9] Henderson, P. B. (2003).Mathematical Reasoning in Software Engineering Education. Vol 46, No. 9. Butler University

[10] Hunter. J. K, (2011). Measure Theory. Department of Mathematics, University of California, Davis.

[11] H. Sackman,(1972). *Planning Community Information Utilities*, AFIPS Press, Higher Education. p34 .Retrieved from http://www.bcf.usc.edu/~fbar/Publications

[12] J.R. Brown, H. Kaspar, M. Lipow, G. McLeod, and M. Merritt, (1978) *Characteristics of Software Quality*, North Holland. Retrieved from http://www.empress- itea.org/publications

[13] James, T.(2006) ed. Women in the Information and Communication Technology Sector in South Africa. Retrieved from http://source.cet.uct.ac.za/svn/cet_website/trunk/files

[14] National Academy Press, (1996) *Software Policies for the Department of Defense.* Retrieved from http://www.nap.edu/openbook.php

[15] R. Turner, Addison Wesley, (2008) *Balancing Agility and Discipline*: *A Guide for the Perplexed*, Transformation. South African Journal of Higher Education. p185-199. Retrieved from http://cit.mak.ac.ug

A Modular-Arithmetic-Based Encryption Scheme

Odule, Tola John
Department of Mathematical Sciences
Olabisi Onabanjo University, Nigeria
tola.odule@oouagoiwoye.edu.ng

Goga, Nicholae
Department of Computer Science
Babcock University, Nigeria.
n.goga@rug..nl

Okolie, Samuel O.
Department of Computer Science
Babcock University, Nigeria
samuelokolie2003@yahoo.com

***Abstract*---**This paper considers a scenario in which a sender who holds a *k*-bit to *k*-bit trapdoor permutation *f* wants to transmit a message *x* to a receiver who holds the inverse permutation f^{-1}; with the condition that encryption should require just one computation of *f*, decryption should require just one computation of f^{-1}, the length of the enciphered text should be precisely *k* and the length *n* of the text *x* that can be encrypted is close to *k*. Our scheme takes the encryption of *x* to be $f(r_x)$, which is a simple probabilistic encoding of *x*. Assuming an ideal hash function and an arbitrary trapdoor permutation, we describe and prove secure a simple invertible enmesh scheme that is bit-optimal in that the length of the string *x* that can be encrypted by $f(r_x$ is almost *k*. Our scheme achieves semantic security, which implies adaptive chosen-ciphertext security and non-malleability.

Keywords: **Asymmetric encryption, provable security, trapdoor permutation, semantic security**

I. INTRODUCTION

Public-key encryption has been around for over thirty years. In its basic form, it is well understood: a public key allows for encryption, while an associated private (secret) key performs decryption. The complication lies in ensuring safe communication over an insecure channel in the presence of a malevolent eavesdropper, without the problem of key distribution and exchange, in a heterogeneous community of users.

In our setup we consider a sender who holds a k-bit to k-bit trapdoor permutation f and wants to transmit a message x to a receiver who holds the inverse permutation f^{-1} We concentrate on the case which arises most often in cryptographic practice, where $n = |x|$ is at least a little smaller than k.

What practitioners want is the following encryption should require just one computation of f decryption should require just one computation of f^{-1} the length of the enciphered text should be precisely k and the length n of the text x that can be encrypted is close to k Since heuristic schemes achieving these conditions exist [1, 2], if provable security is provided at the cost of violating any of these conditions, for instance two applications of f to encrypt message length $n+k$ rather than k, practitioners will prefer the heuristic constructions. Thus to successfully impact practice one must provide provably-secure schemes which meet the above constraints.

The heuristic schemes invariably take the following form: one probabilistically, invertibly incorporates x into a string r_x and then takes the encryption of x to be $f(r_x)$[1]. We call such a process an *invertible enmesh scheme*. We will take as our goal the construction of a provably *invertible enmesh schemes* which allows n to be close to k.

Assuming an ideal hash function and an arbitrary trapdoor permutation, we describe and prove secure a simple *invertible enmesh scheme* that is bit-optimal in that the length of the string x that can be encrypted by $f(r_x)$ is almost k. Our scheme achieves semantic security [3]. This notion is very strong, and in particular, implies ambitious goals like chosen-ciphertext security and non-malleability [6] in the ideal-hash model which we assume.

A. The basic scheme

Recall k is the security parameter, f mapping k-bits to k-bits is the trapdoor permutation. Let k_0 be chosen such that the adversary's running time is significantly smaller than 2^{k_0} steps. We fix the length of the message to encrypt as $n = k - k_0$ bits (shorter messages can be suitably padded to this length). The scheme makes use of a *generator* $G: \{0,1\}^{k_0} \to \{0,1\}^n$ and a *hash function*

[1]It is well-known that a naive enmesh like $r_x = x$ is no good; besides the usual deficiencies of any deterministic encryption, f being a trapdoor permutation does not mean that $f(x)$ conceals all the interesting properties of x.Indeed it was exactly such considerations that helped inspire ideas like semantic security [3] and hardcore bits [4,5].

$H: \{0,1\}^n \to \{0,1\}^{k_0}$. To encrypt $x \in \{0,1\}^n$ choose a random k_0 bit r and set

$$\mathcal{E}^{G,H}(x) = f(x \oplus G(r) \,\|\, r \oplus H(x \oplus G(r)))$$

Here $\|$ denotes concatenation. The decryption function $\mathcal{D}^{G,H}$ is defined in the obvious way, and the pair $(\mathcal{E}, \mathcal{D})$ constitutes what we call the *basic* scheme.

We prove security under the assumption that G, H are *ideal*. This means G is a random function of $\{0,1\}^{k_0} \to \{0,1\}^n$ and H is a random function of $\{0,1\}^n \to \{0,1\}^{k_0}$. The formal statement of our result is in *Theorem 4.1*. It says that if f is a trapdoor permutation and G, H are ideal then the basic scheme achieves the notion of semantic security [3], appropriately adjusted to take account of the presence of G, H.

In practice, G and H are best derived from some standard cryptographic hash function. For example, they can be derived from the compression function of the Secure Hash Algorithm [7] following the methods described in [8].

B. Computational Efficiency of the Proposed Model

The function f can be set to any candidate trapdoor permutation such as RSA [9] or modular squaring [10, 11]. In such a case the time for computing G and H is negligible compared to the time for computing f, f^{-1}. Thus complexity is discussed only in terms of f, f^{-1} computations. In this light our basic encryption scheme requires just a single application of f to encrypt, a single application of f^{-1} to decrypt, and the length of the ciphertext is k, as long as $k \geq n + k_0$. Our scheme requires a single application of f to encrypt, a single application of f^{-1} to decrypt, and the length of the ciphertext is still k, as long as $k \geq n + k_0 + k_1$.

A concrete instantiation of our scheme, using RSA for f and getting G, H from the Secure Hash Algorithm [7], is given in Section 4.1.

C. The ideal hash function paradigm

As we indicated above, when proving security we take G, H to be random, and when we want a concrete scheme. G, H are instantiated by primitives derived from a cryptographic hash function. In this regard we are following the paradigm of [8] who argue that even though results which assume an ideal hash function do not provide provable security with respect to the standard model of computation, assuming an ideal hash function and doing proofs with respect to it provides much greater assurance benefit than purely *ad-hoc*, protocol design.

D. Assessment of the 'Exact security' of our Scheme

We want our results to be meaningful for practice. In particular, this means we should be able to say meaningful things about the security of our schemes for specific values of the security parameter (e.g., $k = 512$). This demands not only that we avoid asymptotics and address security *exactly*, but also that we strive for security reductions which are as efficient as possible.[2]

Thus the theorem proving the security of our basic scheme, as in [12], quantifies the resources and success probability of a potential adversary let her run for time t make q_{gen} queries of G and q_{hash} queries of H, and suppose she could *break* the encryption with advantage ε. It then provides an algorithm M and numbers t', ε' such that M inverts the underlying trapdoor permutation f in time t' with probability ε'. The strength of the result is in the values of t', ε' which are specified as functions of t, q_{gen}, q_{hash}, ε, and the underlying scheme parameters k, k_0, n $(k = k_0 + n)$. Now a user with some idea of the assumed strength of a particular f such as RSA on 512 bits can get an idea of the resources necessary to break our encryption scheme.

[2] Exact security is not new; previous works which address it explicitly include [13, 14, 15 16, 17, 18]. Moreover, although it is true that most theoretical works only provide asymptotic security guarantees of the form *the success probability of a polynomially bounded adversary is negligible* (everything measured as a function of the security parameter), the exact security can be derived from examination of the proof. (However, a lack of concern with the exactness means that in many cases the reductions are very inefficient, and the results are not useful for practice).

II. NOTATIONS AND CONVENTIONS

A. Probabilistic algorithms

We hereby use the notation of [19]. If A is a probabilistic algorithm then $A(x, y, \ldots)$ refers to the probability space which to the string σ assigns the probability that A on inputs $x, y, \ldots$ outputs σ. If S is a probability space we denote its support, the set of elements of positive probability, by $[S]$. When S is a probability space, $x \leftarrow S$ denotes selecting a random sample from S. We use $x, y \leftarrow S$ as shorthand for $x \leftarrow S$; $y \leftarrow S$. For probability spaces $S, T, \ldots,$ the notation $\Pr[x \leftarrow S; y \leftarrow T; \ldots, p(x, y, \ldots)]$ denotes the probability that the predicate $p(x, y, \ldots)$ is true after the ordered execution of the algorithms $x \leftarrow S$ $y \leftarrow T$, etc. PPT is short for *probabilistic-polynomial time*.

In evaluating the complexity of oracle machines we adopt the usual convention that all oracle queries receive their answer in unit time

B. Random oracles

We will be discussing schemes which use functions G, H chosen at random from appropriate spaces (the input and output lengths for G and H depend on parameters of the scheme). When stating definitions it is convenient to not have to worry about exactly what these spaces may be and just write $G, H \leftarrow \Omega$, the latter being defined as the set of all maps from the set $\{0,1\}^*$ of finite strings to the set $\{0,1\}^\infty$ of infinite strings. The notation should be interpreted as appropriate to the context--for example, if the scheme says G maps $\{0,1\}^a$ to $\{0,1\}^b$ then we can interpret $G \leftarrow \Omega$ as meaning we choose G from Ω at random, restrict the domain to $\{0,1\}^a$ and drop all but the first b bits of output.

C. Trapdoor permutations and their security

Our encryption scheme requires a *trapdoor permutation generator*. This is a PPT algorithm $\mathcal{F}$ such that $\mathcal{F}(1^k)$ outputs a pair of deterministic algorithms (f, f^{-1}) specifying a permutation and its inverse on $\{0,1\}^k$.

We associate to $\mathcal{F}$ an evaluation time $T_{\mathcal{F}}(\cdot)$: for all k, all $(f, f^{-1}) \in [\mathcal{F}(1^k)]$ and all $w \in \{0,1\}^k$, the time to compute $f(w)$, given f and w, is $T_{\mathcal{F}}(k)$. Note the evaluation time depends on the setting; for example on whether or not there is hardware available to compute f.

We will be interested in two attributes of a possibly non-uniform algorithm M trying to invert $\mathcal{F}(1^k)$-distributed permutations, namely its running time and its success probability.

Definition 1 Let $\mathcal{F}$ be a trapdoor permutation generator. We say that algorithm M succeeds in (t, ε)-inverting $\mathcal{F}(1^k)$ if

$$\Pr[(f, f^{-1}) \leftarrow \mathcal{F}(1^k);\ w \leftarrow \{0,1\}^k;\ y \leftarrow f(w);\ M(f, y) = w] \geq \varepsilon$$

, and, moreover, in the experiment above, M runs in at most t steps.

RSA [9] is a good candidate function as a secure trapdoor permutation.

III. SEMANTICALLY SECURE ENCRYPTION

We extend the definition of semantic security [3] to the random oracle model in a way which enables us to discuss exact security.

A. Encryption schemes

An asymmetric encryption scheme is specified by a probabilistic generator, $\mathcal{G}$, and an associated plaintext-length function, n. On input 1^k the generator $\mathcal{G}$ outputs a pair of algorithms $(\mathcal{E}, \mathcal{D})$ the first of which is probabilistic. Each of these algorithms has oracle-access to two functions, one called G and one called H. A user i runs G to get $(\mathcal{E}, \mathcal{D})$ and makes the former public while keeping the latter secret. To encrypt message $x \in \{0,1\}^{n(k)}$ using functions G, H anyone can compute

$y = \mathcal{E}^{G,H}(x)$ and send it to i. To decrypt ciphertext y user i computes $x = \mathcal{D}^{G,H}(y)$. We require $\mathcal{D}^{G,H}(y) = x$ for all $y \in [\mathcal{E}^{G,H}(x)]$. We further demand that $\mathcal{D}^{G,H}(y) = *$ if there is no x such that $y \in [\mathcal{E}^{G,H}(x)]$.

An adversary is a possibly non-uniform algorithm A with access to oracles G, H. We assume without loss of generality (w.l.o.g) that an adversary makes no particular G-query more than once and no particular H-query more than once. For simplicity we assume that the number of G-queries and H-queries that an adversary makes don't depend on its coin tosses but only, say, on the length of its input.

B. Semantic security

The following definition will be used to discuss *exact* security. It captures the notion of semantic security [3] appropriately lifted to take into account the presence of G, H.

We consider an adversary who runs in two stages. In the **find**-stage it is given an encryption algorithm $\mathcal{E}$ and outputs a pair x_0, x_1 of messages. It also outputs a string c which could record, for example, its history and its inputs. Now we pick at random either x_0 or x_1, the choice made according to a bit b, and encrypt it under $\mathcal{E}$ to get y. In the **guess**-stage we provide A the output x_0, x_1, c of the previous stage, and y, and we ask it to guess b. We assume w.l.o.g that $\mathcal{E}$ is included in c so that we don't need to explicitly provide it again. Since even the algorithm which always outputs a fixed bit will be right half of the time, we measure how well A is doing by $1/2$ less than the fraction of time that A correctly predicts b. We call twice this quantity the advantage which A has in predicting b. Multiplying by two makes the advantage fall in the range $[0, 1]$ (0 for a worthless prediction and 1 for an always correct one), instead of $[0, 0.5]$.

Definition 2 Let $\mathcal{G}$ be a generator for an encryption scheme having plaintext-length function $n(\cdot)$. An adversary A is said to succeed in $(t, q_{gen}, q_{hash}, \varepsilon)$-breaking $\mathcal{G}(1^k)$ if

$$\varepsilon \le 2 \cdot \Pr\left[\begin{array}{c}(\mathcal{E}, \mathcal{D}) \leftarrow \mathcal{G}(1^k);\ G, H \leftarrow \Omega;\ (x_0, x_1, c) \leftarrow A^{G,H}(\mathcal{E}, \textbf{find}) \\ b \leftarrow \{0,1\};\ \ y \leftarrow \mathcal{E}^{G,H}(x_b);\ \ A^{G,H}(y, x_0, x_1, c) = b\end{array}\right] - 1$$

And, moreover, in the experiment above, A runs for at most t steps, makes at most q_{gen} queries to G and makes at most q_{hash} queries to H.

Note that t is the total running time, that is, the sum of the times in the two stages. Similarly q_{gen}, q_{hash} are the total number of G and H queries, respectively.

IV. THE BASIC ENCRYPTION SCHEME

Let $\mathcal{F}$ be a trapdoor permutation generator and $k_0(\cdot)$ a positive integer valued function such that $k_0(k) < k$ for all $k \ge 1$. The basic scheme $\mathcal{G}$ with parameters $\mathcal{F}$ and $k_0(\cdot)$ has an associated plaintext-length function of $n(k) = k - k_0(k)$. On input 1^k, the generator $\mathcal{G}$ runs $\mathcal{F}(1^k)$ to obtain (f, f^{-1}). Then it outputs the pair of algorithms $(\mathcal{E}, \mathcal{D})$ determined as follows

1. On input x of length $n = n(k)$, algorithm $\mathcal{E}$ selects a random r of length $k_0 = k_0(k)$ It sets $s = x \oplus G(r)$ and $t = r \oplus H(s)$. It sets $w = s \| t$ and returns $y = f(w)$.
2. On input y of length k, algorithm $\mathcal{D}$ computes $w = f^{-1}(y)$. Then it sets s to the first n bits of w and t to the last k_0-bits of w. It sets $r = t \oplus H(s)$, and returns the string $x = s \oplus G(r)$.

The oracles G and H which $\mathcal{E}$ and $\mathcal{D}$ reference above have input-output lengths of $G: \{0,1\}^{k_0} \to \{0,1\}^n$ and $H: \{0,1\}^n \to \{0,1\}^{k_0}$. We use the encoding of

f as the encoding of $\mathcal{E}$ and the encoding of f^{-1} as the encoding of $\mathcal{D}$.

The intuition behind the semantic security of this scheme is as follows. We wish to guarantee that the adversary, given a point y in the range of f, must recover the complete preimage $w = r_x$ of y if she is to say anything meaningful about x itself. Well, if the adversary does not recover all of the first n bits of the preimage, s, then she will have no idea about the value $H(s)$ which is its hash; a failure to know anything about $H(s)$ implies a failure to know anything about $r = H(s) \oplus t$,where t is the last k_0 bits of w, and therefore $G(r)$ and therefore $x = G(r) \oplus s$ itself. Now, assuming the adversary does recover s, a failure to completely recover t will again mean that the adversary fails to completely recover r, and, in the lack of complete knowledge about r, $x \oplus G(r)$ is uniformly distributed and so again the adversary can know nothing about x.

Yet the above discussion masks some subtleties and a formal proof of security is more complex than it might appear. This is particularly the case when one is interested, as we are here, in achieving the best possible exact security.

The following theorem says that if there is an adversary A who is able to break the encryption scheme with some success probability, then there is an algorithm M which can invert the underlying trapdoor permutation with comparable success probability and in comparable time. This implies that if the trapdoor permutations can't be inverted in reasonable time (which is the implicit assumption) then our scheme is secure. But the theorem says more; it specifies exactly how the resources and success of M relate to those of A and to the underlying scheme parameters k, n, k_0 $(k = n + k_0)$.

The inverting algorithm M can by obtained from A in a *uniform* way; the theorem says there is a *universal* oracle machine U such that M can be implemented by U with oracle access to A. It is important for practice that the *description* of U is *small*; this is not made explicit in the theorem but is clear from the proof. The constant λ depends only on details of the underlying model of computation. We write n, k_0 for $n(k), k_0(k)$, respectively, when, as below, k is understood.

Theorem 4.0 Let $\mathcal{G}$ be the basic encryption scheme with parameters $\mathcal{F}$, k_0 and let n be the associated plaintext length. Then there exists an oracle machine U and a constant λ such that for each integer k the following is true. Suppose A succeeds in $(t, q_{gen}, q_{hash}, \varepsilon)$-breaking $\mathcal{G}(1^k)$. Then $M = U^A$ succeeds in (t', ε')-inverting $\mathcal{F}(1^k)$ where

$$t' = t + q_{gen} \cdot q_{hash} \cdot \left(T_{\mathcal{F}}(k) + \lambda k\right)$$

$$\varepsilon' = \varepsilon \cdot \left(1 - q_{gen} 2^{-k_0} - q_{hash} 2^{-n}\right) - q_{gen} 2^{-k+1}.$$

We omit the proof of this *Theorem* for the sake of brevity.

For reasonable values of k (e.g., $k \geq 512$) it will be the case that $k > n >> k_0$. Thus for reasonable values of q_{gen}, q_{hash} we'll have $\varepsilon' \approx \varepsilon \cdot \left(1 - q_{gen} 2^{-k_0}\right)$. Thus the success probability ε' achieved here is good in the sense that it is only slightly less than ε, and close to optimal. Note also that the expression for ε' indicates that A will do best by favouring G-oracle queries over H-oracle queries.

The dominant factor in the time t' taken by the inverting algorithm to compute $f^{-1}(y)$ is the time to do $q_{gen} \cdot q_{hash}$ computations of the underlying f. An interesting open question is to find a scheme under which the number of computation of f is linear in $q_{gen} + q_{hash}$ while retaining a value of ε' similar to ours.

A. A Prototype Model of our Scheme

We provide here a concrete implementation of our encryption scheme, omitting only certain minor details. We use RSA as the trapdoor permutation and construct the functions G, H out of the revised NIST Secure Hash Algorithm [7], although other hash algorithms such as MD5 [20] would do as well.

Let f be the RSA function [9], so $f(x) = x^e \bmod N$ is specified by (e, N) where N is the k-bit product of two large primes and

$(e, \varphi(N)) = 1$. We demand $k \geq 512$ bits, larger values are recommended. Our scheme will allow the encryption of any string msg whose length is at most $k - 320$ bits; thus the minimal permitted security parameter allows 192 bits (e.g., three 192-bit keys to be encrypted). Let $D = \{1 \leq i < N; \quad \gcd(i, N) = 1\} \subseteq \{0, 1\}^k$ D be the set of valid domain points for f.

Our probabilistic encryption scheme depends on the message msg to encrypt, an arbitrary-length string *rand_coins*, the security parameter k, the function f and a predicate $IN\ D(x)$ which should return t**rue** if and only if $x \in D$. Our scheme further uses a 32-bit string key_data, whose use we do not specify here, and a string $desc$ which provides a complete description of the function f; that is, it says "This is RSA using N and e" encoded according to conventions not specified here.

We denote by $\mathrm{SHA}_\sigma(x)$ the 160-bit result of SHA (Secure Hash Algorithm applied to x, except that the 160-bit *starting value* in the algorithm description is taken to be $ABCDE = \sigma$.. Let $\mathrm{SHA}_\sigma^\ell(x)$ denote the first ℓ-bits of $\mathrm{SHA}_\sigma(x)$. Fix the notation $\langle i \rangle$ for i encoded as a binary 32-bit word. We define the function $H_\sigma^\ell(x)$ for string x, number ℓ and 160-bit σ to be the ℓ-bit prefix of

$$\mathrm{SHA}_\sigma^{80}(\langle 0 \rangle \cdot x) \| \ \mathrm{SHA}_\sigma^{80}(\langle 1 \rangle \cdot x) \| \ \mathrm{SHA}_\sigma^{80}(\langle 2 \rangle \cdot x) \| \cdots$$

Let K_0 be a fixed, randomly-chosen 160-bit string, which we do not specify here.

Our scheme is depicted in Figure 1. Basically, we augment the string msg which we want to encrypt by tacking on a word to indicate its length; including $k_1 = 128$ bits of redundancy; incorporating a 32-bit field key_data whose use we do not specify, and adding enough additional padding to fill out the length of the string we have made to $k - 128$ bits. The resulting string x now plays the same role as the x of our basic scheme, and a separate 128-bit r is then used to encrypt it.

ENCRYPT (msg ,*rand_coins*)

$\sigma = \mathrm{SHA}_{K_0}(desc);$

$\sigma_1 = \mathrm{SHA}_\sigma(\langle 1 \rangle);$

$\sigma_2 = \mathrm{SHA}_\sigma(\langle 2 \rangle);$

$\sigma_3 = \mathrm{SHA}_\sigma(\langle 3 \rangle);$

$i \leftarrow 0;$

repeat

$r \leftarrow H_{\sigma_1}^{128}(\langle i \rangle \| \ rand\_coins)$;

$x \leftarrow key\_data \| \langle |msg| \rangle \| 0^{128} \| 0^{k-320-|msg|} \| msg$;

$\bar{x} \leftarrow x \oplus H_{\sigma_2}^{|x|}(r)$;

$\bar{r} \leftarrow r \oplus H_{\sigma_3}^{128}(\bar{x})$;

$r_x = \bar{x} \| \bar{r};$

$i \leftarrow i + 1;$

until in $\mathrm{D}(r_x);$

return $f(r_x);$

Figure 1: A sample implementation of our encryption scheme.

V. CONCLUSION

We comment that in the concrete scheme shown in Figure 1 we have elected to make our generator and hash function sensitive both to our scheme itself, via K_0 and to the particular function f, via $desc$. Such *key separation* is a generally-useful heuristic to help ensure that when the same key is used in multiple, separately-secure, algorithms that the internals of these algorithms do not interact in such a way as to jointly compromise security. The use of *key variants* σ_1, σ_2 and σ_3 is motivated similarly. Our choice to only use half the bits of SHA has to do with a general deficiency in the use of SHA-like hash functions to implement random oracles.

REFERENCES

[1] [1] RSA Data Security, Inc., "PKCS #1: RSA Encryption Standard," June 1991.

[2] [2] M. Bellare and P. Rogaway, "Random oracles are practical: a paradigm for designing efficient protocols." Proceedings of the First Annual Conference on Computer and Communications Security, ACM. 1993.

[3] [3] S. Goldwasser and S. Micali, "Probabilistic Encryption," Journal of Computer and System Sciences **28**, 270-299, April 1984.

[4] [4] M. Blum and S. Micali, "How to generate cryptographically strong sequences of pseudo- random bits," SIAM Journal on Computing **13** (4), 850-864, November 1984.

[5] [5] A. Yao, "Theory and applications of trapdoor functions," Proceedings of the 23rd Symposium on Foundations of Computer Science, IEEE, 1982.

[6] [6] D. Dolev, C. Dwork and M. Naor, "Non-malleable cryptography," Proceedings of the 23rd Annual Symposium on Theory of Computing, ACM,1991.

[7] [7] National Institute of Standards, FIPS Publication 180, "Secure Hash Standard," 1993.

[8] [8] D. Johnson, A. Lee, W. Martin, S. Matyas and J. Wilkins, "Hybrid key distribution scheme giving key record recovery," IBM Technical Disclosure Bulletin, **37** (2A), 5-16,February 1994.

[9] [9] R. Rivest, A. Shamir and L. Adleman, "A method for obtaining digital signatures and public key cryptosystems," CACM 21 (1978).

[10] [10] M. Rabin, "Digitalized signatures and public-key functions as intractable as factorization," MIT Laboratory for Computer Science TR 212, January 1979.

[11] [11] L. Blum, M. Blum and M. Shub, "A Simple Unpredictable Pseudo-Random Number Generator," SIAM Journal on Computing **15**(2), 364-383, May 1986.

[12] [12] Odule, T.J. (2007): Incremental Cryptography and Security of Public Hash Functions. Journal of Nigerian Association of Mathematical Physics, vol. 11 pp. 467-474

[13] [13] O. Goldreich and L. Levin, "A hard predicate for all one-way functions," Proceedings of the 21st Annual Symposium on Theory of Computing, ACM,1989.

[14] [14] R. Impagliazzo, L. Levin and M. Luby, "Pseudo-random generation from one-way functions," Proceedings of the 21st Annual Symposium on Theory of Computing, ACM, 1989.

[15] [15] C. Schnorr, "Efficient identification and signatures for smart cards," Advances in Cryptology - Crypto 89 Proceedings, Lecture Notes in Computer Science Vol. 435, G. Brassard ed., Springer-Verlag,1989.

[16] [16] T. Leighton and S. Micali, "Provably fast and secure digital signature algorithms based on secure hash functions," Manuscript, March _____

[17] [17] S. Evens, O. Goldreich and S. Micali, "On-line/Off line digital signatures," Manuscript, Preliminary version in Advances in Cryptology - Crypto 89 Proceedings, Lecture Notes in Computer Science Vol. 435, G. Brassard ed., Springer-Verlag, 1989.

[18] [18] M. Bellare, J. Kilian and P. Rogaway, "On the security of cipher-block chaining," Advances in Cryptology - Crypto 94 Proceedings, Lecture Notes in Computer Science Vol. 839, Y. Desmedt ed., Springer-Verlag, 1994.

[19] [19] S. Goldwasser, S. Micali and R. Rivest, "A digital signature scheme secure against adaptive chosen-message attacks," SIAM Journal of Computing,**17**(2), 281-308, April 1988.

[20] [20] R. Rivest, "The MD5 message-digest algorithm," IETF Network Working Group, RFC 1321, April 1992.

Digital Jewelry: A Wearable Technology for Enhancing Female Interest in Sciences

Osisanwo F.Y
Computer Science Department
Babcock University, Ilishan-Remo
Ogun State, Nigeria
yosisanwo@yahoo.com

Akinsowon O.A
Computer Science Department
Federal University of Technology, Akure
Ondo State, Nigeria
maggijoke2002@yahoo.com

***Abstract*—** Science Education is a major part of education that enhances the technological and economic development of a society. Over time, science education has been a male dominated branch of study. It is observed that the female gender pick less interest in the science owing to the fact that the field of science is assumed to be a hard nut to crack for females. Most females find themselves in the field of science and still wish they were in a different field. Some study science course and end up not practicing because they think that other fields are more interesting and simpler. The introduction of Information and Communication Technology (ICT) to sciences has done a lot of improvement in the field of science. This paper looks at the values that Digital Jewelry as a wearable device adds to science Education and how these can make the study of sciences and science-related courses interesting for female scientists.

Keywords: Wearable Technology, Wearable Computers, Digital Jewelry, Female Scientist.

I. INTRODUCTION

With the rapid development of technology and economy, numerous devices are being developed to improve the quality of life dramatically. One of these trivial devices is wearable computing device. After many years of research, wearable computing is showing signs of leaving the research laboratory and emerging into the world. Examples of typical consumer applications include body function monitoring products such as the Nike Bluetooth shoe that monitors the user's pace while running, as well as jackets with integrated iPod controls. The emerging field of wearable computing holds much promise in the educational computing arena [4].

In education, information and communication plays a major role in knowledge acquisition and also research works. For scholars in the field of science and technology, quality time is spent on information gathering and communication. This contributes a lot to the period of time spent on research work. Hence, the preference for other fields of study apart from science and technology by some supposed scientist.

According to [2], one of the clearest signs of a society's intellectual health is the strength of its science and math education. Science and math help to spur developments in scientific research and industrial technology, and ultimately lead to a more diverse, robust economy. But in the United States, and even in countries where a greater emphasis is placed on math and science, there is one segment of the population that does not always benefit from the best science and math education the community has to offer: girls and women. It has long been known that there is a "gender gap" in the sciences, which affects the quality and type of education offered to women even up to the advanced college level. Many women who ought to have become major researchers in science or great scientists lost interest in science and gave up the goals mid-way due to some factors or due to circumstances that ensued in the process.

This paper looks at some of the factors that make women lose interest in science and suggest digital jewellery: a wearable technology as a medium of closing the gender gap in sciences by enhancing the interest of girls and women in science subjects starting from high school level to advanced college level and postgraduate level.

II. AIMS AND OBJECTIVE

The main objective of this research is to consider the factors that discourage women from science and propose ways by which the emerging technology "Digital Jewellery" can enhance their interest.

III. WEARABLE COMPUTER DEVICES?

Wearable computers are what most would typically conceive as wearable technology. They boast designs and features that mirror what one may see in futuristic films. [3]

Wearable computers, also known as body-borne computers or wearables are miniature electronic devices that are worn by the bearer under, with or on top of clothing. This class of wearable technology has been developed for general or special purpose information technologies and media development. Wearable computers are especially useful for applications that require more complex computational support than just hardware coded logics. [13]

Wearable computing is the study or practice of inventing, designing, building, or using miniature body-borne computational and sensory devices. Wearable computers may be worn under, over, or in clothing, or may also be themselves clothes e.g. smart clothes [5].

According to [3], wearable computers are probably the most blatant example of wearable technology. Quite literally, wearable computers are computing devices that are designed to incorporate and engage the human body. Two major trends of the wearable computers are *miniaturization* and *multiple functionalities*. In order to incorporate the design and mobility of the human body, the various components of computing devices must not only be broken down and distributed throughout the body but also made tiny enough to be a useful tool rather than a bulky hindrance. Wearable computers contain tiny computing devices that make computation, communication and interactions easier and consistent. Meanwhile these are part of the skills needed in educational sectors for the purpose of study and becoming knowledgeable in a field of study.

A. Application Area of Wearable Technology

This technology that is futuristic is covering a lot of grounds in different fields and as it is still growing it tends to cover more, as there are many application areas of this technology yet more are still adapting this technology, some of the application area include medicine, clothing, manufacturing, training, recreation and communication. Wearable computers unlike desktop or laptop computers allow a much closer association of information with the user [12]

Several areas of wearable technology include, but are not limited to, wearable computers, second skins (or memory-rich clothing), digital jewellery and Personal Area Networks (PANs). Wearable computers are just that, computing devices that have left the desk and entered the world of fashion. Second skins act as additional sensory organs placed on larger surface areas of the exterior of the body, making them much closer to "fashion technology" than wearable computers. Digital jewellery, on the other hand, is definitely "fashion technology"; digital jewellery is the next step in the evolution of wearable technology. [3]

B. Digital Jewellery and Components?

Digital jewellery is the fashion jewellery with embedded intelligence. "Digital jewellery" can help you solve problems like forgotten passwords and security badges. Digital jewellery will be the evolution in digital technology that makes computer elements entirely compatible with the human form. Digital jewellery looks to be the next sizzling fashion trend of the technological wave. The combination of shrinking computer devices and increasing computer power has allowed several companies to begin producing fashion jewellery with embedded intelligence. By the end of the decade, we could be wearing our computers instead of sitting in front of them. Changes PC to Wearable PC [9]

Digital jewellery can be in the form of any kind of non-digital jewellery (bracelet, wristwatch, necklace, ring, piercings, etc.). The concept of digital jewellery is the merely digitalization of non-digital jewellery. These devices could function as global positioning devices, personal data assistants, telephonic devices, video communication tools and more [3]

Digital jewellery can permit a single device to be broken down into various parts and distributed to locations on the body that provide better utility. For instance, IBM created jewellery set at the turn of the century that functioned as a cellular phone. The various jewellery components serve a particular function similar to the various functioning components of a cellular phone. [3][9]

Ear Rings Speakers: these earrings will be the phone's receiver.

Necklace: transfer oratory messages, works like the telephone microphone

Ring is also a functional part of the whole set, providing an alternative numerical pad equipped with LEDs that flash to indicate an incoming call. It can also be programmed to flash different colors to identify a particular caller or indicate the importance of a call.

The same ring that flashes for phone calls could also inform you that e-mail is piling up in your inbox. This flashing alert could also indicate the urgency of the e-mail. The mouse-ring that IBM is developing will use the company's Track Point technology to wirelessly move the cursor on a computer-monitor display [9]

Bracelet Equipped with a video graphics array (VGA) display; this wrist display could also be used as a caller identifier that flashes the name and phone number of the caller. [9]

Figure 1

Figure 1; IBM"s telephonically functional jewellery set [3][9][13]

Figure 2: Image of the ZYPAD wrist wearable computer from Arcom Control Systems [3][9][13]

C. Features of Wearable Computers

1. **Miniaturization**: Instead of remaining as one whole device, computer devices are broken up into their basic components and packaged as various pieces of digital jewellery.
2. **Consistency (constant interaction)**: It's design makes it such that it can be accessed as the need arises since it is to be worn.
3. **Multiple Functionalities**: For example, mobile phones can be on the digital jewellery form by breaking it into earrings (which serve as ear-pieces), necklace (which serves as mouth piece where the microphone is embedded
4. **Multitasking**: IBM's magic decoder ring for example, will flash you when your phone rings and can also notify you when the numbers of unread messages in your e-mail box are increasing.

D. Advantages of Digital Jewellery

1) Security: Looking at the Java ring for example, it locks or unlocks doors or computers just as passwords and keys are used to lock doors and computers. With the Java ring, we don't need to remember passwords or go about with keys.
2) Portability: They have the components of computing devices in small sizes embedded in them. Hence, they are easy to carry about.
3) Acceptability: Looking at the fact that the focus of this trend is more on the device rather than the technology, it is not seen as intruding or introducing a whole new idea. The outlook does not look like a whole new concept, even though it is. For example, asking someone who has worn normal earrings before.

E. Disadvantages of Digital Jewellery

1) Small display screen: Due to size of the device, the display platform is small and can discourage people.
2) Affordability: Considering the cost of acquiring this jewellery can be a discouragement to the people, even though they know that it is needed.
3). Fear of rays being emitted by the jewellery: Another disadvantage is the fear people have for the emission of harmful rays from the devices. So many ailments have been attributed to the exposure of the human body to some of these rays. Hence, this fear may be justifiable except safety measures are put in place.
4). Also, the capacity of the device to keep charge since it would be on-the-go most of the time.

IV. GENDER DIFFERENCES IN INTEREST TOWARDS SCIENCES

Science, by principle, is meant to be progressive. While there has already been a relative paradigm shift towards gender equality, much still needs to be done to accelerate the closing of the ever-present gap in a way that will ensure sustainability. Education about the existence of the imbalance would be a good start. Implementing policies to ensure gender equality in hiring and support for researchers, and establishing tools or media through which gender interest can be enhanced [7]

Girls are considerably less interested in science subjects than boys. One reason may be that science subjects are considered to be generally masculine. Thus, being interested in science may threaten the self-perception of girls as well as the femininity of their self-image.

According to [11], the 'swing away from science' tendency is common in many countries. In Germany, for example, the percentage of students choosing natural sciences as a major subject has decreased by about 50% within the last 20 years. The dwindling number of students in the UK engaging in scientific fields such as biology, physics, statistics, or information technology has become a matter of great social concern as it is well known that members of these scientific disciplines contribute strongly to a nation's economic growth and well-being. Furthermore, the swing away from science may potentially cause more concern in the future as there is an ever-increasing need on the job market for individuals who are highly trained in science subjects.

Thus, a lack of interest in science at the high school level may eventually pose a serious threat to economic prosperity. There is worrying evidence to support this theory: When students were asked to rank school subjects according to personal preference, typically, science subjects are at the lowest ebb. Moreover, compared to other subjects, students' decreased interest in science subjects is particularly marked and this downward trend is most evident in subjects such as physics and chemistry.

Crucially, there are also considerable gender differences in interest levels in science subjects: Boys are far more likely than girls to claim to like science and the decrease in interest in science amongst girls is even more pronounced than amongst boys This is somewhat surprising, as nowadays girls on average perform as well as, if not better than, boys in many science subjects. Furthermore, even if girls have the same level of competence in science as boys, girls will engage considerably less often in science subjects. Consequently, it might be suggested that the gender gap in interest in science could lead, in turn, to an over-representation of boys in science fields and occupations, a trend that can already be observed and one that will most likely continue in the future. [11]

A. Factors Causing Lack of Interest towards Science

Despite efforts to give women greater access to education in science and technology in some countries, research shows they are still significantly under-represented in many degree programmes, especially in engineering, physics and computer science. But even with improved access to science and technology education, women have not increased their numbers in the workforce, study finds. In fact, in some countries, including the USA, the number of women in the science and technology workforce is declining [8]

There are diverse kinds of factors responsible for the gender difference in interest towards science; some of these are discoursed;

1) Individual Interest

This is a relatively stable individual motivational disposition towards specific objects of interest and is characterized by increased attention and (in most cases) positive affect and emotions. The object of interests could be concrete, a particular content, events, or ideas. Individual disposition towards interest is characterized by the tendency of individuals to re-engage with an object of interest and thus represents a relatively stable person– object relationship. Thus, in the science education context, interest in science represents a student's enduring disposition to engage in science school subjects (which represent the objects of interest, e.g., physics). Often girls and women are affected by individual interest as more effort is required of them to have a stable motivation towards science, as most of the sciences are taken in abstract, hence the difficulty to pick interest in the subjects.

2) Situational Interest

By contrast, is triggered by the environment, and is characterized by a state of actual attraction or curiosity as well as increased attention. Situational interest comprises two phases. The environment plays a larger role in making science interesting for women as various environmental factors could readily discourage girls and women, such as

a. The teachers' factor; the perceptions and attitude of teachers teaching in school are discouraging; this attitude often makes girls to drop out of school. Teachers in some rural part of Africa believes that boys will go to university to take courses like medicine, engineering, architecture and so on, while girls will only study to become secretaries, teachers, designers and so on [1]. This perception or attitude of teachers inform how they teach and attend to the girls in classes especially mathematics and science, these end up affecting the performance of the girl child. This then will lead to reduced school attendance and finally lost of interest in science or school altogether.

b. The home factor; balancing family life and personal relationships while achieving a successful career in science is becoming a major challenge for women; this has made some women forgo advancements in their careers in order to devote time to raising a family. The challenges of child care and the demands of running a research laboratory are often seen as incompatible. Women who plan to have children in the future drop out of the academic research race at twice the rate of men. Women are hit hard with family responsibilities just when they need to meet research goals [6].

V. ENHANCING INTEREST OF WOMEN IN SCIENCE WITH DIGITAL JEWELLERY

With the aid of this emerging technology: digital jewellery, interest of women in science can be enhanced based on all the features and advantages that digital jewellery offers. Considering the various categories of women in science in relation to their needs and what digital jewellery can offer the following can be achieved:

a) Researchers:

Female science researchers can be encouraged with this device as most of what is needed by them is information and communication, with digital jewellery communication with various researchers and scholar at any location is made easy. Mails will be exchanged easily, sharing of experiences and interaction is made consistently because of the consistency of the digital jewellery. Multitasking is made possible as several things can be done at once by the researcher once there is need to read message or take a procedure/session the device will alert the researcher to this. With the miniaturized nature of the digital jewellery, PCs are carried around everywhere meaning work continues. Via this device too, electronic data can be gathered for research purpose easily.

b) Teachers/lecturers:

For female science teachers communicating with their students becomes easier, although the eye contact messages may be missing in the teaching process, but the teacher may be home based yet interact with her students frequently, even while on some other task, as long as she can multitask herself. Lecture notes can be communicated to her students; she could read mails from them at any time and respond at any time.

c) Students in Science and maths education

At the elementary and secondary/high school level, digital jewellery can be used to encourage female students to stay in the field of science and technology. It can be used for authentic real-world contexts content. Imagine basing all word problems on data about students' own lives collected by their bracelets! Now that would be authentic. It tends to excite the students into further research. Also, students think about their learning all day, every day, and not just while they're in classes. Wearables integrate technology more seamlessly than even a smartphone does; while a smartphone is designed to be put away in a pocket, a (wearable) jewellery is as much a part of a person's everyday as his/her clothing. With this device, assignments, task can be easily solved by students

At the elementary level, application could be designed and integrate into the device that will play science songs and stories to their hearing, constant interaction with this may inform their interest and choice in science.

d) Students with poor performance

This emerging technology can also assist in stepping up performances of poor students, with the applications integrated or embedded intelligence a personal tutor can be introduce that will teach the poor student repeatedly a concept over a period of time until it is fully taking.

e) Students who dropped based on religious factor/ home factor

Female students who are made to drop out of school base on either religious or home factor can make use of this technology to quietly learn at home. Although in this learning process, physical interaction with other students and teachers together with the benefits attached will be missing.

VI. CONCLUSION

With the portability, non-evasiveness, consistency, miniaturization and other features of digital jewellery, the interest of females in Science Technology and Mathematics can be beefed up. Women who because of family commitments had to drop out of science can now plan their schedule to suite them readily since digital jewellery is always available. All that would be required at this time would just be to know how particular digital jewellery works and suites the need at hand. Then necessary training or required skills for making use of it can be acquired. These technologies can also improve the performance of the girl child in Science Technology and Mathematics (STM) and increase the population of women in the field of Science and at the same time not losing their fashion sense but in fact, enhancing it.

VII. RECOMMENDATION

It is therefore recommended that government, tertiary institution, private school owners, educational agencies, educational policy makers attend to the need of encouraging further development in this area so that it can be affordable ans easily accessible by as many females as are interested in STM at various levels, starting from elementary to tertiary institutions; even up till post-graduate level.

REFERENCES

[1] Angeline Nhantsi (2011), "Girls, Schools and Statistics Education in South Africa", First Africa Young Statisticians conference, retrieved on 15th September, 2011 from http://www.statssa.gov.za/ycs.

[2] Barcodes Inc (2013) "Gender Equity in Science Education", **BarcodesInc**, 200 W Monroe St Chicago IL 60606.

[3] Cody Gwen (n.d) " Wearable Technology: Fashion Meets Applied Science" retrieved from http://student.uscupstate.edu/owensch/FinalPaper.pdf on 3rd March, 2014.

[4] Ngai Grace, Stephen C.F. Chan,, Joey C.Y. Cheung, and Winnie W.Y. Lau (2010) "Deploying a Wearable Computing Platform for Computing Education" IEEE Transactions on Learning Technologies, Vol. 3 No.1

[5] Mann, Steve (2013): Wearable Computing. In: Soegaard, Mads and Dam, Rikke Friis (eds.). "The Encyclopedia of Human-Computer Interaction, 2nd Ed.". Aarhus, Denmark: The Interaction Design Foundation. Available online at http://www.interaction-design.org/encyclopedia/wearable_computing.html

[6] Pappas Stephanie (2013)" Science Gender Gap: Five Reasons Women Trail Men In Science", Live Science retrieved from http://www.huffingtonpost.com/2013/03/07/science-gender-gap-five-reasons-women-men_n_2827567.html.

[7] Payam Z. (2013) "Closing the Gender Gaps for Women in Science", Immpress Magazine.

[8] Research Information (2012) "Gender inequality continues in science and technology", retrieved from http//www.researchinformation.info.

[9] Schwartz, Ephraim (2000) "Wireless world takes James Bond-like twist with wearable digital jewellery" Trade Publication, InfoWorld;08/21/2000, Vol. 22 Issue 34, p59B

[10] Simon T. Perrault, Hind Gacem, Eric Lecolinet, Yves Guiard (2013) "Augmenting digital jewelry with advanced display capacities" Workshop on Displays Take New Shape: An Agenda for Interactive Surfaces, April 28, 2013, Paris, France.

[11] Sylvie Kerger, Romain Martin and Martin Brunner (2011) "How can we enhance girls' interest in scientific topics?" British Journal of Educational Psychology.

[12] Woodrow Barfield, Thomas Caudell (2001) " Fundamentals of Wearable Computers and Augmented Reality" in Woodrow Barfield, Thomas Caudell (eds), Lawrence Erlbaum Associates Inc. USA.

[13] Wikipedia (2014) "Wearable Computers" *en.wikipedia.org/wiki/Wearable_computer* retrieved on 4th March 2014

Embedded Systems in Healthcare Systems

Maitanmi, S. O
Computer Science, Department
Babcock University
Ilisan Remo, Ogun State
maitanmi@yahoo.com

Idowu, S. A
Computer Science, Department
Babcock University
Ilisan Remo, Ogun State
Saidowu07@gmail.com

Abstract: One of the biggest achievements nowadays is to increase people's life expectancy with constant, adequate and efficient healthcare systems. This cannot be achieved without mentioning embedded systems which are part of human life. For instance, one cannot imagine life without mobile phones for personal communication. Its presence is virtually unavoidable in almost all facets of human endeavor. The emergence of embedded systems development has not only been associated with terms such as *smaller*, *faster*, *smarter*, and *connected*, but also a reduction in term of costs. Despites the advancement recorded in health sector through the implementation of technological innovations with the convergence of biotechnology, nanotechnology, manufacturing technology, communication technology and device, sensor technologies are making breathtaking transformations in healthcare delivery and creating new health care paradigms. The paper aims at giving a general overview of the impacts of embedded systems in healthcare with respects to Cyber Physical Systems**.**

Keywords: Embedded system, healthcare, Cyber Physical Systems (CPS), Sensor.

I. INTRODUCTION

Health and education are the two important prerequisites for human capital development, and have been demonstrated to be the basis of individual's economic productivity. Good health is a critical input into poverty reduction, economic growth and long-term economic development which support the old saying that 'health is wealth' [7].

Health sector in any country has been recognized as the primary engine of growth and development [2]. But despite this, the health sector has witnessed various instability that has negatively reduced the progress recorded at various times. Some of these persistent challenges in healthcare management are: gross inadequate infrastructural support, acute shortage of electricity supply, inadequate potable water, inadequate diagnostic laboratories, lack of willingness to spend on health treatments by patients, and a total absence of a community-based integrated system for disease prevention, corruption of leaders, illiteracy, very low budget on health, inadequate mechanisms for families to access health care, shortage of essential drugs and supplies, fake drugs in circulation, and inadequate supervision of health care providers.

II. EMBEDDED SYSTEM

Embedded system is any device that includes a computer but is not itself a general purpose computer. "An embedded system is a special-purpose computer system, which is completely captured by the device it controls". It is a small computer system that is generally hidden inside the equipment (machine, electrical appliance, electronic gadget) to increase the intelligence of the equipment for better or more efficient functionality. Embedded system involves both the software and hardware co-development. Embedded systems are often easier to understand in terms of smart devices, intelligent or automated devices. Typical examples are; mobile phones, washing machines, wristwatches among others. Embedded system involved the inbuilt of tiny devices called microprocessors which in respect would lead to the increase in smartness and intelligence of automated devices.

An embedded system has specific requirements and performs pre-defined tasks, unlike a general-purpose personal computer.

The meaning of embedded system may not be well defined or explained without mentioning telemedicine. A research carried out by Marwedel [5] considered "telemedicine **as a process, rather than a technology:** telemedicine connects patients and healthcare professional in a chain of care". An interaction with Marwedel [5] the author of Embedded Systems textbook revealed that telemedicine is an aspect of cyber physical system.

A. CHARACTERISTICS AND BENEFITS OF EMBEDDED SYSTEMS

Embedded systems are usually designed to perform selected functions at a low cost. The system may need to be very fast for some functions, but most of its other functions will probably not need speed. We can thus define a smart device by the following attributes.

1. Computational power: The small amount of computing power required by these devices is provided by simple 8-bit controller or by high end 64-bit microprocessor.
2. Processing power: The processing power of microprocessor include the followings:
3. Memory: These devices possess some amount of memory that can be used by the

processor and also to remember the user data and preferences.

4. Real-time: These devices have to respond to the user's input in a specified period of time.
5. Communication: The device must be able to receive inputs given by another device in the environment, process it and provide some tangible output to other devices or users.
6. Dynamic decisions: The system should be able to change its next course of activity based on the change of input from sensors or surroundings. This distinguished it from artificial intelligent because of the presence of inference engine and specific programming languages which are involved when building artificial intelligence such programming languages are Prolog and Lisp to mention a few. .

BENEFITS OF EMBEDDED SYSTEMS

It can satisfy the strict requirement of functionality,

- reliability,
- cost,
- volume,
- and power consumption of the particular application. Years back to power a cathode ray tube monitor requires a high voltage but recently flat screen monitors requires a very lower voltage.

B. HEALTHCARE MANAGEMENT

Present, Emerging and Future System of Medical Device Systems

According to the CPS lee [4], it was observed that over the last 20 years or more, designs for medical devices have evolved from analog to digital systems. Today, software, microprocessor, sensor, and actuator technologies are ubiquitous in these devices. Most devices contain embedded systems that rely on a combination of proprietary, commercial-off-the-shelf (COTS) and custom software or software-of-unknown-pedigree (SOUP) components.

Emerging medical device systems may be characterized as those systems that are pushing the envelope of regulated technology. Examples are closed-loop or networked device systems. Open loop is used to describe systems designed to allow ready interconnection and interoperation with other systems; closed loop refers to those designed to be used in isolation.

As technological advances and innovation increases, medical device systems also decrease in size while increasing in capabilities, it is reasonable to expect that future devices could evolve into ubiquitous supervisory-control, patient-centric systems performing autonomous, cooperative, and coordinated actions as summarized in table 1.

Table1. Developmental changes in Health care system. Source: [4]

	Current	**Emerging systems**	**Future**
Architectures	Device architectures are highly proprietary, not interoperable, and rely heavily on professionals to provide inputs and assess outputs.	Emerging medical device architecture is beginning to provide wired and wireless interfaces to facilitate networked communication of device (patient) data.	The future research claimed to be heading toward the merging of physiological, biological, engineered, and physical systems for health care for example, biomechanical systems.
Development Methods	Range from older methods such as structured programming to object-oriented programming paradigms where objects are instantiated at run-time	It is aimed that emerging development methods must scale across the device industry's entire problem set as well as its diverse development skill levels. The ability for third parties to trust all aspects of design and development results is key to device innovation.	It is likely that a greater dependence on trust in, physical and biological models will become necessary.
Research challenges	Platform and implementation technologies are needed that support science and engineering based design, development, and certification.	-Synthesizing medical information. As medical systems evolve into cyber-physical systems, the notions of feedback and interoperability will most likely emulate the practice of medicine itself -QoS guarantees. -Dynamic management of large-scale systems. The workshop participants' report touched on the notion of dynamic medical-system configurability in the context of plug-and-play (PnP) systems. -Trust in abstractions. Need for more robust means of modelling and formal methods-based checking and verification of medical-device designs -The security challenge.	-Consumer demands and needs: An already high and still-growing cost is associated with traditional (hospital, clinic) care settings. Such as home care, assisted living, and commoditized convenient care settings. -Networking: Future computing would be interconnected in complex open system with heterogeneous components. -Sensor fusion: Academic, government, and industry research today is exploring sensors that can be implanted or worn (e.g., attached to clothing), communicatin

	Current	Emerging systems	Future
		(software/hardware)	g either wirelessly or through networks woven into fabrics, and used for gait analysis, detection of falling, and monitoring for abnormal mobility and activity levels. -Prosthetics: The growing demand for devices for people with debilitating injuries or disabilities is driving innovation in prosthetics of all kinds. -Control and Feedback:

ADVANTAGES OF CPS TO HEALTH MANAGEMENT

- Improve Patient safety
- Produce complete and accurate medical data records
- Reduce errors
- Context awareness
- Initiate rapid deployment
- Introduce safety interlocks

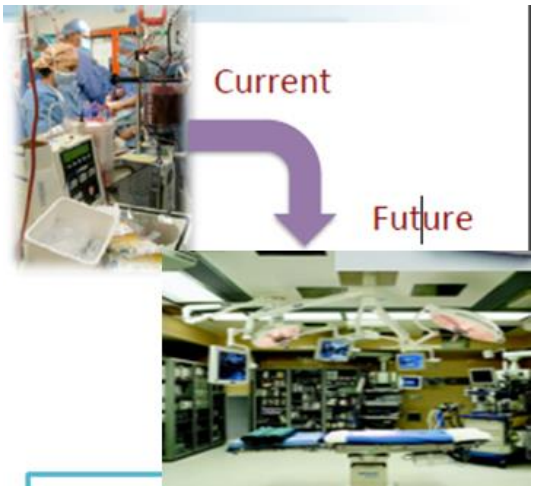


Fig. 1. Benefits of CPS to health care systems. Source: [4]

Clinical care: Hospitalized patients whose physiological status requires close attention can be constantly monitored using CPS noninvasive monitoring. This type of solution employs sensors to collect comprehensive physiological information and uses gateways and the cloud to analyze and store the information and then send the analyzed data wirelessly to caregivers for further analysis and review as demonstrated in figure 3.

Remote monitoring: There are people all over the world whose health may suffer because they do not have ready access to effective health monitoring. But small, powerful wireless solutions connected through the Internet are now making it possible for monitoring to come to these patients instead of vice-versa as seen in fig. 2.

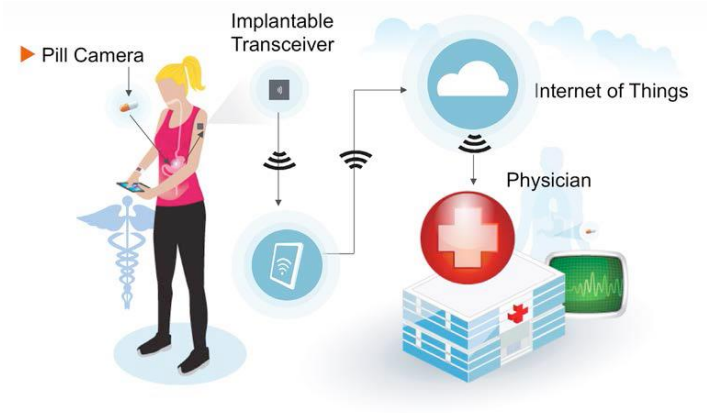
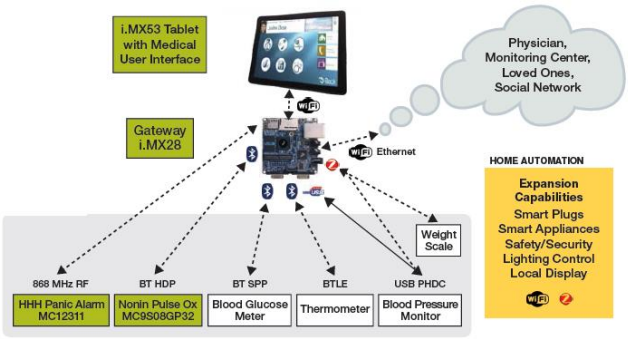


Fig. 2. Remote monitoring System. Source: [4]

Early intervention/prevention: Healthy, active people can also benefit from CPS in monitoring of their daily activities and well-being. A senior living alone, for example, may want to have a monitoring device that can detect a fall or other interruption in everyday activity and report it to emergency responders or family members. For that matter, an active athlete such as a hiker or biker could benefit from such a solution at any age, particularly if it's



available as a piece of wearable technology.

Fig. 3. Early intervention using smart devices. Source: [5]

III. IMPLEMENTING SMART MEDICAL DEVICES IN MEDICAL CYBER PHYSICAL SYSTEMS

- Anesthesia Unit Monitor: An anesthesia machine administers anesthesia to patients. Anesthesia is given in 90 percent of surgical interventions by two ways, intravenously or inhalational gas anesthesia [6].
- Continuous Positive Airway Pressure (CPAP) Machine: The most common treatment for sleep apnea is a method of pushing air through the airway called continuous positive airway pressure (CPAP). The main goal for this device is to provide constant positive pressure to the respiratory system in order to prevent muscles from obstructing the airway
- Defibrillators: An automated external defibrillator (AED) is a portable device used to restore normal heart rhythm to patients in cardiac arrest by delivering an electrical shock to a patient thought the chest wall. Cardiac arrest is an abrupt loss of the heart function.
- Digital Stethoscope: The digital stethoscope is a step forward from its counterparts, with the capability of filtering and amplifying the signal.
- Electrocardiograph (ECG): The Electrocardiograph (ECG) Accessory utilizes a function-specific sensor that acquires and monitors the user's heart rate and passes the signals to the smart phone application to display the heart rate data and optionally sends it to your healthcare provider [6].
- Hearing Aids: A hearing aid is a small electronic device worn in or behind the ear which amplifies incoming sounds for the hearing impaired.
- Hospital Admission Machine: A hospital admission machine is an acquisition system that assesses the health of a person to increase efficiency through automation and administration procedures.
- Infusion Pumps: An infusion pump is an electronic device capable of delivering fluids such as blood, medication and nutrients to a patient via the circulatory system, subcutaneously, or epidurally in a controlled way.
- Multi-parameter Patient Monitor: A multi-parameter patient monitor measures blood pressure, oxygen saturation, temperature and heart electrical activity.
- Power Patient Beds: Powered patient beds offer a number of functions for patients and health care professionals. These highly integrated beds may contain a vital signs monitor, heart rate monitor, ECG, infusion pump, bed tilt controls, bed wheel controls and a communication interface.
- Ventilators and Respirators: A ventilator (also known as a respirator) is a machine designed to mechanically move air in and out of the lungs to intermittently or continuously assist, or control pulmonary ventilation.

IV. CONCLUSION

The paper has thus far given the general overview of embedded systems as related to healthcare management systems. It is therefore no doubts from the aforementioned points that developments through smart devices and other technological advancements by the use of CPS has helped in salvaging more lives than past two decades.

REFERENCES

[1] Federal Ministry of Health (FMOH) [Nigeria]. Strategic Plan for Rolling Back Malaria in Nigeria 2001-2005. Abuja, Nigeria: Federal Ministry of Health Unpublished 2001.

[2] Herfon, "Health Reform foundation of Nigeria Annual Report" unpublished 2011.

[3] E. A Lee, and S. A Seshia, "Introduction to Embedded Systems - A Cyber-Physical Systems Approach", 2011

[4] E.A Lee, "Cyber physical systems: design challenges. Technical report", EECS Department, University of California, Berkeley, January (CPS Steering Group) 2008 Cyber-physical systems executive summary. In: Cyber-physical systems summit. http://varma.ece.cmu.edu/summit/CPS-Executive-Summary.pdf

[5] Lee I. Medical-Device Cyber Physical Systems PRECISE Center Department of computer and Information Science University of Pennsylvania TCES Workshop November, 2010

[6] P. Marwedel, "Embedded System Design" by Springer 2006

[7] D. Work, A. Bayen, and Q. Jacobson, "Automotive cyber-physical systems in the context of human mobility "National workshop on high-confidence automotive cyber-physical systems. 2008

[8] J. Sachs, "The Strategic Significance of Global Inequality," The Washington Quarterly, Summer: 191.

[9] Sen, A. 1999. Development as Freedom. New York: Alfred A. Knopf. 2001

A Medical Information System for Data Migration in Ghana Using Health Level 7 (HL7) Message Standard Protocol

Dominic Damoah
Bethel University
MONGU – ZAMBIA

Nicolae Goga
Molecular Dynamics Group
University of Groningen
Netherlands
Politehnica University of Bucharest,
Romania

Mukuze Nelson
Faculty of Science
Department of Computer Science
Valley View University
Accra-Ghana

Abstract- This project aims at the development of a web based message generation and validation system that utilizes HL7 version 2 standards and message formats to encode healthcare data and produce their correspondents in xml format to enable their migration from one database to another independent of the database design. The project was limited to the HL7 ADT (Admission, Discharge, and Transfer) message type. At the end of the project two modules were developed: The first module forms the HL7 messages, validates them, and transforms them into their xml format using the C# programming platform in Microsoft Visual Studio 2008. The second module migrates the xml formatted messages to the receiving database using OPENXML, an xml migration technology also supported by Microsoft Visual Studio 2008 C#. The back-end is based on the Microsoft SQL Server 2005.

***Keywords*—HL7, Interoperability, Openxml, Health Information System, ADT. Inter-Hospital, Intra-Hospital**

I. INTRODUCTION

Health is an important aspect of human life and an important indicator of the development level of a country. Improved health care for the patient goes beyond the provision of the adequate IT infrastructure; it involves adhering to protocols and standards to enhance interoperability in the health sector. Interoperability refers to the ability of systems used in health care to communicate with one another and share medical information as the need arises. Health Level 7 (HL7) is a message standard protocol that is widely used for medical information exchange. The project is intended to explore the feasibility of implementing and deploying a centralized medical informatics system for Ghana using health level 7 (HL7) message standard protocols. This project focuses on the encoding and decoding of HL7 messages for the purpose of transmitting medical information from one database to another independent of the database design. In this project the technology used in the encoding and decoding of HL7 messages is XML (Extended Markup Language). The project also provides a practical understanding of the use of web services and other web technologies like SOAP (Simple Object Access Protocol) to enhance communications over the network of connected health information systems and other health related IT equipment.

II. RELATED WORKS

The Malawi National Electronic Data System Task Force developed a national HL7 standard for messaging for the ART M&E dataset, based on HL7 message standard version 2.5. According to this specification a master card called ART Master Card is used to record most patients' demographical, clinical and treatment information. All variables are mapped in HL7 message structure [1].

In Romania the Integrated Management System for medical information using Hl7 standard – SIMIMED was developed with the main objective of establishing communication between hospital information systems and other information systems used in healthcare system. As shown in figure 1 SIMMED contains four important modules designed to provide services. These are integrated health information system, datacenter, the web application and the admin tool [2].

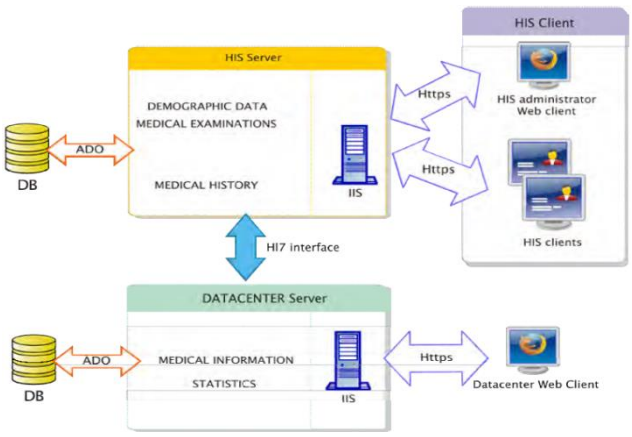


Figure 1. SIMIMED system architecture [3]

ENCODING MEDICAL INFORMATION

The SIMIMED provides an HL7 interface that allows medical information to be exchanged between the various health information systems in various hospitals and the datacenter. This exchange is made possible by the implementation of web services loaded by the internet information server (IIS). The internet information server provides a number of services including web services. These vendor-neutral technologies like XML (Extensible Markup Language) make it possible for heterogeneous systems to communicate or exchange information. In the case of the SIMIMED web services loaded by the internet information server (IIS) allow the databases of the different health information systems to exchange medical information even though their designs are based on different database engines like ORACLE, MySQL, SQL, Microsoft SQL server, etc.

III. CRYSTALLIZATION OF THE PROPOSED SYSTEM

This section offers an overview of the Ghanaian health system (organization, structure, and challenges), and the National Health Insurance Scheme (NHIS). This overview contributes to the understanding of the physical environment in which the distributed medical informatics system is to be deployed. The first part offers an in-depth analysis of the laws that govern health information privacy and security; the current status of health information systems (HIS); and network infrastructures in Ghanaian health centers. This analysis provides the technical basis for the implementation of an information system for medical information exchange in Ghana. The second section deals with the designs and implementation. The focus remains on the exchange of medical information between different database engines.

A. THE HEALTHCARE SYSTEM REGULATIONS

The constitution of Ghana provides that, "The state shall safeguard the health, safety and welfare of all persons in employment, and shall establish the basis for the full deployment of the creative potential of all Ghanaians" [3]. The current nationwide health insurance scheme is based on the parliamentary bill "ACT 650 and LI 1809" (National Insurance Act), which was passed into law in August 2003. The National Health Insurance Scheme (NHIS) was formally launched in December 2004 [4] The Health Ministry is responsible for policy planning processes and information management, particularly concerning the areas of financing, human resources and infrastructure [5]. The Ghana Health Service (GHS) is in charge of transport, equipment and infrastructure provision, information dissemination and provision of support for the design of policies and strategies to the Ghana Health Service Council. The activities of the various organs under the Ghana Health Service are coordinated and administered by the Ghana Health Service Council supervised by the Minister of Health. Its main objectives are to "implement approved national policies for health delivery in the country, increase access to improved health services and manage prudently resources available for the provision of health services" [2]. There are other allied organizations that provide health Services such as the National Health Insurance secretariat and the auditing service come under the council [6].

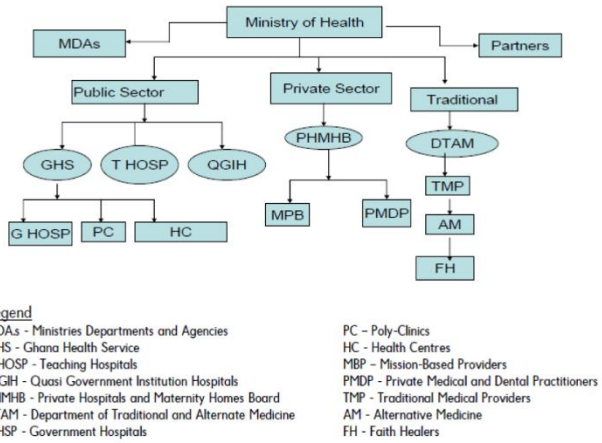


Figure 2: The Structure of Health Providers

As shown in the figure 2 the largest health provider is the public sector and this is followed closely by private sector. The government of Ghana is solely responsible for Ghana health Service, Teaching Hospital and Quasi Government Institution Hospital. The private sector is dominated by the Mission Based Providers supported by the private Medical and Dental Practitioners.

B. HEALTH DELIVERY STRUCTURE

Table 1 shows the Health Delivery Structure currently in operation in Ghana

Table 1: Level of Treatment, Institutions and Their Roles

Treatment levels	Role	Type of institution	Main personnel in charge	Basis of treatment	Reim-bursements	Operators	Demo-graphics covered
Primary	Preventive, information and curative	Community health centre	Community health officers	First point of contact	NHIS, "Cash-and-carry" and gratuitous	Government, mission and others'	Towns and surrounding areas
	Prevention and information	Rural health centre	Rural health worker				Rural areas
Reproductive health care	Birth attendance	Maternity home	Midwives	First point of contact	"Cash-and-carry"	Government, mission and others	Rural, towns, urban and metropolitan
Secondary	Curative	District hospital	General medical personnel	First and second point of contact on referral	NHIS, "Cash-and-carry"	Government and mission	
-	Training of Health workers	Training Centre	Public health trainers	-	-	Government and Mission	Rural areas and towns
Secondary and tertiary	Curative	Regional Hospital	General medical personnel	Second and third point of contact	NHIS, "Cash-and-carry"	Government and Mission	Towns Urban and rural areas
Primary and secondary	Curative, training and preventative	Polyclinic	General medical personnel	First and second point of contact	NHIS, "Cash-and-carry"	Government	Urban and Metropolitan Centres
First, second and tertiary		Teaching Hospital		First, and Second point of contact or referral			

C. HEALTHCARE CHALLENGES

Among other socio-economic priorities, health is one of the issues that are at the forefront of the Millennium Development Goals (MDGs), which Ghana hopes to achieve by 2015. Goal 4 of the MDGs is to reduce under-five mortality rate to two-thirds by 2015; Goal 5 is to reduce the maternal mortality ratio by three-quarters by 2015, and Goal 6 is to reduce infection rates of HIV/AIDS, malaria, and other communicable diseases associated with hygiene and environment by 2015. The main challenge in achieving the health-related MDGs, is blamed on the overall coverage and inaccessible healthcare services in poor neighborhood.

A study in 2005 identified health information and data handling as a core part of healthcare delivery and said, that "the single most frequently quoted constraint to improved health information is human resources availability and capacity at all levels, national, regional and district." According to the study, people with advanced skills in health information and data handling were seriously needed in the system. The study also said the Ghana Bureau of Statistics which handles these data is overstaffed with under-qualified personnel lacking the proper skills for health information management [7].

D. NATIONAL HEALTH INSURANCE SCHEME (NHIS)

The main objective of the National Health Insurance Scheme (NHIS) is to ensure high availability of basic healthcare services to all Ghanaians or residents who are registered under the scheme. Three types of insurance schemes exist under the National Health Insurance Scheme: First is the District-Wide (Public) Mutual Health Insurance scheme through which the workers of the public sector contribute directly into the health insurance system;

Second is the "Private Mutual Health Insurance scheme" through which workers in the informal sector and unemployed pay their contributions;

Third is the "Private Commercial Health Insurance Scheme reserved for payment of contribution through companies and multi-national companies [4].

Within these schemes, the health insurance programme offers the following benefits package:

Full Out and In Patient treatment for both surgery and medical conditions including the cost of feeding;

Full payment for medicine if within the approved list. Payments for referrals if within the approved list.

The exclusion list stipulates, "Appliance, prostheses, rehabilitation, dentures, organs aids, cosmetics surgery, assisted reproduction, HIV retroviral drugs, hormone and organ replacement therapy, heart and brain surgery other than accidents and diagnosis and treatment abroad, dialysis for chronic renal failure and cancers" [4]. As at June 2009 registered members of National Insurance Scheme were13,840,198 representing about 67% of the 2004 estimates. Of this number only1693672 have not received their cards. This means 12,146,526 were issued identification card just shy of 88% registered members have received their cards. A total of about 721,163 pregnant women were under free maternal care [8].

E. CURRENT STATUS OF MEDICAL INFORMATICS IN GHANA

Medical informatics as herein considered is the use of computer technology to help organize, analyze, manage and use information in improving health care. Its chief goals are to develop standards and clinical guidelines, formalize health terminology, enhance the usability of health information systems to facilitate communication and manage the flow of information from physicians to patients and vice versa [9].

Hospitals and other healthcare providers in Ghana acquire health information from various sources such as patients, statistical institutions, and other healthcare institutions. However, possessing health information in itself is not much beneficial until it is organized into a form useful to all stakeholders in the health sector. To solve the problem of health information organization, healthcare providers in Ghana are increasingly computerizing their health information into a repository. Processing of health information is also one major hurdle for Ghanaian Health centers. Indeed, access, storage, retrieval and processing of patients' health records are usually done manually and not effective. Most health centers just gather and use data only for treatment and temporarily usage, forgetting decision making purposes and secondary use of data.

F. MEDICAL INFORMATION PRIVACY AND SECURITY

Accurate and comprehensive health Information is critical to the quality of healthcare delivery and to the physician-patient relationship. However the privacy and security of

health information must be guaranteed by the healthcare providers and other stakeholders in the usage of health information. The existing laws and regulations in Ghana do not adequately address concerns relating to Health Information.

The nature of health information and medical records allows details of individual records to be captured as part of the overall statistics. The management of this primary source of information goes a long way in protecting the individual and allows informed decisions to be made at the management level. Inaccuracies in the information, or its improper disclosure, can affect an individual's dignity [5].

There are concerns raised about patients' privacy since it has the potential to facilitating and compromising personal identification information. The concerns raised are very legitimate as the individual has right to privacy.

Medical Profession Regulation requires that to ensure patient privacy the medical practitioner must seek the patients consent before information obtained in the course of the patient-doctor professional relationship could be disclosed. In the same vain the law gags the practitioner after disclosing such information in the interest of the public by law enforcement agencies this time without the consent of the patient. Aggregate reports on outbreaks of diseases is not regarded as evading privacy. It is intended for the good of society especially where it helps in the management of epidemics or prevalence. Finally in Ghana the disclosure of any information about an AIDS patient to any other person by health practitioners is strictly prohibited.

G. HEALTH INFORMATION SYSTEMS (HIS) IN GHANA

Hospitals and other healthcare providers in Ghana use a paper-based system to store virtually all health information. Paper-based health information storage makes access, retrieval and processing of patient health records ineffective. Many times when visiting hospitals we wait for hours in the queue while the nurses are busy searching for folders in a heap of folders; the case is even worse when a patient has forgotten the hospital card and folder number. Few hospitals and health providers have functional databases, to store health information. Even when the information is stored in databases they are disparate and do not exchange information. Therefore, even though one's personal information may be stored in a database one may still have to carry a folder to the doctor's office since the database containing ones records cannot be accessed by the doctor electronically. Also until very recently, databases such as Medline were available only to experts. Today few private hospitals such as TRUST Hospital have fully functional information systems that automate the key activities and transactions in the entire hospital. The only problem these hospitals face is the lack of interoperability which makes it impossible for information systems from various hospitals or other health institutions to exchange health information. Rural areas have the lowest level of health care provision and hence the greatest problems in public health. The lowest level of health care provision is as a result of the poor medical infrastructure. Health information systems are inexistent in these areas. Finally, Patients who are the main stakeholders in the health sector do not have access to their health records because information stored in a paper-based system is not usually sharable and it remains at a single location (at the health institution). Also the few health information systems that are functional in some private hospitals are not accessible from outside the boundaries of the hospital. In conclusion, health information systems in Ghana are not very efficient, reliable, accessible and available. Among the main reasons are factors related to the communication infrastructure, information sharing, and continuous training of health professionals.

IV. THE DISTRIBUTED MEDICAL INFORMATION SYSTEM (DMIS) FOR GHANA

A. REQUIREMENTS: UNDERLYING NETWORK INFRASTRUCTURE

For the implementation of an effective distributed medical informatics system for Ghana there should be an underlying network infrastructure to support the communication between the various stakeholders in the health sector. Currently few health centers (approximately 50%) in the Accra metropolis have individual network infrastructures. These network infrastructures are mostly based on the bus network topology. Even though the bus topology infrastructure is easy to implement it presents a number of disadvantages that make it inappropriate especially when performance degrades as additional computers are added, and it is easy to hack [10]. The rest of the health centers do not have any network infrastructure. In rural areas the case is even worse as technology is almost non-existent in the health centers. Therefore, there is a need for all health centers in the country to deploy effective individual and unified network infrastructure. After health centers have deployed their networks individually, there is a need to implement a nationwide wide area network (WAN) based on any effective network technology. This WAN should be the underlying network infrastructure for the Distributed Medical Information System (DMIS) for Ghana.

B. FUNCTIONAL HEALTH INFORMATION SYSTEMS

Health information systems (HIS) are at the heart of the Distributed medical information system (DMIS) for Ghana. The current status of health information systems is not good because practically one or two private hospitals can boast of a functional HIS. In the public sector there are no functional health information systems at all.

C. FEATURES: HL7 (HEALTH LEVEL 7 MESSAGING STANDARD PROTOCOL)

As healthcare organizations contemplate integrating their information systems, it is becoming apparently clear that many of these systems are not compatible. Hence a standard protocol to exchange data will obviously be needed especially if we are to integrate medical institutions in Ghana. The best and most widely used standards in healthcare tried and tested is the Health Level 7 (HL7) Message Standard protocol, a globally recognized electronic data interchange and messaging standard for exchanging healthcare data. HL7 Version 3 is XML-based and is becoming widely accepted as the right alternative for the replacement of HL7 Version2 since it requires off-shelf tools (XML technologies).

D. AUTHENTICATION

The medical informatics system using HL7 messaging standard protocols would support security in terms of authentication of both health practitioners and patients. For the identification of patients their NHIS ID card could be used. When a NHIS ID card is used a magnetic stripe reader could be used to display information about a member's eligibility on the computer thus offering a way to authenticate each patient before granting access to the medical information system. The Ghana health service holds a database of all health practitioners in Ghana. Therefore, it would be appropriate to let the Ghana Health Service (GHS) authenticate health practitioners before they are allowed to interact with the DMIS.

E. COMPLIANCE TO THE LAW

The implementation of the DMIS using HL7 messaging standard protocols will also take into consideration the laws that govern the exchange of medical data. For example it would be unlawful for AIDS patients' medical records to be stored in a database which is accessible by any health center since such records are highly confidential (just between the practitioner and the AIDS patient). To cater for the confidentiality of patients' records Access Control is used especially person involved is an AIDS patient.

F. ACCESS CONTROL

Maintains and enforces authorization rules, based on the identity of the requester. The access control mechanism is Role Based. This component allows patients to specify which details they wish to shield from specific practitioners, or categories of practitioners [11]. In the case of the DMIS for Ghana this component allows the AIDS patients to shield their AIDS related information from being accessed by other practitioners other than their own practitioners.

THE ARCHITECTURE OF THE (DMIS)

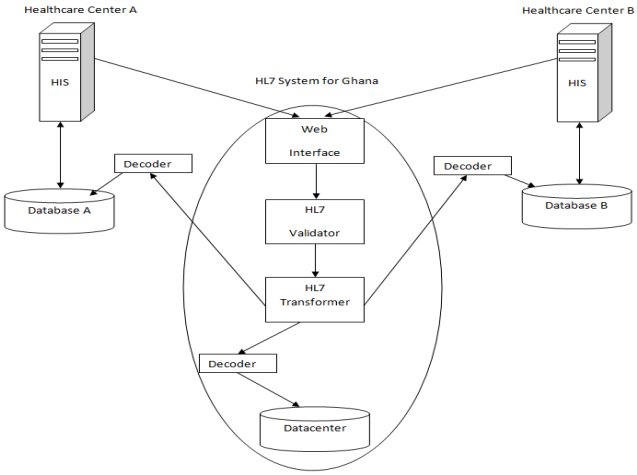


Figure 3. The Distributed Medical Information System

G. DMIS COMPONENTS

1) HIS (Health Information System)

The HIS automates the processes of capturing, organizing, and storing medical information. It is located at each health provider or center. The HIS is implemented using the client-server architecture whereby the HIS client provides an interface for the administrator and the patient to enter medical information into the DMIS (see figure 3). The HIS server provides functionality to the DMIS especially in terms of storage. The database is implemented using any database engine such as SQL, ORACLE, MySQL, etc… The communication between client and server is based on the hypertext transfer protocol (https). The HIS may be implemented using a variety of programming languages. The HIS must contain different modules that take care of the different aspects of the health process. For example the HIS may contain the patient demographic module which contains the graphical controls to add/edit/search demographic data of a patient. Using this module one can add/modify visits, document upload/download documents, add/modify transfers or other type of information (allergies, antecedents, and medical history), health insurance, and patient discharge.

2) HL7 System for Ghana

The HL7 system allows all the Health information systems from various health institutions to connect and share information. The HL7 system is made up of modules that manage messages and documents, and exchange them using a wide variety of transport mechanisms.
The HL7 system has the following modules:

3) Web Interface

This web application allows access to the datacenter. It is the portal through which the HIS clients and HIS administrators access health information stored in the datacenter.

4) HL7 Validator

The HL7 validator provides a complete set of rules that check whether the primitive type values in the information meet the constraints defined in the HL7 specification [12].

5) HL7 Transformer

Information systems communicate using different HL7 versions (Version 2 and Version 3) and different HL7 semantics (SNOMED-CT, ICD-10, LOINC, ETC…). The HL7 transformer performs complex HL7 message transformations so that one system's HL7 output becomes another's HL7 input.

6) Datacenter

The datacenter is a centralized database which receives and stores information from the various HIS at the various health institutions. The datacenter stores information that will allow patient identification in addition to medical information statistics that will be used for research purposes. The storage of patient identification information will contribute mainly to the creation of a Master Patient Index (MPI) which will improve healthcare delivery in Ghana.

H. ENCODING AND DECODING MEDICAL INFORMATION IN THE DMIS FOR GHANA

The encoding and decoding of medical information for data migration is the main focus (see figure 4). It is expected that the databases at the various healthcare institutions are designed using various database engines such as SQL, ORACLE, MySQL, etc. Also the datacenter which is a large centralized database can be designed using any database design engine. Exchanging medical information between databases with different designs becomes problematic. HL7 decoders solve the problem by decoding HL7 messages into XML based messages. XML is a language that is platform independent, that is XML based messages are readable by any database independent of the design and database engine. In the DMIS architecture, HL7 Version 3 messages are wrapped in SOAP (Simple Object Access Protocol) messages and transmitted between the various health information systems. SOAP is relatively lightweight, is based on XML, and is designed for the exchange of information in a distributed computing environment. There is no concept of a central server in SOAP; all nodes are peers. The protocol is made up of a number of distinct parts. The first is the envelope, used to describe the content of a message and some clues on how to go about processing it. The second part consists of the rules for encoding instances of custom data types. This is one of the most critical parts of SOAP: its extensibility. The last part describes the application of the envelope and the data encoding rules for representing remote procedure calls (RPC) and responses, including the use of HTTP as the underlying transport [13]. The encoding of HL7 messages takes place at the sending node and the decoding takes place after the HL7 messages have been validated and transformed.

V. DESIGN AND IMPLEMENTATION

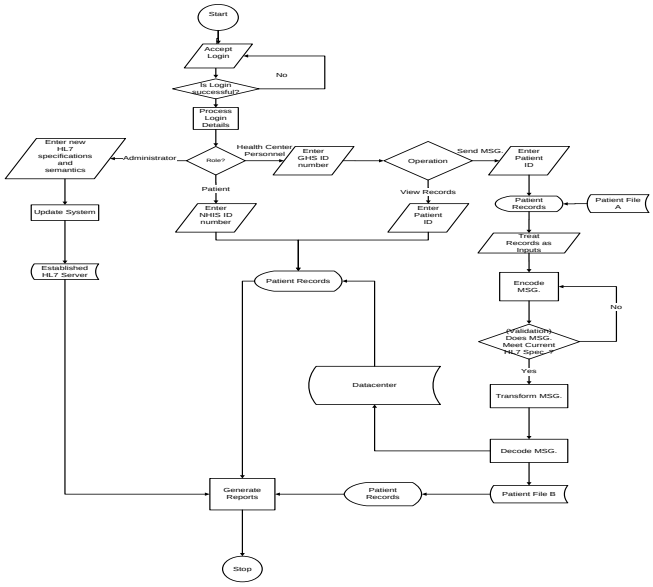


Figure 4. Flow Chart of the DMIS

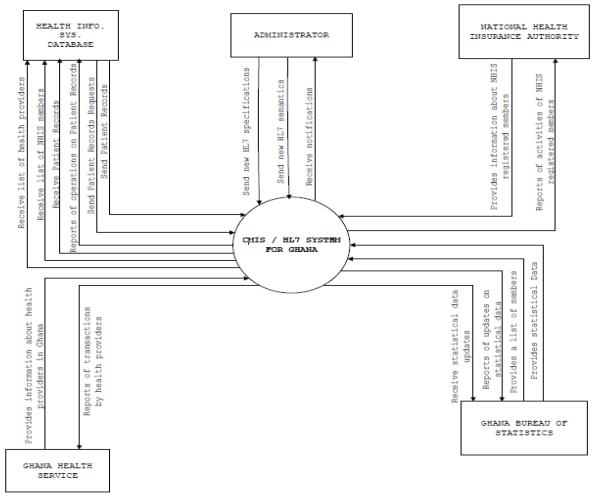


Figure 5: Context Level Data Flow of the DMIS

The context level diagram gives the broad abstraction of the entire system. This overview gives a quick view of how data flows in the DMIS. The level 1 data flow diagram shows the entities involved and how data flows between them. These are well illustrated in figure 5 and figure 6. The data diagram for DMIS displays the flow of data and processes that alter the data from information producers to information consumers.

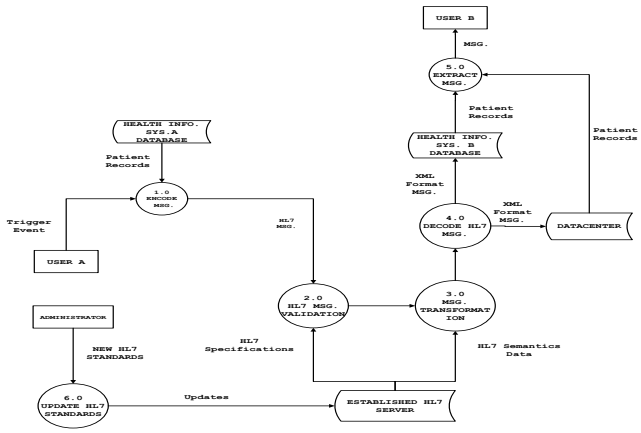


Figure 6: Data Diagram of the DMIS

A. IMPLEMENTATION

1) Forming of Message

First of all, an instance of the message type is created; from which all the segments will be referenced. SQL queries are used to retrieve the desired fields from the databases into the corresponding fields in the message structure. Below in figure 7 is a demonstration of how this is done.

```
//Instance of ADT_A01 message
NHapi.Model.V251.Message.ADT_A01 adt = new NHapi.Model.V251.Message.ADT_A01();
conn.Open();
{
//Instance of PID segment
PID pid = adt.PID;

//Assigning Fields from database into corresponding fields in the message structure
pid.GetPatientName(0).FamilyName.Surname.Value = reader1["PatientLastname"].ToString() + " " + reader1["PatientFirstname"].ToString();
pid.GetPatientIdentifierList(0).IDNumber.Value = txtpid.Text;
pid.DateTimeOfBirthTime.SetLongDate(DateTime.Now.AddDays(20));
pid.GetPatientAddress(0).City.Value = reader1["City"].ToString();
pid.GetPatientAddress(0).StreetAddress.StreetName.Value = reader1["StreetAddress"].ToString();
pid.GetPhoneNumberHome(0).TelephoneNumber.Value = reader1["PhoneNumber"].ToString();
pid.MaritalStatus.Text.Value=reader1["MaritalStatus"].ToString();
pid.Religion.Text.Value=reader1["Religion"].ToString();
pid.AdministrativeSex.Value=reader1["Sex"].ToString();



}

reader1.Close();
conn.Close();
```

Figure 7 Assigning the PID segment with data retrieved from the database

2) Validating the Message

After the message structure fields have been assigned with data, we are assured the HL7 message has been formed. But there is the need to validate this message to test its conformance to the HL7 syntax. To do this, nHapi makes use of the PipeParser Method to validate and check for conformance.

An instance of the PipeParser is created and then the variable is used to validate the HL7 message. In figure 6 is a demonstration of how this can be achieved.

```
PipeParser parse = new PipeParser();
string encodemessage = parse.Encode(adt);
string msh = encodemessage;
txthl7.Text = msh;
```

Figure 8 Validating the Message

From Figure 8, the PipeParser variable (parse) is used to validate the ADT_A01 HL7 message which has been represented by adt and then the encoded message is displayed in a textbox.

3) Creating the corresponding XML Message

After the message has passed valid for conformance, it is then expected that it be transformed into its corresponding XML format which is supported on all platforms. nHapi does this with the implementation of the DefaultXMLParser method. Figure 9, shows how this is performed.

```
DefaultXMLParser xml = new DefaultXMLParser();

encodemessage = xml.Encode(adt);
string strFilePath = encodemessage;
txtxml.Text = strFilePath;
```

Figure 9: Creating the corresponding XML Message

4) HL7 & XML Message Output

After message has been transformed and in both XML and HL7 formats, they are displayed in their respective textboxes. Figure 10 illustrates a sample message.

HL7

```
MSH|^~\&|Admission iHospital|VVU Hospital|Admit mHospital|Ridge
Hospital|20110510174422||ADT^A01|10||2.5.1|123
EVN|A01|20110510174422.5645
PID|||101||Boafo Kay||201105301744|Male|||&No.2^^Accra||9876543|||^S|^C
NK1|201|PJ Paa|^Brother|&P. O. Box 234^^Oyibi|985678909876
PV1|100|||R|||^Kwadwo^Debrah|^Dominic^Damoah|^PHILIP^MAIYO|LAB
TEST|||||||^ESTHER^BOAFO||100|||||||||||||||||||5||^^23||10/12/2010 12:00:00 AM
```

XML

```
<ADT_A01 ><MSH><MSH1>|</MSH1><MSH2>^~\&amp;</MSH2><MSH3>Admission iHospital</MSH3>
<MSH4>VVU Hospital</MSH4><MSH5>Admit mHospital</MSH5><MSH6>Ridge Hospital</MSH6>
<MSH7>20110510174422</MSH7><MSH9>ADTA01</MSH9><MSH10>10</MSH10><MSH12>2.5.1</MSH12>
<MSH13>123</MSH13></MSH><EVN><EVN1>A01</EVN1><EVN2>20110510174422.5645</EVN2></EVN>
<PID><PID3>101</PID3><PID5>Boafo Kay</PID5><PID7>201105301744</PID7><PID8>Male</PID8>
<PID11>No.2Accra</PID11><PID13>9876543</PID13><PID16>S</PID16><PID17>C</PID17></PID>
<NK1><NK11>201</NK11><NK12>PJ Paa</NK12><NK13>Brother</NK13><NK14>P. O. Box
234Oyibi</NK14><NK15>985678909876</NK15><PID>101</PID></NK1><PV1><PV11>100</PV11>
<PV14>R</PV14><PV17>Kwadwo  Debrah</PV17><PV18>Dominic  Damoah</PV18><PV19>PHILIP
MAIYO</PV19><PV110>LAB TEST</PV110><PV117>ESTHER  BOAFO</PV117><PV119>100</PV119>
<PV140>5</PV140><PV142>23</PV142><PV144>10/12/2010 12:00:00 AM</PV144><PID>101</PID>
</PV1></ADT_A01>
```

Figure 10 HL7 & XML Message Output

To query the XML document the OPENXML technology first declares a variable and loads the XML file into that variable- DECLARE @xmldata varchar(8000)

Then a pointer to delete the XML file from memory when the querying is over is also declared as DECLARE @hand int

After a pointer has been declared OPENXML selects the XML file from any location on the local machine and allocates it to the variable that was declared in the following code:

SELECT @xmldata= xmlData FROM (SELECT * FROM OPENROWSET (BULK 'C:\newMsg.xml', SINGLE_CLOB) AS xmlData) AS fileImport(xmlData)

Once the XML file has been stored in a variable, manipulations upon the file is now possible. To prepare the XML file for querying OPENXML uses this execute command: EXEC sp_xml_preparedocument @hand OUTPUT, @xmldata. After the execute command the rest of the commands are typical SQL commands. The first set of commands specify the table fields in which the data is to be inserted. For example: INSERT INTO MessageDetails. The second set of commands selects the XML document tag names whose values are to be picked. SELECT MSH3, MSH4, MSH5, MSH6, MSH7, MSH8, MSH9, MSH10 …

After the names of the element tags to be picked are specified the OPENXML technology is ordered to pick the values inside these tags. The command also specifies the root element tag and the direct parent element tag to the group of elements whose values are to to store into the HL7 database.

When OPENXML picks the values from the element tags in the XML document it has to specify the data types of the fields in the database that will receive the values as shown in figure 11.

```
WITH (MSH3 varchar(50), MSH4 varchar(50),MSH5 varchar(50),MSH6
varchar(50),MSH7 varchar(50), MSH9 varchar(50),MSH10 varchar(50),MSH12
varchar(50),MSH13 int)
```

Figure 11: A Snapshot of OPENXML Readings

The snapshots below show how OPENXML insert the values in the MSH segment of the HL7 message and stores them in the corresponding tables in the HL7 database. To store the other HL7 segments namely EVN, PID, NK1, and PV1. the process is followed using OPENXML into insert records into other segments. These are illustrated in figures 12, 13, and 14

.

```
INSERT INTO EventDetails

SELECT EVN1,EVN2 FROM openxml (@hand, '/ADT_A01/EVN' ,2)

WITH (EVN1 varchar(50),EVN2 varchar(50))
```

Figure 12: EVN Segment Querying using OPENXML

```
INSERT INTO PatientDetails

SELECT PID3,PID5,PID7,PID8,PID11,PID13,PID16,PID17

FROM openxml (@hand, '/ADT_A01/PID' ,2)

WITH (PID3 int,PID5 varchar(150),PID7 varchar(50),PID8 varchar(10),PID11
varchar(50),PID13 int,PID16 char(1),PID17 char(1))
```

Figure 13: PID Segment Querying using OPENXML

```
INSERT INTO PatientVisitDetails

SELECT PV11,PV14,PV17,PV18,PV19,PV110,PV117,PV119,PV140,PV142,PV144,PID

FROM openxml (@hand, '/ADT_A01/PV1' ,2)

WITH (PV11 int,PV14 char(1),PV17 varchar(150),PV18 varchar(150),PV19
varchar(150),PV110 varchar(50),PV117 varchar(150),PV119 varchar(50),PV140
int,PV142 int,PV144 varchar(50),PID int)
```

Figure 14: PV1 Segment Querying using OPENXML

VI. RECOMMENDATIONS

HL7 is a medical information messaging protocol that is being adopted and implemented as the world standard for medical information exchange. In spite of its widespread HL7 remains almost unknown in sub-Saharan Africa.

The first recommendation is to educate the health sector stakeholders about the benefits of implementing a standard in exchanging medical information.

The second recommendation is to build upon the work done to finally implement HL7 system for Ghana for the numerous benefits it represents for an improved patient care in this country.

The benefits are as follow:

- High accurate accessibility of medical reports.
- High Patient medical records availability.
- Cost-Effectiveness and Low overhead cost of transportation and communication of healthcare data.
- Interoperability-No need to rule out old systems, since they can be integrated to HL7 system.
- Its universality allows collaboration in medical research with medical systems in the developed world.
- Integration at the intra-hospital level and also at the inter-hospital level

Last sending and receiving of HL7 messages conformant to the standard is just one face of the successful implementation the HL7 standard in Ghana. This represents the inter-hospital communication. The other face that is equally important is the intra-hospital communication which will enable data in each hospital to be organized and stored in a consistent manner. So when there is a need to communicate between hospitals the process will be very effective since effective inter-hospital communication requires the existence of a good underlying intra-hospital communication. Therefore, the recommendation is that emphasis should also be put on the development of effective hospital management systems in order to provide a strong basis for inter-hospital communication thus improving patient care in Ghana and pioneering HL7 establishment in the Sub-Saharan Africa.

This paper resulted in the successful implementation of the ADT (Admission, Discharge, and Transfer) message type using

the version 2 of HL7. Therefore further work is to be undertaken in implementing other message types and adopting the latest version of the HL7 standard.

REFERENCES

[1] Ministry of Health, HIV/AIDS Unit, Health Management Information Unit, "Anti Retroviral Therapy Message Profiling Specification for HL7 Standard in Malawi.," no. v1.3, 2007.

[2] C. A. Orza Bogdan, " Integrated medical system using DCOM and HL7 standards. New Advanced Technologies," pp. 231-246, 2010.

[3] Republic of Ghana, Republic of Ghana, 1992 Constitution, Article 36/10, Accra, 1992.

[4] Health Economics and Policy Network in Africa:, The Implementation of the National Health Insurance Scheme (NHIS) Ghana: Experience Sharing,, 2007.

[5] MINISTRY OF HEALTH, HEALTH MATRIX NETWORK., "LEGAL AND POLICY FRAMEWORK FOR HEALTH INFORMATION AND HEALTH DATA REPORTING," Accra, 2008.

[6] V. P. Abdallah Salisu, "HEALTH CARE IN GHANA," Accra, 2009.

[7] Health Metric Network:, Ghana health information system: Profile and summary preliminary assessment,, 2005.

[8] M. S. Mensah, "STATUS OF NHIS AND NEW STRATEGIC DIRECTION.," pp. http://moh-ghana.org, 2009.

[9] K. Powell, " What is Medical Informatics?," 2010, pp. http://www.wisegeek.com/what-is-medical-informatics.htm.

[10] Spiritus-Temporis.com , "Bus network," pp. http://www.spiritus-temporis.com: , 2005.

[11] R. Spronk, "AORTA, the Dutch national infrastructure. Ringholm Whitepaper .," 2008.

[12] M. Krasser, "HL7 processing tutorial," p. http://repo.openehealth.org: http://repo.openehealth.org/confluence/display/ipf2/HL7+processing+tutorial, (2010, 03 29). .

[13] R. Englander, JAVA AND SOAP, O' Reilly., 2002.

E-Commerce Security Model Approach in Mobile Agents

Solanke Ilesanmi O.
Department of Computer Technology
Yaba College of Technology
Yaba Lagos, Nigeria
*solankesanmy@gmail.com

Adetoba Bolaji T.
Department of Computer Technology
Yaba College of Technology
Yaba Lagos, Nigeria
tiwabj@yahoo.com

Omotunde Ayokunle A
Department of Computer Science
Babcock University
Ilisan Remo, Ogun State, Nigeria.
ayo_omotunde@yahoo.com

Abstract-With the advancement and the rapid growth in information and communication technology, the internet, coupled with the increased in the complexity of network environment has brought about a greater level of uncertainty and insecurity in an e-commerce environment. An enterprise wants to keep their customers within their reach by establishing secured e-commerce transaction methods that will improve on the traditional methods of doing business. However, for the application of mobile agent to be successful in e-commerce environment, the issue of security of mobile agent needs to be addressed. This paper explores some e-commerce and mobile agent security and integrity threats. Also, this study presented mobile agent encryption algorithm and a secured e-commerce model for mobile agent based on trusted third party.

Keywords: e-commerce, mobile agents, security, host.

I. INTRODUCTION

As a result of the introduction of mobile technology and platform communication, language exchange has given a new era to business to realizing a better approach to e-commerce. The mobile Agent is a new technology in distributing computing. The application of mobile agents provides advantages in arrears such as responsiveness, adaptability, convenience, mobility and independence; it enhances the search of resources online and online automated trading. Moreover, the issue of security in an e-commerce environment cannot be overemphasized, it needs an urgent security issues to be resolved in the e-commerce research study.

This paper firstly explores some e-commerce security threats, challenges and also the integrity of mobile agent's code and data when moving through their path especially in a platform that

Cannot be trusted. Also, this paper presented an algorithm for a mobile agent message encryption and a secured e-commerce model for mobile agents based on trusted third party. The study also shows that e-commerce transaction risk and malicious interference in mobile agent's path can be reduced to the barest minimum if not totally avoidable.

II. ISSUES OF SECURITY IN MOBILE AGENTS

There are so many uncertainties with the issue of mobile agents; a widely acceptance and successful application of mobile agents in e-commerce will help to enhance security issues. The structure of mobile agents comprises of three basic components: Code Application Segment, data application, status application segment. For the status application segment, it includes the agent's data logs and current implementation status. The security of mobile agent exist within the following context, whether mobile agents has a strong protection against malicious attack in their host which comprises of integrity of the implementation process, data agents and mobile agent code application. Second, whether the function of the host can be hidden by mobile agents on e-commerce security platform for the protection of privacy of mobile agents. Lastly, whether a private key can be signed and authenticated remotely without revealing the identity of the owner.

A. ISSUES OF SECURITY AND INTEGRITY OF MOBILE AGENT HOST

The host has absolute control on mobile agents. As a result of the activities of the host, it is difficult to avert sabotage activities that may be conducted through a malicious host. The malicious activities and threats are integrity threats, confidentiality threats, availability threats, and repudiation threats [3]. In this approach of securing mobile agent code and data, the security and integrity of the mobile agent on host will be protected and secured by allowing the owner of the mobile agent to verify how correct the implementation of the mobile agent is and return the verified correct and complete data. Based on the verification by mobile agents, the mobile agents can detect malicious host or behavior and issue warning, feedback report and avoid counter attack. The integrity protection approach for mobile agents can be categorized into the following:

1). *Digital Signatures*: The digital signatures can detect interference on the mobile agent data or code. This technique makes a signed code of mobile agents detectable when there is interference or when a signed code is tampered with [3]. Sometimes, it can be difficult to provide appropriate security for mobile agent data especially when they pass through malicious platform. However, it is possible to detect a mobile agents code and data that has been interfered with through the use of a trusted platform such as digital signatures. A secured and encrypted digital signature can be attached to the mobile agent code and data, this approach will make the mobile agent platform to be able to verify the integrity of the agent code and authenticate agent owner [7]. Also the generated code can detect if mobile agent code and data has been interfered with after reporting back to the sender's trusted environment or platform.

2). *State Assessment*: it plays a major role in the security of mobile agent's code and data by ensuring that changes in mobile agent executions are monitored. This approach was designed for multi-hop agent architecture, in which an agent moves in different path and through several platforms in order to perform their task [3]. Further damages and interference with mobile agent's code and data can be aborted or minimized through detection of malicious platform by state assessment [3]. A state assessment is used to implement a set of privileges to be requested for based on the state of the mobile agent's code and data when it arrives at a new platform. When a state assessment is implemented, malicious modification to the state of the mobile agents code or data can be detected and the appropriate counter actions taken.

3). *Register Path History of Mobile Agent*: The registering of the path taken by mobile agents is an approach that is used for keeping records and learning about mobile agent path logs with respect to the platform they visited. Registering path history of mobile agents assist in enforcing the integrity of the agent migration path, and this approach also assist in identifying a host that has been interfered or tampered with during migration. An improved security protocols will reduce attack through the application of distributed or through tracking of the migration path of mobile agent which is done through the authorization of a trusted third party using signatures [3].

4). *Fault Tolerance*: Mobile agent's data roam on uncertain networks. During the movement of a mobile agent, many unpredicted issues may arise such as server Error, Communication breakdown. The more the mobile agents move through its path, the more susceptible it is to malicious act. Making mobile agents fault tolerance is an important task, with this it can quickly detect error and make recovery, it can implement a primary back up, replicated agents with voting by distributing out mobile agents with the same initial state, and after executing the same task, they can vote to get result [1].

5). *Result Encapsulation:* The result encapsulation can be used for the confirmation and verification of computational results for all the platforms that has been visited by the mobile-agents [3]. The result of each platform visited by mobile agents should be encapsulated so that attack or interference can be traced back to a particular platform. Irrespective of the mobile agent structure, encapsulation can be implemented in order to encrypt messages moving through the path for confidentiality. Secondly, attaching digital signatures and message authorization code in order to maintain integrity of the messages sent across mobile agent path. Thirdly, encapsulate partial results with time record and digital signatures through a trusted third party. This approach also helps to verify the results returned by mobile agents from all the path and destination that the agent has visited.

B. MOBILE AGENTS PROTECTION TECHNIQUES AND TRUSTED THIRD PARTY SECURITY POLICY.

Mobile Agent should be protected against malicious platform. The following techniques can be used to build trust and making sure that security and integrity of mobile agents are maintained on their path, and also confidentiality can be maintained by making sure mobile agents are not compromised.

1). *Separation of Right:* This approach can be used to split task between mobile agents by granting them separate rights and privileges. This minimizes the malicious damages that may be done to the mobile agents. The most critical components of the mobile agents task are executed by agents on trusted platform, while other task that are less-critical are performed by agents sent to platforms that cannot be trusted [3]. Although separation of critical task platform from non-critical platform does not eliminate attack or compromise of the mobile agents.

2). *Encryption Function Evaluation*

This approach is theory base. The essence is to allow mobile Agent moving in a Trusted Path even when the Trusted Path (TP) is not without the doubt of malicious compromise or untrusted activities. This approach can give the implementation process of the TP necessary instruction through the cryptographic function method.

2.1) *Mobile Agents Encryption Function Algorithm*

This approach can be analyzed as follows:

S: Sender of message
M: Message sent.
E: Encryption Agent of the sender.
D: Destination of message.
T_p : Trusted Path of the message

- M sends messages M(a,b) using function f(x,y) to encrypt (a,b) as (e,j) because M does not want D to view the content of the message sent.
- D received messages (a,b) represented and encrypted as (e,j) and returned same as (e,j) to M.
- M decrypts the message (e,j) returned by D, and decrypt the message using function f(x,y) and gets (a,b).

For this approach, if the function of f(x,y) which has been used to encrypt message (a,b) as (e,j) has a digital signature, then Mobile Agents can encrypt information without Trusted Path unaware.

3). *Limited Privileges:* Most mobile agents has a private key that are encrypted with a symmetric key, the symmetric key can be divided into several key fields in order to secure and deliver the mobile agent data. Limited privileges or limited authorization allows a mobile agent data collect certain number of keys before authorization can be granted.
Figure 1 illustrates a secured e-commerce model for mobile agents based on trusted third party. It is designed to secure the agent from threat of a malicious host.

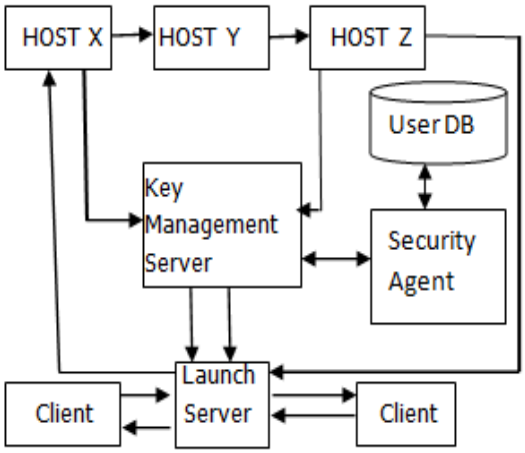


Fig. 1 Flowchart of secure e-commerce channel model

- The client makes a launch request to the key management server.
- The launch server initiates a query that sets agents access path.
- Launch server generates keys.
- Launch server grant request and register the mobile agent public key on the key management server.
- Launch server approves the movement of mobile agent to the network.
- HOST X requests the public key of the mobile agent from the key management server which will be retrieved from the database through the security agent query.
- Mobile agent moves through its path to HOST Y and Z and repeat the same process in X.
- Mobile agent returns through its path to the launch server, and then the launch server decrypts the results that has been returned to it and verifies the digital signatures of the returned results.
- Launch server verifies if the mobile agent's path has been tampered with for the purpose of integrity and report back to the clients.
- After the process, the public key will be deleted by the launch server which makes the whole process on the key management server complete.

In this model approach, the use of SSL (Secured Socket Layer) protocols can be used to ensure the security of the information and migration of mobile agents from one channel to another through a secured path. The longer the encryption keys, the more complex it becomes, and this makes it to become harder to be broken into my malicious host, this makes the mobile agent's environment more secured and reliable.

III. CONCLUSION

The mobile agent has contributed immensely to the development of e-commerce both in transaction and communication. As a significant distributed technology, it is an independent executable program that can move independently in a heterogeneous network environment in accordance with certain set of networking protocols. This study has revealed some challenges faced by mobile agents, it also discussed some of the difficulty in protecting mobile agents against malicious platform or interference. The continuous refinement of e-commerce security based on mobile agent will continue to enhance the performance and security of e-commerce and also contribute a significant improvement to the future development.

REFERENCES

[1] Li, Yan, Min Fu, and Lina Yu. "E-commerce security model construction based on Mobile Agent." In *Networking and Digital Society (ICNDS), 2010 2nd International Conference on*, vol. 2, pp. 55-58.2010.
[2] Wei, Dai, and Ji Wei. "Research on the Security of an Improved E-commerce Model." In *E-Business and E-Government (ICEE), 2010 International Conference on*, pp. 2534-2537, 2010.
[3] Ssekibuule, Richard. "Mobile agent security against malicious platforms." *Cybernetics and Systems: An International Journal* 41, no. 7 (2010): 522-534.
[4] Sander, Tomas, and Christian F. Tschudin. "Protecting mobile agents against malicious hosts." In *Mobile agents and security*, pp. 44-60. Springer Berlin Heidelberg, 1998.
[5] Ganesan, Ramachandran, and Kanniappan Vivekanandan. "A Novel Hybrid Security Model

for E-Commerce Channel." In *Advances in Recent Technologies in Communication and Computing, 2009. ARTCom'09. International Conference on*, pp. 293-296, 2009.

[6] Li-xiao, Geng, Zeng Zhen-Xiang, and Zhang Xue-Min. "Research on PKI-Based E-Commerce Security Mechanism." In *Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007. International Conference on*, pp. 3545-3548,2007.

[7] Ghesmati, Simin, Majid Sate, and Abbas Asosheh. "A 2-level model for e-commerce security." In *Internet Security (WorldCIS), 2011 World Congress on*, pp. 263-268.2011.

[8] Jing, Wang, and Qin Bingwen. "The security model design for collaborative mobile e-commerce based on wireless middleware." In *Information Science and Engineering (ICISE), 2010 2nd International Conference on*, pp. 2471-2474, 2010.

[9] Xu, Bing, and Shiyi Xie. "Research of session security management in E-Commerce system." In *Information Engineering and Electronic Commerce, 2009. IEEC'09. International Symposium on*, pp. 796-799, 2009.

[10] Tian, Zhaofu, Ningning Xu, and Wuliang Peng. "E-Commerce Security: A Technical Survey." In *Proceedings of the 2008 Second International Symposium on Intelligent Information Technology Application-Volume 02*, pp. 956-960.2008.

Fingerprint Analysis – A Framework for Gender Classification

Komolafe Oyindolapo
Department of Computer Science
Babcock University, Nigeria
oyindolapokomolafe@yahoo.com

Onuiri Ernest
Department of Computer Science
Babcock University, Nigeria
ernestonuiri@gmail.com

Abstract— Security remains a challenge in the global community and as information technology continues to extend human capacity, by means of information accessibility and dissemination, our tendency to commit all manner of crime also increases. This leaves security agents, especially those who rely on forensics to have their work cut out, in the process of investigation. One novel method of fingerprinting is that which pertains to gender classification, an important step in forensic anthropology, used to identify the gender of a suspected criminal. This, in turn helps to streamline the search to a particular gender as well as enhancing the efficiency with which the investigation is carried out. This research proposes the combination of principal component analysis, singular value decomposition and discrete wavelet transforms in the framework for gender classification. It is expected that results obtained from this combination will be more precise and efficient in determining gender based on fingerprint analysis.

***Keywords*: gender classification, principal component analysis, singular value decomposition, discrete wavelet transforms, fingerprinting**

I. INTRODUCTION

Fingerprinting constitute a trusted and tested biometric technology, considered legitimate proof of evidence in courts of law all over the world. Based on the varieties of the information available from the fingerprint, it is possible to process the identity along with gender, age and ethnicity of a subject [1]. With increasing need for reliable security mechanisms, identification and authentication methods have evolved into a key technology. Fingerprints are constant and individualistic and form the most reliable criteria for identification [2].

A fingerprint is an impression of the friction ridges of all parts of the finger. A friction ridge is a raised portion of the epidermis on the digits or on the palmar and plantar skin, consisting of one or more connected ridge units of the friction ridge skin. Fingerprint identification algorithms are well established and are being implemented all over the world for security and person identity. Very few attempts have been made to classify the gender from an obtained fingerprint. This is helpful for anthropologists, for classifying gender from the fingerprints obtained from excavated articles and for crime investigators for minimizing the range of the suspects [3].

Fingerprints collected at a crime scene can be used to identify suspects, victims and other persons who touched the surface, fingerprint scans can be used to validate electronic registration, cashless catering and library access especially in schools and colleges. Fingerprint verification and identification algorithms can be classified into two categories: image-based and minutiae-based. Image-based methods include methods involving optical correlation and transform-based features. The gender of the person can be judged using the fingerprint of that concern person based upon the count of the ridges of the fingerprint. The age and ethnicity can also be determined [4]. Within a single individual, the breadth of finger print ridges varies within the hand and between hands, but the differences are quite small on the order of 0.05mm and less. The gender of each person must be based on different combinations of these features ridge count, ridge thickness to value thickness ratio (RTVTR), fingerprint pattern type, white lines count, pattern type concordance between the corresponding left-right fingerprints, and ridge count asymmetry between the left-right corresponding fingerprints. Male's and female's fingerprints are characterized by an average rightward asymmetry in the ridge count [4]. The average ridge count is slightly higher in males than in females, with high standard deviation among subjects of both genders [3].

The performance of a fingerprint feature extraction and matching algorithm depends critically upon the quality of the input fingerprint image. While the 'quality' of a fingerprint image cannot be objectively measured, it roughly corresponds to the clarity of the ridge structure in the fingerprint image. This paper focuses on the combination of principal component analysis (PCA), singular value decomposition (SVD) and discrete wavelet transforms(DWT) in the framework for gender classification to produce a more precise and efficient result for gender classification.

II. LITERATURE REVIEW

A system that used Support Vector Machines (SVM) for classification of features extracted was found reliable. The features are ridge count, RTVTR, white lines count, ridge count asymmetry and pattern type concordance. The concept of SVM was based on decision planes that defines decision boundaries. The system was tested with four fingerprints, the prints were converted into binary image which contains only

two values of either 0 or 1. The binarized image was skeletonized and later thinned. The ridge thickness and valley thickness was calculated, and the average ratio was calculated. It was discovered that females have high ridge counts than males. The results confirmed a difference in fingerprint image quality across age groups. Inherent feature issues, such as poor ridge flow, and interaction issues, such as inconsistent finger placement, have an impact on captured fingerprint quality, which eventually affects overall system performance [4].

A real time system whose input is fingerprint digital image also suffices. Discrete Wavelet Transform (DWT) and Principal Component Analysis (PCA) were used to classify gender. The fingerprint digital image was subjected to preprocessing for enhancement and to remove noise. The processed image undergoes two level of feature extraction: wavelet decomposition for obtaining frequency domain feature and PCA for the spatial level. The 2D DWT was used to get images in results in four decomposed sub-band images referred to as low–low (LL), low–high (LH), high–low (HL), and high–high (HH). With each sub-bands representing different image properties. The feature vectors obtained from DWT and PCA are combined and a graph is plotted. The graph showed a demarcation between Male and Female prints meaning to say that there are differences between male and female fingerprints even when analyzed using frequency domain analysis. The gender classification approach used was Minimum Distance Classifier which uses the Euclidian distance measure for finding the distance between the query and the database. The result has a 70% overall success rate [3].

A novel method of gender classification using level 6 DWT and SVD is also worthy of mention. The method considered the frequency features of the wavelet domain and the spatial features of the singular value decomposition. The spatial features include the internal structure of the fingerprint images and the fusion of the features with the frequency features, produced improved performance in gender classification. The level 6 DWT was selected as optimum level for the gender classification by analyzing the results obtained for the database used for training and testing and the database used other than the training and testing. The classification is achieved by extracting the energy computed from all the sub-bands of DWT combined with the spatial features of non-zero singular values obtained from the SVD of fingerprint images. K-Nearest Neighbor (KNN) was used as a classifier. This method was experimented with the internal database of 3,570 fingerprints in which 1,980 were male fingerprints and 1,590 were female fingerprints. Finger-wise gender classification was achieved at 94.32% accuracy for the left hand little fingers of females and 95.46% for the left hand index fingers of males. Gender classification for any finger of males tested was attained at 91.67% and 84.69% for females respectively. Overall classification rate of 88.28% was achieved [1].

The analyses of the effectiveness of physical biometrics (thumbprint) in order to determine age and gender in humans was carried out in this research**.** The application was created using Microsoft visual studio integrated development environment for creating, documenting, running and debugging programs written in a variety of .NET programming languages. A fingerprint scanner was interfaced with the system to collect fingerprint images. Multi-Layer Perceptron (MLP) model technique was used, and results were obtained for age & gender, RTVTR & Ridge Count for trained finger prints that were sampled (100 females and 100 males). It was observed that the females had a higher ridge thickness to valley thickness ratio compare to males. The Males had a slightly higher ridge count compared to females, there is no particular relationship between the age of subjects and their fingerprint pattern, as it does not change (only as a result of accident or mutation) [8].

Another novel method to estimate gender by analysing fingerprints is with the use of fast Fourier transform (FFT), discrete cosine transform (DCT) and power spectral density (PSD). The fingerprint image is given to the database identification system, FFT transforms the given input and generates the output, then DCT transforms the given input and generates the output and PSD transforms the given input and generates the output. Comparing the decisions by all the transforms, if two decisions are male, the result is announced as male and if two decisions are female, the result is announced as female. 110 males and females were tested with the system, it was discovered that the proposed algorithm produced an accurate decision of 90% for female and 79.07% for male [9].

III. METHODOLOGY

The fingerprint verification will be divided into two parts: minutiae extraction and minutiae matching.

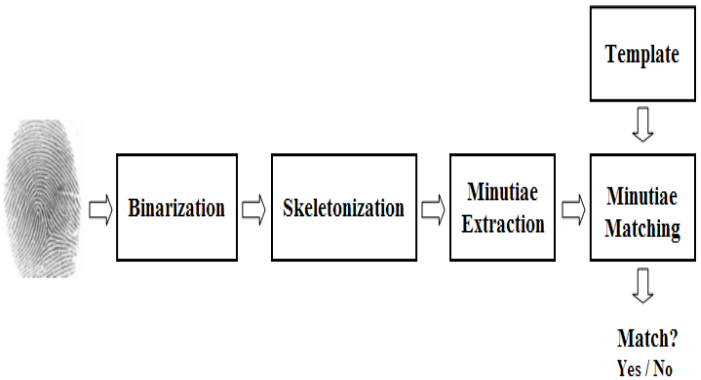


Figure 1: system design source: [4]

Minutiae extraction

The extracted fingerprint images from the input device are sometimes of poor quality because of variation in impression conditions, ridge configurations, skin conditions,

acquisition devices and non-cooperative attitude of subjects [5]. Image preprocessing techniques that includes segmentation, thinning, noise reduction, image enhancement and error correction are used on the raw image to get enhanced image.

Segmentation: segmentation is the process of breaking an image into groups, based on similarities of the pixels. The foreground region is the area containing the ridges and valleys in the fingerprint. A variance threshold is done to discard the background regions that are of no interest in the minutiae extraction.

Noise reduction: In order to reduce the noise pixel-wise wiener filtering is proposed. The filter is based on the estimated local statistics from a local neighborhood $'a'$ of size 3×3 of each pixel, and is given by equation 1 below:

$$w(m_1,m_2) = r + \frac{\mu^2 - \sigma^2}{\mu^2}(I(m_1,m_2) - \tau) \quad \text{Equation 1}$$

Where σ^2 is the noise variance, τ and μ^2 are the local mean and the variance respectively, the grey level is represented in $(m_1,m_2)\epsilon\, a$

Image enhancement: histogram equalization is used for contrast expansion because it defines a mapping of gray levels q into grey level p. the distribution of gray level p is uniform.

Binarization: image binarization converts an image of up to 256 gray levels to a black and white image. Binarization is done through filtering the fingerprint image in spatial domain with specially designed filters whose orientation matches the local ridges orientation in the fingerprint. After the operation, ridges in the fingerprint are highlighted with black color while furrows are white.

PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA was invented in 1901 by Karl Pearson. PCA involves a mathematical procedure that transforms a number of possibly correlated variables into a number of uncorrelated variables called principal components, related to the original variables by an orthogonal transformation. This transformation is defined in such a way that the first principal component has as high a variance as possible, and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to the preceding components. PCA involves the calculation of the Eigen value decomposition of a data covariance matrix or singular value decomposition of a data matrix, usually after mean centering the data for each attribute [7].

PCA algorithm is used on the fingerprint image, to obtain a feature vector in spatial domain. Every fingerprint in the database undergoes the PCA for obtaining the eigenvector.

DISCRETE WAVELET TRANSFORM (DWT)

The fingerprint image will undergo discrete wavelet transformation for obtaining the feature vector. DWT has been frequently used in feature extraction and it is a suitable tool for a classification system because it decomposes images into sub-bands that are localized in frequency and orientation and the decomposition permits the isolation of images into different frequency ranges. DWT is used for feature extraction, de-noising, compression, face recognition, and image super-resolution. DWT has a complete theoretical framework and great flexibility for choosing bases and the low computational complexity [6].

The energy of all the sub-band coefficients is used as feature vectors individually which is called as sub-band energy vector (E). The energy of each sub-band is calculated by using equation 2.

$E_k = \frac{1}{RC}\sum_{i=1}^{R}\sum_{j=1}^{c}|x_k(i,j)|$ - Equation 2, where $x_k(i,j)$ is the pixel value of the kth sub-band and R and C is width and height of the sub-band respectively.

SINGULAR VALUE DECOMPOSITION (SVD)

The Singular Value Decomposition (SVD) is an algebraic technique for factoring any rectangular matrix into the product of three other matrices.

GENDER CLASSIFICATION WORKFLOW

Steps to be taken for gender classification entails the following:

1. The fingerprint image undergoes preprocessing of image enhancement, binarization and noise reduction.
2. The processed fingerprint image undergoes Wavelet Decomposition of level 6 and stored.
3. The fingerprint image from the preprocessing stage also undergoes PCA Eigen-vector feature extraction.
4. The two results of DWT and PCA are combined to get the features for a particular fingerprint
5. This fingerprint feature vector is classified using the minimum distance classifier and this undergoes the gender classification.
6. Calculate the RTVTR.
7. Subject the fingerprint image from the preprocessing to SVD.
8. Compare the results of SVD, RTVTR and DWT and PCA to determine the gender.

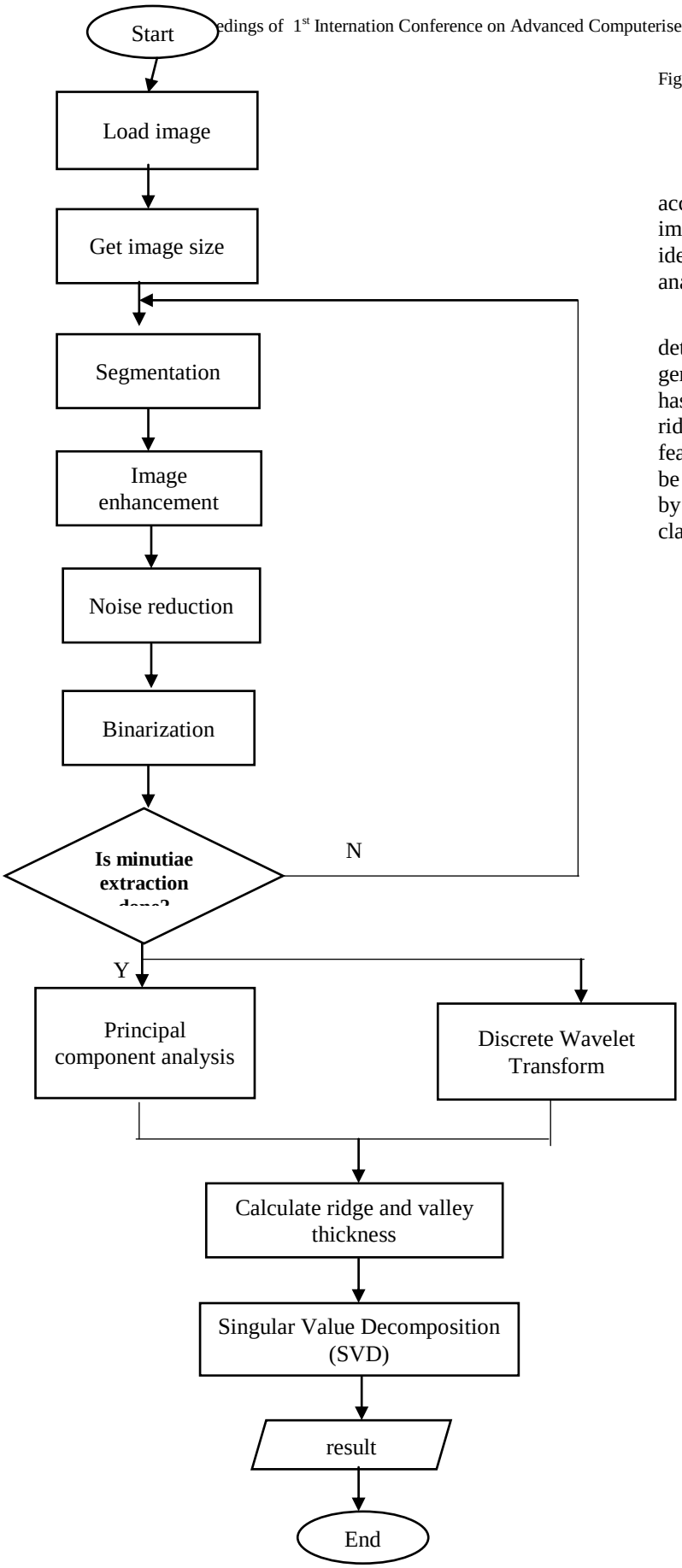


Figure 2: Flowchart of the fingerprint analysis process

IV. CONCLUSION

Fingerprint evidence is undoubtedly the most reliable and acceptable evidence till date in the court of law. Due to the immense potential of fingerprints as an effective method of identification an attempt has been made in the present work to analyze their correlation with gender of an individual.

The traditional methods uses ridge related parameters to detect gender and not much work has been done to detect gender using frequency domain analysis. Hence, an attempt has been made by proposing a method that utilizes both the ridge parameters, frequency domain analysis and spatial features of the fingerprints. The above proposed method can be implemented using MATLAB. The work can be improved by adding neural network algorithms and fuzzy logic to the classification of gender.

REFERENCES

[1] Gnanasivam P. & Muttan. S. (2012): "Estimation of Age through Fingerprints Using Wavelet Transform and Singular Value Decomposition" International Journal of Biome Biometrics and Bioinformatics (IJBB), Volume (6): Issue (2): pp 58-67

[2] Pillay, V.V (2009). Textbook of Forensic Medicine and Toxicology. 15th ed. Hyderabad: Paras Medical Publishers, 2009: 53-94.

[3] Tom R.J and Arulkumaran T,(2013): Fingerprint Based Gender Classification Using 2D Discrete Wavelet Transforms and Principal Component Analysis

[4] Ponnarasi S.S. and Rajaram M (2012): Gender Classification System Derived from Fingerprint Minutiae Extraction. *International Conference on Recent Trends in Computational Methods, Communication and Controls (ICON3C 2012) Proceedings published in International Journal of Computer Applications® (IJCA)*

[5] Meenakshi A, Gaganpreet K and Ravi S.R (2012): Minutiae feature based algorithm for fingerprint recognition and Verification. VSRD international journal of Electrical, electronics and communication engineering VSRD-IJEEECE, Vol 2(5), 279-287

[6] Bai-Ling Z, Haihong Z., and Shuzhi S.G (2004), "Face Recognition by Applying Wavelet Subband Representation and Kernel Associative Memory", IEEE Transactions on neural networks, vol. 15, no. 1, 2004, pp.166-177.

[7] Dean, T., Allen, J., and Aloimonds, Y. (1995), *Artificial Intelligence: Theory and Practice*, Addison-Wesley, Reading, MA, pp. 409-469, 1995.

[8] Ojo O, Omidiora E.O., Yekini N.A, Tubi T.O (2012): Analysis, Design and Implementation of Human Fingerprint Patterns System "Towards Age & Gender Determination, Ridge Thickness to Valley Thickness Ratio (RTVTR) & Ridge Count on Gender Detection. *(IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 1, No. 2, 2012*

[9] Kaur R. and Mazumbar S.G (2012): Fingerprint Based Gender Identification Using Frequency Domain Analysis. International Journal of Advances in Engineering & Technology, ©IJAET ISSN: 2231-1963

Gait Recognition for Real Time Video Surveillance

Akinyokun O.C. and Akintola K.G.
Computer Science Department,
Federal University of Technology Akure,
Ondo State, Nigeria

***Abstract*-** This paper presents a framework for the automatic detection and recognition of human beings from video cameras via smart visual systems that automatically sense and correctly recognize human using gait by means of machine vision techniques. Such systems require low response time in terms of image processing and acceptable recognition accuracy. Initial human detection is addressed by background subtraction techniques using parallel–processed Kernel Density Estimation (PKDE). Temporal tracking of the objects' trajectories is performed by employing a spatial body tracking system designed as a multi-part colour histogram-based tracker. In gait recognition, the Principal Component Analysis (PCA) algorithm is implemented. An experiment was performed in computer laboratory of Federal University of Technology Akure where a camera was installed in the Laboratory to capture students while entering the lab. The gait detection algorithm performs well and reduces the computational time. Some future areas of practical application of such system include access control to facilities like lecture rooms, Automated Teller machines, and attendant management systems.

Keywords: PCA, KDE, A Haar-like feature

I. INTRODUCTION

The need for securing physical properties and assets is becoming increasingly important. Recently technology became available to allow verification of the true identity of criminals. This technology is called Biometrics. Biometric access control is automated methods of verifying or recognizing the identity of a living person on the basis of some physiological and behavioral characteristics. Recognizing gaits in videos is a fundamental for intelligent vision-based systems for human monitoring, identity recognition and activity analysis. To be able to recognize humans, in a surveillance scenario, robust, efficient and fast gait detection and recognition algorithms are required. Motion information is proposed in this work for fast gait detection using the Principal Component Analysis (PCA) is used for gait recognition.

This paper describes a solution for human identification using biometrics designed for use in facility access control. The gait recognition system is based on the eigengaits method introduced by Turk and Pentland (1991). Eigenvector-based methods are used to extract low-dimensional subspaces which tend to simplify tasks such as classification. The system that will be used in access control to facilities should be able to robustly detect and recognize people based on their gaits.

Eigengaits method is the implementation of Principal Component Analysis (PCA) over gait images. The Eigengaits method tries to find a lower dimensional space for the representation of the gait images by eliminating the variance due to non-gait images. In this method, the features of the studied images are obtained by looking for the maximum deviation of each image from the mean image. This variance is obtained by getting the eigenvectors of the covariance matrix of all the images.

II. APPROACHES TO HUMAN RECOGNITION BY GAIT

There are two major approaches used for gait recognition.

a. Model-based Approaches. This usually involves modeling the gait data collected from walking person for identification. Spencer and Carter [2002] use angles between two joints to model human gait. Nowlan [2009] adopts a method that exploits the biometric property of human gait to identify people. A single sensor composed of an accelerometer and gyroscope is used to record gait characteristics. Using a Direct Form II Transpose, gait cycles are extracted and compressed to their characteristic features. A classifier is then used for the gait recognition.

In Bobick and Davies [2001], a gait-recognition technique that recovers static body and stride parameters of subjects as they walk is presented. This approach is an example of an activity-specific biometric: a method of extracting identifying properties of an individual or of an individual‟s behavior that is applicable only when a person is performing that specific action [Bobick and Davies, 2001]. The set of static body parameters measured as a person walks, are four distances: the vertical distance between the head and foot (BP), the distance between the head and pelvis (BL), the distance between the foot and pelvis (B..), and the distance between the left foot and right foot are used. A motion-capture system that uses magnetic sensors to

capture the three-dimensional position and orientation of the limbs of the subjects as they walk along a platform is utilized.

In Zhang and Vigjer [2007], Human Gait Recognition System at Saggital Plane is presented. In the project, a novel 2-step, model-based approach to gait recognition by employing a 5-link biped locomotion human model was proposed. The gait features were first extracted from image sequences using the Metropolis-Hasting method. Hidden Markov Models are then trained based on the frequencies of these feature trajectories, from which recognition is performed. The approach is robust to different type of clothes the subjects wear. The method is applied to both the USF Gait Challenge data-set and CMU MoBo data-set, and achieved recognition rate of 61% and 96%, respectively. The results suggest that the recognition rate is significantly limited by the distance of the subject to the camera.

The model-based gait feature extraction step is insensitive to noise, cluttered background or even moving background. Furthermore, this approach also minimizes the size of the data required for recognition compared to model-free algorithms. It is evident that model-based approaches are view-invariant and scale-independent. However, model-based approaches perform poorly compared to model free approaches.

b. Model-Free Approaches. Model-free approaches focus on using the properties of the silhouette extracted from the video. Such properties include the contour, the raw pixels, the width. Model-free approaches are insensitive to the quality of silhouettes and have the advantage of low computational costs compared to model-based approaches. However, they are usually not robust to viewpoints and scale.

Wang *et al*. [2002] built a PCA-based gait recognition system. For each image sequence, an improved background subtraction procedure is first used to accurately extract spatial silhouettes of a walker from the background; Then, eigenspace transformation to time-varying silhouette shapes is performed to realize feature extraction.The nearest neighbor classifier using spatio-temporal correlation or the normalized Euclidean distance measure is finally utilized in the lower-dimensional eigenspace for recognition, and some additional personalized physical properties are selected for the validation of final decision. Experimental results on a small database show that the proposed algorithm has an encouraging recognition rate with relatively lower computational cost.

In Zhang *et al.* [2011], a new approach to gait recognititon based on dynamic template matching is proposed. This approach is suitable for video surveillance systems requiring high accuracy and low computational complexity. Human appearance is first extracted from the silhouettes in video sequences. The contour is unfolded in clockwise manner and every silhouette in image sequences is converted to a normalized distance of one dimension. A distance based on Dynamic time warping is then applied to measure the similarity between the probe sequences and gallery ones with predetermined parameters. In the experiment performed, a recognititon rate of about 92% was achieved. This approach is not only suitable for video surveillance systems but is simple and requires low memory storage. However, Hidden Markov Models have been found a better classifier than the Dynamic Time Warping, so there is the need to investigate the HMM for the gait classification.

Fazenda *et al.* [2005] present people recognition using gait. In this project a gait recognition algorithm based on the averaged silhouette of a person over a gait cycle is developed. A binary silhouette of the walking person is obtained by background subtraction; the binary silhouettes are then aligned and averaged over each gait period. The Euclidean distance between the averaged silhouettes of a number of persons is used for recognition purposes. Experimental results, using both lateral and oblique views, show very promising recognition rates.

This paper will be based on the use of eigengait for gait recognition in real time.

III. OBJECT DETECTION USING KERNEL DENSITY ESTIMATION

Kernel density estimation (KDE) is the most used and studied nonparametric density estimation method. The model is the reference dataset, containing the reference points indexed natural numbered. In addition, assume a local kernel function centered upon each reference point, and its scale parameter (the bandwidth). The common choices for kernels include the Gaussian: and the Epanechnikov kernel (Elgammal etal., 1991)

The Gaussian Kernel is given by:

$$K_N = (2\pi)^{-\frac{d}{2}} \exp\left(-\tfrac{1}{2}\|x\|^2\right) \qquad (1)$$

The Epanechnikov kernel is given by:

$$K_E = \begin{cases} -\frac{1}{2}c_d^{-1}(d+2)(1-||x||^2 & if\ ||x|| < 1 \\ 0 & otherwise \end{cases} \qquad (2)$$

Let $x_1, x_2, \ldots \ldots x_n$, be a random sample taken from a continuous, univariate density *f*. The kernel density estimator is given by,

$$\hat{f}(x;h) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{x-x_i}{h}\right) \quad (3)$$

K is the function satisfying $\int k(x)dx = 1$ which is refered to as the Kernel, h is a positive number, usually called the bandwidth or window width.

The practical implementation of this algorithm is discussed in these subsections of section 3.

A. Histogram computation.

The first 100 initial frames in the video sequence (called learning frames) are used to build stable distributions of the pixel RGB mean. The RGB intensities of each pixel position is accumulated for the 100 frames and we calculate the cumulative sum of the average intensities i.e (sum of (RGB)/3) is computed over 100 frames. A histogram of 256 bins is constructed using these pixel average intensities over the training frames. . The sum is then normalized to 1. That is we divide each histogram bin value with the accumulated sum to get a normalized histogram as shown in figure 1.

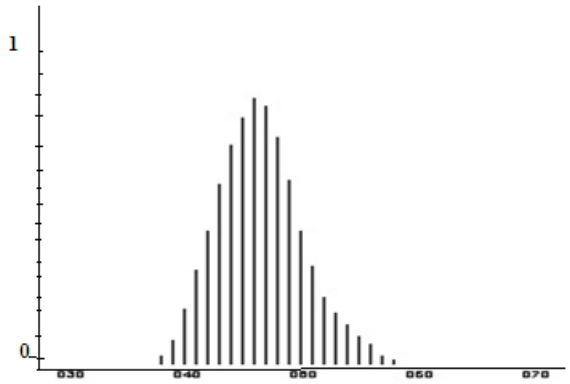


Figure 1. An Histogram of typical pixel location.

B. Threshold Calculation.

Threshold is a measure of the minimum portion of the data that should be accounted for by the background.

For more accuracy in our segmentation, we use different threshold for each histogram bins.

The pseudo- code for the Threshold calculation is given below

```
1  For each H[i]
2          Get sum of H[i]
3          Peak[i]=max(H[i])
4          Pth[i]=Peak[i]/2
5          Calculate sum2(H[i] > Pth[i])
6                  If(sum2(H[i] > Pth[i])  is less than 0.95 of sum of Hi
7                          Pthi=Peak[i]/2
8                          go to 5
9                  else
10                 threshold=Pth[i]
```

C. Foreground/ Background detection.

For every pixel observation, classification involves determining if it belongs to the background or the foreground. The first few initial frames in the video sequence (called learning frames) are used to build histogram of distributions of the pixel means. No classification is done for these learning frames. Classification is done for subsequent frames using the process given below. Typically, in a video sequence involving moving objects, at a particular spatial pixel position a majority of the pixel observations would correspond to the background. Therefore, background clusters would typically account for much more observations than the foreground clusters. This means that the probability of any background pixel would be higher than that of a foreground pixel. The pixel are ordered based on their corresponding value of the histogram bin.

The Background detection Algorithm

(Frames 1—N is used for training modeling the background).
Read frames 1 -- N
For each pixel
    Calculate the value of (r+g+b)/3
    Locate the corresponding bin value in the histogram of the pixel.
    Increment the bin of this value by 1
Increment the surrounding bandwidth pixels by fraction of 1
Normalize the histogram value by dividing each bin value by the sum of bins.
Calculate the adaptive threshold as given in figure xx.
Read the Next frame after N.
For each pixel
    Read the intensity of RGB of the pixel.
    Calculate the value of (r+g+b)/3
    Locate the corresponding bin value in the histogram of the pixel.
    Test if the value is < threshold
        Classify the pixel as foreground
    Else
        Classify the pixel as background

IV. A SPATIO-COLOR HISTOGRAM ALGORITHM FOR SCALABLE HUMAN OBJECT TRACKING

The proposed algorithm is composed of two stages. First is the appearance correspondence mechanism. Once detected, Appearance models are generated for objects appearing in the scene. The model is the estimate of probability distribution of colour of pixel colours. Multiple models are developed for a single object. These models are then used in subsequent frames to match the set of currently detected models and that of target models. In the second phase occlusion and object merge and separation are handled. The foreground object detected in previous stage is passed to the object tracker. This

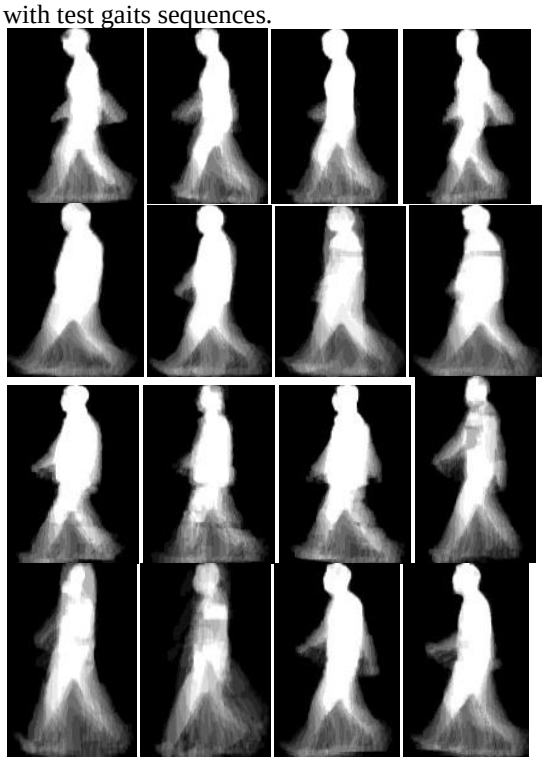
information is the appearance model of the object.
We adopt a multi-part tracking algorithm in our system. That is, we segment each silhouette into upper-body area and lower-body area and generate a histogram of colures in HSV color space for each region. This approach is good enough at discriminating individuals because of varying intensity in identical objects with similar color and occlusion. Our approach makes use of the object color histograms of previous frame to establish a matching between objects in consecutive frame. Our method is also able to detect object occlusion, object separation and label the object appropriately during and after occlusion [Akintola and Tavakollie, 2011]

V. MODELING GAIT USING GAIT ENERGY IMAGE (GEI)

Given series of gait within a cycle, , GEI is used as gait representation for the gait information in spatial and temporal domain. The silhouettes extracted from background modeling are used to construct gait energy image. Suppose each $I_{n,t(x,y)}$ is a particular pixel located at position (x, y) of t(t = 1, 2,…, T) image from n (n = 1, 2,…, N) gait cycle. All the silhouettes are normalized along both horizontal and vertical directions to a fixed size. Assuming that width and height of GEI are W and H representatively, GEI is defined as

$$g(x,y) = \frac{1}{T}\sum_{n=1}^{N}\sum_{t=1}^{T} I_{n,t}(x,y) \quad (4)$$

where T is the number of frames in gait sequence. I is a silhouette image at frame t, x and y are the image coordinates. The following methods were then used to match the templates with test gaits sequences.



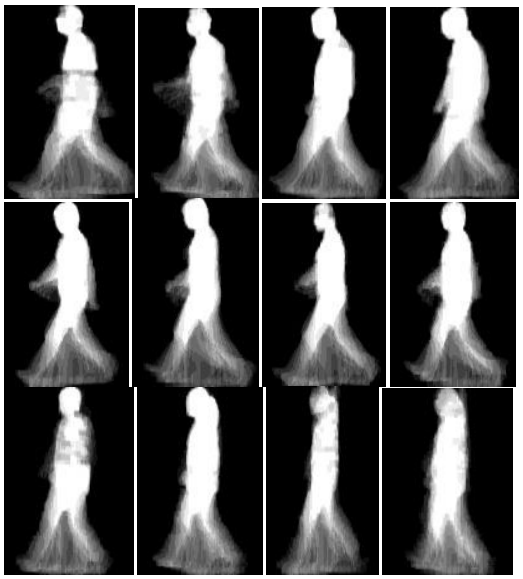


Figure 2 Sample gait energy images.

VI. GAIT RECOGNITION USING PRINCIPAL COMPONENT ANALYSIS (PCA)

The Eigengait space is obtained by applying the eigengait method to the training images. Later, the training images are projected into the Eigengait space. Next, the test image is projected into this new space and the train image projection with the minimum distance from the test image projection is the correct match for that test gait.

A. Training Operation in PCA

Let I be an image of size(Z*x* , *Zy)* pixels, then the training operation of PCA algorithm can be expressed in mathematical terms as follow:

a. Convert the training image matrix *I* of size (Z*x, Zy*) pixels to the image vector Γ of size (P×1) where

$$P = Zx \times Zy \quad (5)$$

(that is, the train image vector Γ is constructed by stacking each column of train image matrix I)

b. Create a training set of training image vectors such that its size is $P \times M$ where M is the number of training images.

$\Gamma_{PxM} = \{\Gamma_1\ \Gamma_2 \ldots \Gamma_M\}$ where, Γ_i represents the image vector of i^{th} training images

c. Compute arithmetic average (mean face,Ψ) of the training image vectors at each pixel point given by:

$$\Psi_{p\times 1} = \frac{1}{M}\sum_{i+1}^{M}\Gamma_i \quad (6)$$

d. Obtain mean subtracted vector (Φ) by subtracting the mean face (Ψ) from each of the training image vector as given below:

$$\Phi_{i} = \Gamma i - \Psi \quad (7)$$

e. Create the difference matrix (*A*) which is the matrix of all the mean subtracted vectors and is given by:

$$A_{PxM} = \{\Phi_1\Phi_2 \ldots \Phi_M\} \quad (8)$$

f. Compute the covariance matrix (X) given by:

$$Xp \times p = A.A^T \quad (9)$$

Compute the eigenvector and eigenvalue of the covariance matrix (X). The dimension of covariance matrix (X) is $P \times P = (Zx \,.\, Zy) \times (Zx \,.\, Zy)$. For an image of typical size, the cost of computing eigenvector of a matrix of such huge dimension is computationally intensive. An efficient way of solving this problem is analyzed below.

Let a matrix be defined as:

$Y_{M \times M}$ such that

$$Y = A^T . A \quad (10)$$

$$Y.\upsilon_i = \mu_i.\upsilon_i \quad (11)$$

$$(A^T.A).\upsilon_i = \mu_i.\upsilon_i \text{ (substituting } Y = A^T.A \text{ into 3.61)} \quad (12)$$

(Premultiplying both sides by matrix A)

$$(A.A^T.A).\upsilon_i = A.\mu_i.\upsilon_i \quad (13)$$

(as μ_i is a scalar value, the term can be rearranged.)

$$A.A^T.A.\upsilon_i = \mu_i.(A.\upsilon_i) \quad (14)$$

$$X.(A.\upsilon_i) = \mu_i.(A.\upsilon_i) \quad (15)$$

where $X = A.A^T$ (16)

$$X.\vartheta i = \mu_i.\vartheta i \quad (17)$$

where $\vartheta i = (A\,\upsilon_i)$

which is of the form $B.I = \lambda I$ where I and λ are the eigenvector and eigenvalue of B. Hence, it is concluded that ϑ and μ_i is one of the eigenvectors and eigenvalues of matrix $X = A.A^T$ respectively. Thus the eigenvector ϑi $(A.\upsilon i)$ of covariance matrix $X = A\,A^T$ is obtained first by computing the eigenvector (v_i) of matrix $Y = A^T.A$.

Therefore, instead of using P eigenfaces, M << P eigenfaces can be used for the eigenface projection. This technique is computational efficient. Eigenvectors are vectors pointing in the direction of the maximum variance and the eigenvalue of the eigenvector is directly proportional to the variance the eigenvector represents, that is, the larger the eigenvalue the larger variance the eigenvector represents. Hence, the eigenvectors are sorted with respect to their corresponding eigenvalues. The eigenvector having the largest eigenvalue is marked as the first in the eigenvector matrix. This is then followed by the second eigenvector having the second largest eigenvalue and so on until the desired number of eigenvectors is selected.

The training images are projected into the eigenfaces space to obtain a weight vector to be used to represent the image in lower dimensional space. Each weight is obtained by the dot product of each image with each of the eigenvector as:

g. Determine the projection of a training image of the eigenvectors as given below:

$$\omega_k = \upsilon^T_\kappa \cdot \Phi = \upsilon^T_k \cdot (\Gamma - \Psi)$$

$$k=1,2,\cdots,M \quad (18)$$

h. Determine the weight matrix, which is the representation of the training images in eigenfaces space:

$$\Omega_{Mx1} = [\omega_1 \omega_2 \cdots \omega_M]^T \quad (19)$$

The training operation of PCA algorithm ends with the computation of the weight matrix. At this point, the images are just composed of weights in the eigenfaces space, simply like they have pixel values in the image space. The important aspect of the eigenfaces transform lies in this property. Each image is represented by an image of size in (Z_x , Z_y) the image space, while the same image is represented by a vector of size (M × 1) in the eigenfaces.

B. Recognition operation in PCA

When a new probe(test) image is to be classified, it is also mean subtracted and projected onto the eigenfaces space and the test image is assumed to belong to the nearest class by calculating the Euclidean distance of the test image projections to that of the training image projections.

Let T be an image of size (Z_x, Z_y) pixels, then the recognition operation of PCA algorithm can be expressed in mathematical terms as:

a. Convert the test image matrix T of size (Z_x, Z_y) pixels (the size of test image must be same as that of the training images) to the image vector of size (P×1) where $P = Z_x \times Z_y$ (i.e: the test image vector is constructed by stacking each column of the test image matrix T)

b. Obtain mean subtracted test image vector (Φ_T) by subtracting the mean face (computed during the training session) from the test image vector as given below:

$$\Phi_T\,(P \times 1) = \Gamma_T - \Psi \quad (20)$$

Determine the projection of a test image on each of the eigenvectors as given below:

$$\omega_k = \upsilon^T_\kappa \cdot \Phi_T = \upsilon^T_k \cdot (\Gamma - \Psi)$$

$$k=1,2,\cdots,M \quad (21)$$

c. Determine the weight matrix Ω -which is the representation of the test image in eigenfaces space by:

$$\Omega_{Mx1} = [\omega_1 \omega_2 \cdots \omega_M]^T \quad (22)$$

d. Compute the value of similarity function of given test image for each training image. The similarity of test image with the ith training image is defined as:

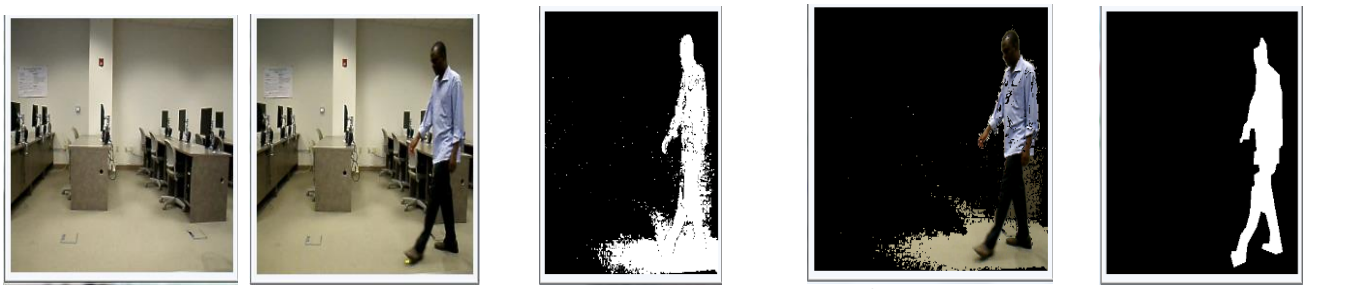
$$\delta_i = \|\Omega_T - \Omega_{\Psi_i}\| = \sqrt{\sum_{k=1}^{M'} (\Omega_T - \Omega_{\Psi_i})^2} \quad (23)$$

where (δ_i) is the Euclidean distance (L2 norm) between projections of images in face space. The training face image which has the minimum distance (L2 norm) is the face that images matches closely with the test face.

VII. EXPERIMENT

A video camera was stationed inside a computer laboratory FUTA to capture the gaits of students entering the laboratory. Four images each of fourteen individuals were then used to train an eigengaits recognition algorithm.

Figure 4.0 Stages in background subtraction



(a) background **(b)Moving objects moved** **(c) Noisy Mask with shadow** **(d) Colored foreground with shadow** **(e)Foreground Mask with shadow removed and morphological operation performed**

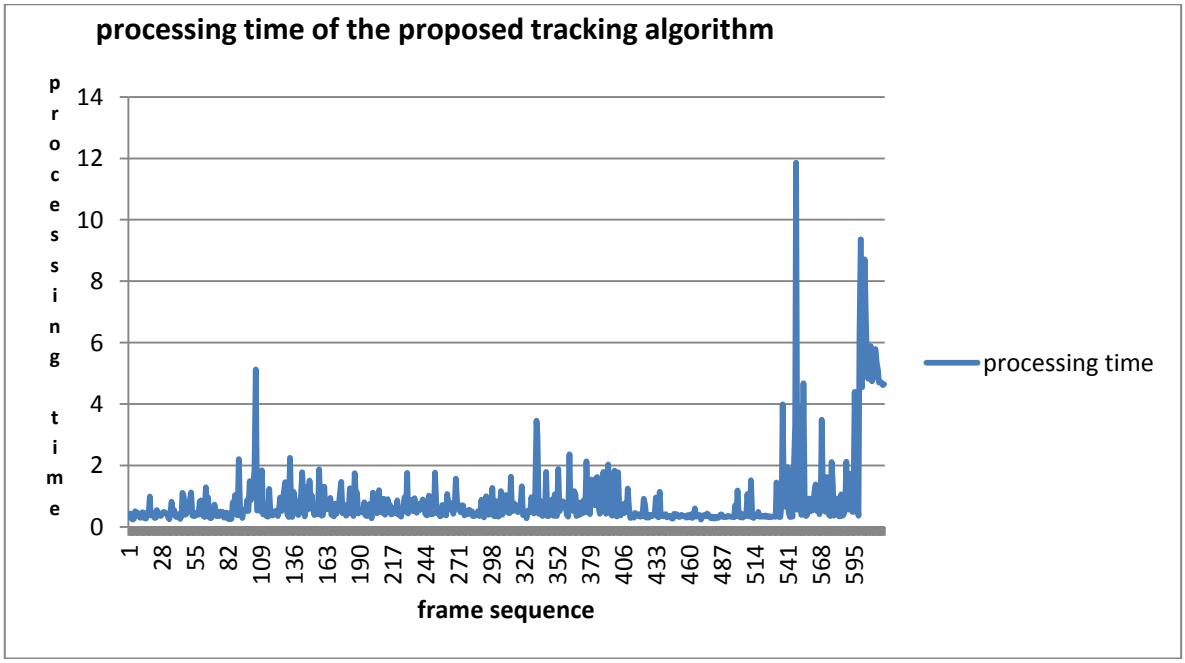


Figure 4.0 Tracking time per frame

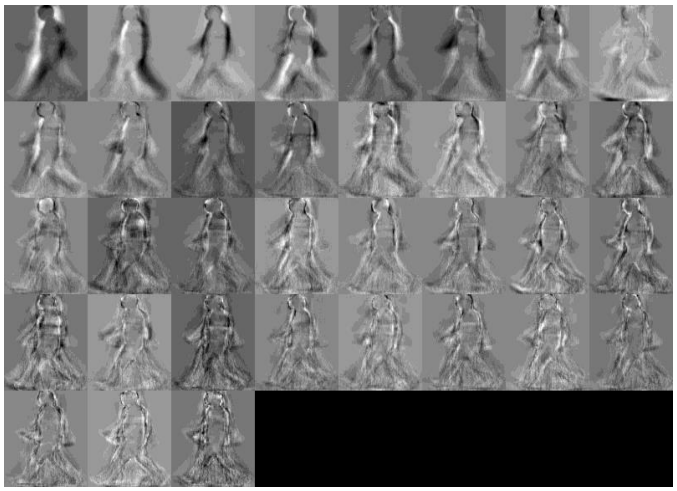


Figure 5 Some of the Eigengaits of the training gait images

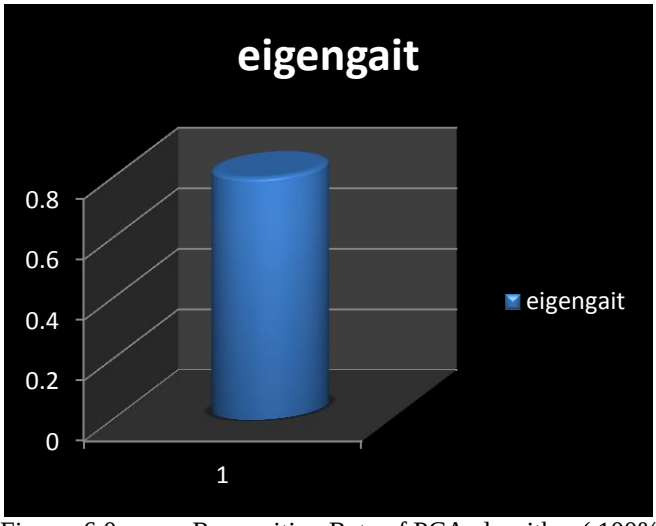


Figure 6.0 Recognition Rate of PCA algorithm (100%)

VIII. CONCLUSION

We have implemented gait detection and recognition algorithm using motion information and Gait Energy Image (GEI) with Principal Component Analysis (PCA). This will be highly appreciated in a surveillance environment where time is a critical requirement. It can be seeing that Principal Component Analysis works fine in recognizing such gaits as we can achieve good accuracy of 80'% with the recognition. Future work can be done on this study by using different other classifiers such as neural networks and support vector machines and K Nearest Neighbour. The system has provided an hybridized probabilistic models to enhance fast human recognition using human gait as features.

REFERENCES

[1] Akintola K.G, Tavakollie A. (2011) Robust Foreground detection in Videos using Adaptive Colour Histogram Thresholding and shadow removal 7th International Symposium on Visual Computing Las-Vegas, NV, September, 2011.

[2] Anik K. Jain, Arun Ross, Salil Prabhakar (2004) "An Introduction to Biometric Recognition" IEEE transactions on circuits & systems for Video Technologies (2004).

[3] Bobick A. F., and Johnson A.Y., (2001) Gait Recognition Using Static, Activity-SpecificParameters, CVPR, 8-14 December 2001, Kauai, HI,USA, pp. 423-430

[4] Bobick A.F. and Davis J.W., The recognition of human movement using temporal templates. IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI), Vol. 23 issue 3, pp257- 267, March 2001

[5] Bouchrika I. and Nixon M.S. (2008) Gait Recognition by Dynamic Cues 19TH IEEE International Conference on Pattern Recognition, (ICPR, 2008) 8-101 December, 2008, Tampa Florida USA .

[6] Caetano T.S and Barone D.A.C A probabilistic Model for the Human Skin Color, 11th International conference on image analysis and processing 26-28 September, 2001 (ICIAP01), Palermo Italy, pp.279-283

[7] Collins, R. T., Lipton, A. J., Fujiyoshi, H., & Kanade, T.(2000). A system for video surveillance and monitoring: VSAM final report. Technical report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, May 2000.

[8] Comaniciu D., Ramesh V., and Meer P..(2000) Real-time tracking of non-rigid objects using mean shift. In Proc. Conf. Comp. Vision Pattern Rec., pages II:142–149, Hilton Head, SC, June 13-15, 2000.

[9] Elgammal A., Duraiswami R, Harwood D., and Davis L.S (2002) Background and foreground modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance, PROCEEDINGS OF THE IEEE, VOL. 90, NO. 7, JULY 2002.

[10] Fazenda, J., Santos D., and Correia P., (2005) Using gait to recognize people Eurocon 2005, Serbia and Montenegro Belgade November 22-24, 2005.

[11] Fazli S., Moradi H., Bouzari P.H.,(2009) Particle Filter based Object Tracking with Sift and Color Feature proceedings of IEEE Second International Conference on Machine Vision 28-30 Dec., Dubai, United Arab Emirate

[12] Kale A., Cuntoor N., Yegnanarayana B., Rajagopalan A.N, and Chellappa R. (2003) Gait analysis for human identification In Proceedings of AVBPA'2003. pp.706~714. Guldford, UK

[13] Kale A., Rajagopalan A.N., Cuntoor N. and Kruger V. (2002) Gait-Based Recognition of Humans using Continuous HMMs Proceedings of fifth International conference on Automatic Face and Gesture Recognition (FGR'02) May 20-21 Washington DC, USA.

[14] Turk M. Pentland A. (1991). "Eigenface for Recognition", Journal of Cognitive NeuroScience Vol. 3 pp 71-86

[15] Modi R.V and Mehta T.B. (2011) Neural network based approach for recognition of human motion using stationary camera, International journal of computer applications (975-887) vol 25 –no 6 -July 2011

[16] Spencer N.M and Carter J.N. (2002) Viewpoint invariance in automatic Gait Recognition. Proc. AutoID 14-15 March, 2002 Wetchester Marriot, 670 white plain road Tarry Town, New York, USA

[17] Katiyar R., Pathak V.K (2010) Clinical gait data analysis based on Spatio-Temporal Features International Journal of Computer Science and Information Security, IJCSIS, Vol. 7 No. 2, February 2010, USA. ISSN 1947 5500

[18] Nowlan M. F., (2009) Human Identification via Gait Recognition using accelerometer Gyro forces, CSPC-536Networke embedded systems and sensor networks a project.Computer Science department Yale University

[19] Viola, P. and Jones M., (2001). Rapid object detection using boosted cascade of simple features. In: Proceedings IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, HI Hawaii USA.

[20] Wang L., Hu W., Tan T., (2002) A new attempt towards gait-based human identification ICPR2002 The 13th IEEE International Conference on Patteren Recognition Qubec.

[21] Zhang R Vigjer c M (2007) Human gait recognition at sagital plane, Image and vision computing vol., 25 issue 3, March 2007 pp 321-330.

A Text-Independent Speaker Recognition Framework for Detecting Criminals

Amusa, A. I.
Department of Computer Science,
Federal University of Agriculture,
Abeokuta, Nigeria.
amusa.afolarin@yahoo.com

Sodiya A.S.
Department of Computer Science,
Federal University of Agriculture,
Abeokuta, Nigeria.
sinaronke@yahoo.uk

Onashoga S.A.
Department of Computer Science,
Federal University of Agriculture,
Abeokuta, Nigeria.
bookyy2k@yahoo.com

Aborisade D.O.
Department of Computer Science,
Federal University of Agriculture,
Abeokuta, Nigeria.
aborisadeda@funaab.edu.ng

Abstract- In order to enhance security and safety in most countries, there is an urgent need to develop efficient techniques for combating kidnapping and other crimes that pose threat to human lives. This paper presents a text-independent speaker recognition system for detecting criminals. Mel frequency - Perceptual Linear Predictive (MF-PLP) and Perceptual Linear Predictive (PLP) are used for the features extraction. Spectral subtraction approach and Cepstral Mean Normalization method were used for noise reduction. Vector quantization and Autoassociative Neural Network (AANN) were used as classifiers for efficient detection. The system is designed to recognize speech pattern by matching input signal with stored samples. The proposed framework shows a system that can be used to improve the performance of the speaker identification even in noisy condition and generally assists in reducing crime rate in nations.

Keywords: speaker's recognition, speech recognition, text independent, voice identification.

I. INTRODUCTION

In recent years, there has been a high increase in crimes specifically kidnapping and these pose threat to the safety and security of human lives. Although speaker recognition has improved significantly in recent times, there is not a suitable system that can solve the problem of kidnapping which is a major challenge in the society. Humans have used body characteristics such as face, voice gait, etc. for thousands of years to recognizing people for detecting criminals [1]. Alphonse Bertillon, chief of the criminal identification division of police department in Paris, developed and then practiced the idea of using a number of body measurements to identify criminals in mid-19 century. [1]. A wide variety of systems require reliable personal recognition schemes either confirm or determine the identity of an individual requesting their services, the purpose of such schemes is to ensure that the rendered services are accessed only by legitimate user and thus unauthorized user.

A biometric system is essentially a pattern recognition system that operates by acquiring biometric data from an individual, extracting a feature set from the required data and comparing this features set against the template in the database. [1]. Thus, biometric systems fall under the ambit of technology designed and used specifically for measuring and analyzing the characteristics of a person. [1]. Any physiological and behavioral characteristic of a person can be used as a biometric feature as long as the following criteria are taken into account. [16].

• Universality: Each person will have a characteristic which is distinguished from another person in question.

• Distinctiveness: Any two people should be sufficiently different in terms of their characteristics.

• Permanence: The characteristic should be sufficiently invariant over a reasonable period of time;

• Collectability: Measuring the characteristic quantitatively should be possible;

• Performance: This refer to the achievable recognition accuracy and speed, the resources required to achieve the desired performance, as well as the operational and environmental factors that affect performance;

• Acceptability: It indicates the extent to which people are willing to accept the use of a particular biometric identifier, i.e., the characteristic, in their daily lives;

• Circumvention: This reflects how easily the system can be bypassed using fraudulent methods. Even though reliable methods of biometric personal identification like finger-print analysis and retinal or iris scan do exist, however, the validity of forensic fingerprint evidence has recently been challenged by academics, judges and the media. [1]. While fingerprint identification was an improvement over earlier systems, the subjective nature of matching, along with

relatively high error rate has made this forensic practice controversial. [1]. As far as Iris recognition is concerned, it is very difficult to perform at a distance larger than a few meters and depends on the cooperation of the person. The initial investment in setting these systems is relatively high. [14]. In contrast, voice recognition systems have the following advantages over biometric identification systems: Users can enroll themselves over the telephones rather than having them enrolled in person to deliver a fingerprint or iris scan.

- The technology also requires no special data-acquisition system, other than a microphone. In case of signature verification systems, a specialized digital pen tablet acts as the data acquisition system whereas in Iris recognition systems, an Iris reader is deployed. Thus, hardware costs are reduced to a minimum.
- The voiceprint generated upon enrolment is characterized by vocal tract (pitch, volume and tone) which is a unique physiological trait. A cold does not affect the vocal tract, so there will be no adverse effect on accuracy levels. Only extreme conditions such as laryngitis can hinder the optimal performance of the system.
- Voice recognition offers relatively low perceived invasiveness as compared to iris recognition, face recognition and signature verification.

Speaker recognition is the process of validating a user's claimed identity using characteristics extracted from their voices. [1]. Speaker recognition is a process that enables machines to understand and interpret the human speech by making use of certain algorithms and verifies the authenticity of a speaker with the help of a database [8]. No two individuals sound identical because their vocal tract shapes, larynx sizes and other parts of their voice production organs are different. [1]. In addition to these physical differences, each speaker has a distinctive manner of speaking, like the use of a particular accent, rhythm, intonation style, pronunciation pattern, choice of vocabulary etc. . Speaker recognition can be divided into two specific tasks: identification and verification [9] If a speaker claims a certain identity his or her speech is used to identify if he or she is who he or she claims to be, this is called *verification* or *authentication* [12] and identification is trying to identify a speaker or comparing a speaker voice pattern with other voice patterns in the database to get the speaker. The two nodes modes of operation for speaker identification namely; closed-set and open-set. Closed set means set of known voice and Opened set means set of unknown voice which is referred to as *impostors*. [12]. The closed set of speaker identification can be considered as a multiple-class classification problem. It can also be divided into three categories based on the types of utterances 1.Text dependent 2.Text independent 3.Text prompted. In Text independent the user can utter an arbitrary phrase (unconstraint), whereas Text dependent a fixed voice password is uttered and Text prompted users are asked to repeat phrase [12]

The various algorithms used in speaker recognition which are called *classifiers* are to process and store voice print includes: frequency estimation, hidden Markov model, Gaussian mixture model, pattern matching algorithms, neural networks, matrix representation, vector quantization and decision tree. There are many features depicting the characteristics of the vocal tract which are called *features extraction* such as; MFCC, PLP, MF-PLP, Spectral subtraction approach, Cepstral Mean Normalization method and they are used in both identification and verification. Basically this framework is focusing on using MF-PLP, PLP and Spectral subtraction approach together with Cepstral Mean Normalization for features extraction and vector quantization for classifier in getting a suitable system for detecting criminals.

II. PROBLEM STATEMENT

The main purpose of this framework is to produce a suitable an efficient system for solving crimes incidents in the society specifically kidnapping. With this system it will reduce the rate of kidnapping in our society and also help the government, judiciaries and the force men in getting criminals which other biometrics and related works have been unable to solve in recent years.

III. RELATED WORK

This paper work makes use of multilevel cryptography to preserve data integrity while in transit or storage. Encryption and decryption follow a transform based approach layered with pseudorandom noise addition whereas for pitch detection, a modified version of the autocorrelation pitch extraction algorithm is used. [1]

The experiment results shows that the proposed algorithm can decrypt the signal test with exponentially reducing Mean Square Error over am increasing range of SNR. [1]

Furthermore, it outperforms the conventional algorithms in actual identification tasks even in noisy environments. [1].

The primary objective of this work was to implement a robust and secured voice recognition system using minimum resources offering optimum performance in noisy environments. [1]. The system was implemented using three levels of encryption for data security and autocorrelation based approached to find the pitch of the sample.

The result of this work pointed to improve the performance in adverse SNR environments and at the same time high accuracy rate of the pitch detect algorithm makes it very powerful in both clean and noisy environment. This paper was proposed based on features extracted from speech recorded using close speaking microphone in clean and noisy environment.

This system recognizes the speaker from a number of acoustic features that include linear predictive coefficient, linear predictive cepstral coefficient (LPCC) and Mel-frequency cepstral coefficients (MFCC). [14]

RBFNN and AANN are two modeling techniques used to capture the features, RBFNN model enable nonlinear transformation followed by linear transformation to achieve a higher dimension in the hidden space and the AANN is used to capture the distribution of the acoustic feature vectors in the feature space. [14]. The experiment result shows that, the performance of AANN performs better than RBFNN and also gives better accuracy of 94.93% for various features both clean and noisy environment. After all the experiment in this paper work, we are made to realized AANN has a better accuracy than RBFNN for features both clean and noisy environment and also made us realize that in the future throat microphone can be used for speaker recognition. This paper combined MF-PLP and PLP, considering the limitation of PLP, which present the effectiveness of perceptual features and iterative clustering approach for performing both speech and speaker recognition and its evaluation on clean text speeches. [5]. Showing the comparison between MF-PLP and PLP for speaker independent isolated digit recognition in term of their individual accuracy, Average accuracy of the MF-

PLP feature is better than PLP feature for speaker independent case. [5]. Where MF-PLP gives an accuracy of 91% for identifying speakers and it also gives low value rate of (Fast Acceptance Rate) % FAR, (Fast Rejection Rate) % FRR and(Equal Error Rate) %EER, which makes MF-PLP better feature for speaker recognition. [4]. The general use of both features is to get a better result considering the limitation of PLP and another important point is that this speaker recognition system is evaluated on the identical messages for all enrolled speakers and its reveals that perceptual features indicates the characteristics of speaker rather than the spoken content due to the formation of the training speech for the purpose of text independent recognition. [4]. These perceptual features depict the characteristics of the speech due to the formation of training in the case of speaker independent speech recognition. [4]. [12] gave solution to the performance of automatic speaker recognition degrades in the presence of additive noise and convolutive noise causing severe difficulties in various environments. This paper addresses the problem of additive noise and convolutive noise in speech but also has its own limitations which are the performance of the noise suppression system is based on accuracy on the background noise estimate, so better methods are needed to silence or speech detection and further research can be conducted to address various issues involved in implementing the system in real-time and issues which are still inherent to subtractive type of algorithms.

The goal of this research work is to create a speaker recognition system and apply to speech of an unknown speaker. [12]. Based on the research work much care and time was taken in getting an efficient speaker recognition system; however this task has been challenged by high variant input speech signals, due to the fact that people voice changes with time, health conditions, speaking rate, varied microphones and channels that people use can cause difficulties because of all these difficulties there is a need for a well suitable system for speaker recognition. [12]. The objective of modeling techniques is to generate speaker models using speaker specific feature vectors. Such models will have enhanced speaker-specific information at reduced data rate. This is achieved by exploiting the working principles of the modeling techniques. [12] The proposed approach could be used with conjunction with other ones like face recognition, for better security and increasing the area of the particular application. [12].

IV. METHODOLOGY

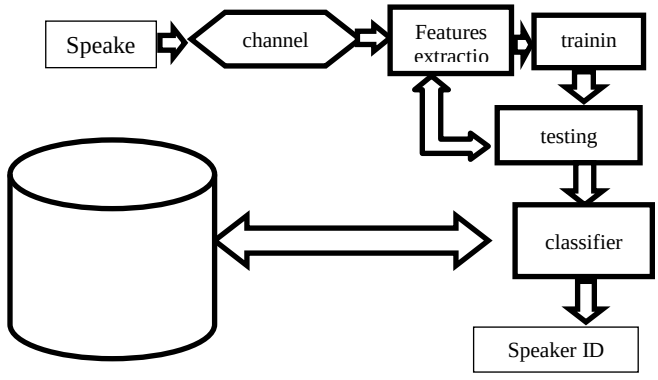


Fig 1: Proposed Architecture for Text-Independent for speaker recognition system

A. **Speaker**: A speaker is a person who gives a speech or a lecture. In the proposed speaker recognition system, the speaker is the known or unknown user who speaks through the channel for the means of extracting the voice.

B. **Channel**: A channel is an intermediary between the speaker a system, the speech can be recorded or maybe during a chat. The devices used for channel are; headphones, standard microphones etc.

C. **Data collection**: To implement this system, test data is needed. The database of all the voices were collected in different environments of different languages with little noise and unpleasant noise. Some of these voices were collected on recording speakers form the market, halls and even microphones etc. The speech data was collected in multiple sessions. A microphone was used to record all the data clips and we were able to record in clips of 20 speakers, each of the speaker data consist of 5 speech clips of varying lengths. The data is spitted into three categories which the vocal tracts will pass through namely; training, tuning and testing. A diagram is shown below of collected recorded clips for database. One major factor that affected the result was the distance between the speaker's mouth and the microphone which resulted to performing another experiment on all voices gotten through microphones.

Figure 2: Recorded clips for databases

D. **Features extraction:** The short term speech spectrum can be classified into two component which are namely; glottal pulse shape and harmonic peak due to periodicity of voiced speech. [6].The periodicity of the voice speech is decided by excitation source and also reflects the characteristics of the speaker. The Spectral enveloped is shaped by formants which reflect the resonances of the vocal tract. The variations among speakers are indicated by formant location and bandwidth. [5]. PLP and MF-PLP extraction. The detailed procedure PLP (perceptual linear predictive cepstrum) and MF-PLP (Mel frequency perceptual linear predictive cepstrum) extraction is showed in in Fig 2

Compute power spectrum of windowed speech

• Perform grouping to 25 critical bands in bark scale or Mel scale for sampling frequency of 16Hz.

• Perform loudness equalization and cube root compression to stimulate the power law of hearing.

• Perform Inverse fast Fourier transform (IFFF).

• Perform Linear Predictive (LP) by Levinson-Durbin procedure

• Convert LP coefficient into cepstral coefficients.

The relationship between frequency in Bark and frequency in Hz

$$f(bark)6 * \operatorname{arcsinh}\left(\frac{f(Hz)}{600}\right) - Eq.1$$ [6]

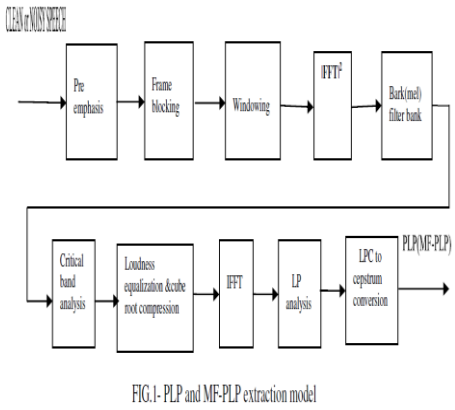


FIG.1- PLP and MF-PLP extraction model

Figure.3: PLP and MF-PLP extraction model [5]

Algorithms

Hybrid algorithm is used in to remove the additive and convolutive noise. It is used and modified in this work to remove additive noise and convolutive noise that is extracted from the channel. Cepstral coefficient are estimated by applying DCT (Discrete Cosine Transform) to the spectral coefficients

Compute Cepstral mean of all N cepstral vector using the formula.

$$1/N \sum_{n=1}^{N} x[n] \quad (2)$$

$$y[n] = x[n] - 1/N \sum_{n=1}^{N} x[n] \quad (3)$$

The hybrid algorithm designed has the following steps:

• Divide the signals into frames

• Apply the windowing function to each frame

• Apply spectral subtraction method on each frame. The enhanced signal is estimated after removing the additive noise.

• Apply CMS method on the enhanced signal obtained
• Cepstral coefficient are estimated after removing the convolutive noise
• These cepstral coefficients can be used as an input for speaker recognition system.

E. **Training phase:** In this phase, a model was created which is the background model. The background model is a large pool, where the wave files are converted into different formats so it will be easy to analyze. These wave files are continuous signal which are broken down into discrete parameter vectors. Here each vector is about 12cm long, because we are assumption that the duration of vector is stationary. It's often strictly used but its reasonable approximation to make use of. Here we use a proposed hybrid algorithm format. The hybrid algorithm designed has the following steps:

• Divide the signals into frames
•Apply the windowing function to each frame
•.Apply spectral subtraction method on each frame. The enhanced signal is estimated after removing the additive noise.
• Apply CMS method on the enhanced signal obtained
• Cepstral coefficient are estimated after removing the convolutive noise
• These cepstral coefficients can be used as an input for speaker recognition system.

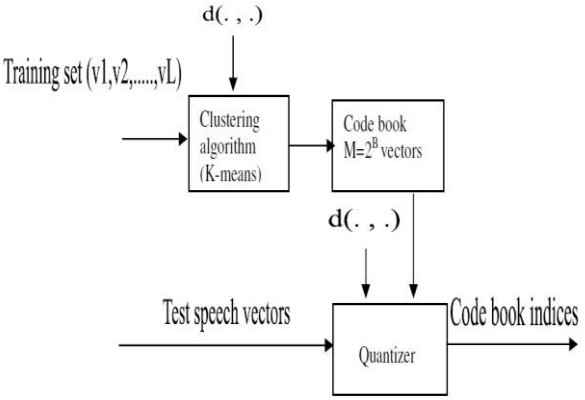


Figure 4: VQ training and classifier structure [14]

F. **Testing phase:** Once we have gotten the threshold values and the speaker model to use then we can test the remaining files. With this test, we will be sure if all the analysis done in the training phase is accurate enough. While calculating the threshold in the tuning phase the remaining speaker files are tested. To be sure the system is accurate while verifying users then we need to test the threshold values in two ways- false alarms and false rejections. If the likelihood is an imposter file is higher than the threshold of the speaker being tested then the system accept the imposter as the speaker. On the other hand sometime a speaker's file may not have a likelihood value higher than the threshold, there is a false alarm. Which means the speaker is falsely identified as an imposter, this is called false rejection. An optimal threshold value would minimize both values and keep the ERROR RATE low.

G. Classifier:

Vector Quantization: VQ approach is used in this system, due to ease of implementation and very high accuracy. VQ is a process of mapping vectors from a large vector space to a finite number of regions in that space [13]. Each region is called *cluster* which are represented by its Centre called a *code word*. The collection of all code-words is called a codebook. [12]. Below is a diagram that illustrates the conceptual of this recognition process. In the diagram below, we used only two speakers and two dimensions of the acoustic space as shown in the diagram. The circles refer to the acoustic vectors from the speaker 1 while the triangles are from the speaker 2. In the training phase, a speaker-specific VQ codebook is generated for all known speaker by clustering his/her training acoustic vectors. The result codewords (centroids) are shown in the diagram by black circles and black triangles for speaker 1 and 2, respectively. [13]. The distance from a vector to the closest codeword of a codebook is called a VQ-distortion [12]. In the recognition phase, input utterances of unknown voices are vector-quantized using each trained codebook and the total VQ distortion are computed. The speakers corresponding to the VQ codebook with smallest total distortion are identified.

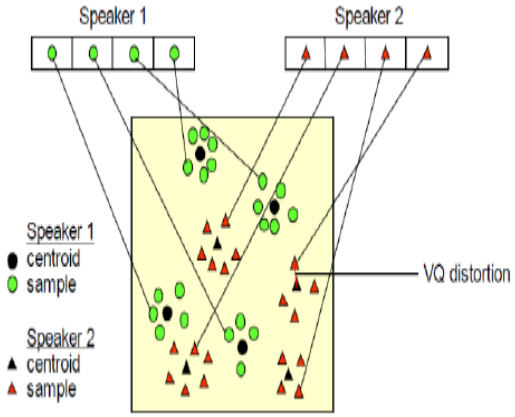


Figure.5 Conceptual diagram for vector quantization codebook formation.

Autoassociative neural network model:

They are feed forward neural networks performing an identity mapping of the input space and are used to capture the distribution of all data. [7]. Considering five layers of AANN as shown in the diagram below which have three different hidden layers. These layers are called the

dimension compression hidden layer, as these layers causes the inputs to go through a dimension compression process. [7]. In this proposed network the second and fourth layers have more units than the input layer. The third layer has fewer units than the first and fifth. [14]. The process unit in the first and the hidden layers are nonlinear, the unit in the second compression or hidden layer can be called linear or nonlinear. The activation function in the second, third and fourth layer are nonlinear. [14]. The structure of an AANN model used in our study is 39L 78N 4N 78N 39L for MF-PLP for capturing the distributed acoustic functions, where L denotes a linear unit. [14]. The integer value indicates the number of units used in that layer. The non-linear units use tanh(s) as the activation function where s is the activation values of the unit. [14]. A back propagation learning algorithm is used to adjust the weights of the network to minimize the mean square error for each features vectors [17]

During training the target vectors are same as the input vectors. The AANN trained with a data set will capture the subspace and the hyper-surface along the surface of maximum variance of the data. [15]..

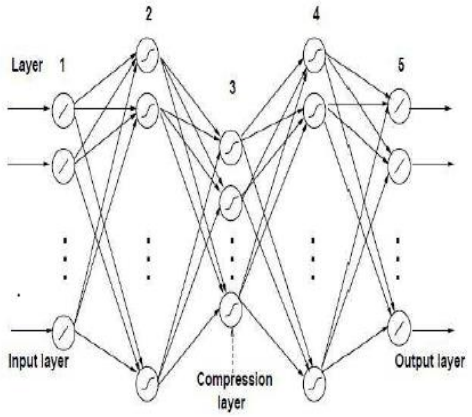


Figure 6.: Autoassociative neural network

V. CONCLUSION

In this paper we have proposed a text-independent speaker recognition system using MF-PLP and Hybrid algorithms features extracted from the voice signal or speech data that can be used to represent each speaker and Vector Quantization and AANN are the models used. It is expected that the characteristics of the speech data would be collected from clean and noisy environments using close speaking microphone and headphones etc. we intend to capture speaker information present using source features such as; AANN and Vector Quantization, AANN methods provides robustness against channel and also gives a better accuracy and Vector Quantization is also used to ease implementation and very high accuracy. We also intend to use hybrid algorithm and MF-PLP basically for additive and convolutive noise reduction to get a better result in detecting a criminal.

REFERENCES

[1] Aman Chadha, Divya M. Jyoti, Mani Roja "Text-independent Speaker Recognition for low SNR Environments with Encryption" Department of Electronics and Telecommunication, 2011

[2] .A.D Andrews, M.A Kohler and J.P Campbell, "Phonetic Speaker Recognition" Euro Speech, 2001

[3] A.Higgins., L. Bahler. and J.Porter, "Speaker Verification using Randomized 4.Phrase Prompting" Digital Signal Processing vol. 1 pp. 89—106, 1991

[4] A.K Jain, A.Ross., Prabhakar. S, "An introduction to biometrics recognition," IEEE Transaction on circuits and systems for Video Technology, Vol. 14, Jan. 2004.

[5] A.Revathi, R.Ganapathy, Y.Venkataramani. "Text Independent Speaker Recognition and Speaker Independent Speech Recognition using Iterative Clustering Approach" International Journal of Computer Science and Information Technology, Nov. 2 2009

[6] B.Yegnanarayana, K.S Reddy, S.P Kishore, Source and System Features for Speaker Recognition using a model " In IEEE Int. Conf. Acoust., Speech and Signal Processing, Salt Lake City , Utah, USA. Pp. 409-412, 2001.

[7] B. Yegnanayarana, S. Kishore, ANN: an alternative to GMM for pattern recognition Neural Networks, 15.419-469, 2002.

[8] D.A Reynolds, "Special Issues on Speaker Recognition, Digital Signal Processing" vol. 10, Jan 2000.

[9] L.R Rabiner ,"A Tutorials on hidden Markov models and Selected Applications IN Speech Recognition" Proceeding of the IEEE, 1989

[10] M..Carey .M, E.Parris and J.Bridle, "A Speaker Recognition System" using Alphanets" ICASSP, pp. 397—400, May 1991.

[11] N.Tadokoro., T.Kosaka , M. Kato and M. Kohda "Improvement of Speaker Vector-Based Speaker Verification" 2009 Fifth International.

[12] Parul, R.B Dubey "Automatic Speaker Recognition System" International journal of Advance Computer Research, Dec. 2012

[13] Rekha Nair, Nirmala Salam, Ashutosh Singh, Ganesh Joshi "An efficient method for additive and convolutive Noise Reduction" Internal journal of Electronics and Computer Science Engineering, 2011

[14] R.Visalashi., P Dhanalaksmi "Speaker Recognition in Clean and Noisy Environment using RBFNN and AANN" Journal of Theoretical and Applied Technology, July 31, 2014

[15] R. Clarke., "Human Identification in Information Systems: Management Challenges and Public Policy Issues" Information Technology & people, Vol.7, No. 4, 6-37, 1994

[16] S. Laxman. and P,S Sastry, "Text-dependent speaker using speaker specific Compensation" Dept. Electrical Engineering India Institute of Science Bangalore 560012 INDIA

[17] R .Kinnunen, H. Lib, "An Overview of Text-Independent Speaker Recognition: from Feature to Supervectors" 1-3,2008

[18] Y..Goto ,T. Akatsu, M. Katoh, T. Kosaka and Kohda "An investigation on Speaker Vector Based Speaker Identification under noisy Condition" Graduate School of Science and Engineering, Yamagata Uni. Tcn15987@st.yamagata-u.ac.jp, tatsuyal17akatsu@yahoo.co.jp,{katoh, tkosaka, kohda}@z.yamagata-u.ac.jp.

[19] Y. Zhou, J. Wang., X. Zhang. "Research on Adaptive Speaker Identification Based on GMM" PLA University of Science and Technology Nanjing China, 210007 zhouyuhuan250@sina.com.

Design and Implementation of Text to Speech System for Visually Impaired People

Afe Oluwaseyi F.
Department of Computer Science,
Lead City University, Ibadan.
afeseyi@yahoo.com

Oladunmoye M.
Department of Computer Science,
Lead City University, Ibadan.
bjarchimat@gmail.com

Oyekunle V.B.
Department of Computer Science,
Lead City University, Ibadan.
bolanleoyekunle@yahoo.com

Abstract- Computers are fast becoming our ways of life and one cannot imagine life without computers in today's world. Systems are built to perform a lot of operation such as Reservation of ticket in a railway station, Electronics library system, Electronics voting system e.t.c. Since computers are used in every possible field today, it becomes an important issue to understand and build these computerized systems in an effective way.
Technology defines the world as the vacuum of excellence where dynamism and stability reign. With this trend, network simplification is adopted across all platforms of the day in day out activities to actualize the convenience and modality in dealing with a complex network design and architecture. Wireless technology brings light to the world of the inter-connectivity as it reduces the stress of drawing wires across the globe. More also it reduces the troubleshooting of wired lines technology as in a case of partial contact or bridge of network.
Speech synthesis has long been a vital assistive technology tool and its application in this area is significant and widespread. It allows environmental barriers to be removed for people with a wide range of disabilities. The longest application has been in the use of screen readers for people with visual impairment, but text-to-speech systems are now commonly used by people with dyslexia and other reading difficulties as well as by pre-literate youngsters. They are also frequently employed to aid those with severe speech impairment usually through a dedicated voice output communication aid.
The aim of this research work is to identify the use of Text To Speech System technology and the objectives include the analysis of user environment of text to speech technology, application that allow a caller to access the system over the phone and to enable the caller to be presented with a selection of options.

I. INTRODUCTION

Information technology (IT) is the study, design, development, implementation, support or management of computer-based information systems, particularly software applications and computer hardware"(. IT deals with the use of electronic computers and computer software to convert, store, protect, process, transmit, and securely retrieve information.
A text-to-speech system (or "engine") is composed of two parts: a front-end and a back-end. The front-end has two major tasks. First, it converts raw text containing symbols like numbers and abbreviations into the equivalent of written-out words. This process is often called text normalization, pre-processing, or tokenization. The front-end then assigns phonetic transcriptions to each word, and divides and marks the text into prosodic units, like phrases, clauses, and sentences. The process of assigning phonetic transcriptions to words is called text-to-phoneme or grapheme-to-phoneme conversion. Phonetic transcriptions and prosody information together make up the symbolic linguistic representation that is output by the front-end. The back-end—often referred to as the synthesizer—then converts the symbolic linguistic representation into sound.

II. THE AIM AND OBJECTIVES

The aim of this research work is to identify the use of Text To Speech System technology and the objectives include the analysis of user environment of text to speech technology, application that allow a caller to access the system over the phone and to enable the caller to be presented with a selection of options.

III. LITERATURE REVIEW

A. What Is Speech Synthesis?

Speech synthesis is the artificial production of human speech. A computer system used for this purpose is called a speech synthesizer, and can be implemented in software or hardware. A text-to-speech (TTS) system converts normal language text into speech; other systems render symbolic linguistic representations like phonetic transcriptions into speech.
Synthesized speech can be created by concatenating pieces of recorded speech that are stored in a database. Systems differ in the size of the stored speech units; a system that stores phones or diphones provides the largest output range, but may lack clarity. For specific usage domains, the storage of entire words or sentences allows for high-quality output. Alternatively, a synthesizer can incorporate a model of the vocal tract and other human voice characteristics to create a completely "synthetic" voice output.
The quality of a speech synthesizer is judged by its similarity to the human voice and by its ability to be understood. An intelligible text-to-speech program allows

people with visual impairments or reading disabilities to listen to written works on a home computer. Many computer operating systems have included speech synthesizers since the early 1980s.

B. Applications

Speech synthesis has long been a vital assistive technology tool and its application in this area is significant and widespread. It allows environmental barriers to be removed for people with a wide range of disabilities. The longest application has been in the use of screenreaders for people with visual impairment, but text-to-speech systems are now commonly used by people with dyslexia and other reading difficulties as well as by pre-literate youngsters. They are also frequently employed to aid those with severe speech impairment usually through a dedicated voice output communication aid.

Sites such as Ananova and YAKiToMe! have used speech synthesis to convert written news to audio content, which can be used for mobile applications.

Speech synthesis techniques are used as well in the entertainment productions such as games, anime and similar. In 2007, Animo Limited announced the development of a software application package based on its speech synthesis software FineSpeech, explicitly geared towards customers in the entertainment industries, able to generate narration and lines of dialogue according to user specifications. The application reached maturity in 2008, when NEC Biglobe announced a web service that allows users to create phrases from the voices of Code Geass: Lelouch of the Rebellion R2 characters.

TTS applications such as YAKiToMe! and Speakonia are often used to add synthetic voices to YouTube videos for comedic effect, as in Barney Bunch videos. YAKiToMe! is also used to convert entire books for personal podcasting purposes, RSS feeds and web pages for news stories, and educational texts for enhanced learning.

Software such as Vocaloid can generate singing voices via lyrics and melody. This is also the aim of the Singing Computer project (which uses GNU LilyPond and Festival) to help blind people check their lyric input.

C. CHALLENGES

1) Text Normalization Challenges

The process of normalizing text is rarely straightforward. Texts are full of heteronyms, numbers, and abbreviations that all require expansion into a phonetic representation. There are many spellings in English which are pronounced differently based on context. For example, "My latest project is to learn how to better project my voice" contains two pronunciations of "project".

Most text-to-speech (TTS) systems do not generate semantic representations of their input texts, as processes for doing so are not reliable, well understood, or computationally effective. As a result, various heuristic techniques are used to guess the proper way to disambiguate homographs, like examining neighbouring words and using statistics about frequency of occurrence.

Recently TTS systems have begun to use HMMs (discussed above) to generate "parts of speech" to aid in disambiguating homographs. This technique is quite successful for many cases such as whether "read" should be pronounced as "red" implying past tense, or as "reed" implying present tense. Typical error rates when using HMMs in this fashion are usually below five percent. These techniques also work well for most European languages, although access to required training corpora is frequently difficult in these languages.

Deciding how to convert numbers is another problem that TTS systems have to address. It is a simple programming challenge to convert a number into words (at least in English), like "1325" becoming "one thousand three hundred twenty-five." However, numbers occur in many different contexts; "1325" may also be read as "one three two five", "thirteen twenty-five" or "thirteen hundred and twenty five". A TTS system can often infer how to expand a number based on surrounding words, numbers, and punctuation, and sometimes the system provides a way to specify the context if it is ambiguous. Roman numerals can also be read differently depending on context. For example "Henry VIII" reads as "Henry the Eighth", while "Chapter VIII" reads as "Chapter Eight".

Similarly, abbreviations can be ambiguous. For example, the abbreviation "in" for "inches" must be differentiated from the word "in", and the address "12 St John St." uses the same abbreviation for both "Saint" and "Street". TTS systems with intelligent front ends can make educated guesses about ambiguous abbreviations, while others provide the same result in all cases, resulting in nonsensical (and sometimes comical) outputs.

2) Text-to-phoneme challenges

Speech synthesis systems use two basic approaches to determine the pronunciation of a word based on its spelling, a process which is often called text-to-phoneme or grapheme-to-phoneme conversion (phoneme is the term used by linguists to describe distinctive sounds in a language). The simplest approach to text-to-phoneme conversion is the dictionary-based approach, where the program stores a large dictionary containing all the words of a language and their correct pronunciations. Determining the correct pronunciation of each word is a matter of looking up each word in the dictionary and replacing the spelling with the pronunciation specified in the dictionary. The other approach is rule-based, in which pronunciation rules are applied to words to determine their pronunciations based on their spellings. This is similar to the "sounding out", or synthetic phonics, approach to learning reading.

Each approach has advantages and drawbacks. The dictionary-based approach is quick and accurate, but completely fails if it is given a word which is not in its dictionary. As dictionary size grows, so too does the memory space requirements of the synthesis system. On the

other hand, the rule-based approach works on any input, but the complexity of the rules grows substantially as the system takes into account irregular spellings or pronunciations. (Consider that the word "of" is very common in English, yet is the only word in which the letter "f" is pronounced [v].) As a result, nearly all speech synthesis systems use a combination of these approaches.
Languages with a phonemic orthography have a very regular writing system, and the prediction of the pronunciation of words based on their spellings is quite successful. Speech synthesis systems for such languages often use the rule-based method extensively, resorting to dictionaries only for those few words, like foreign names and borrowings, whose pronunciations are not obvious from their spellings. On the other hand, speech synthesis systems for languages like English, which have extremely irregular spelling systems, are more likely to rely on dictionaries, and to use rule-based methods only for unusual words, or words that aren't in their dictionaries.

D. Evaluation challenges

The consistent evaluation of speech synthesis systems may be difficult because of a lack of universally agreed objective evaluation criteria. Different organizations often use different speech data. The quality of speech synthesis systems also depends to a large degree on the quality of the production technique (which may involve analogue or digital recording) and on the facilities used to replay the speech. Evaluating speech synthesis systems has therefore often been compromised by differences between production techniques and replay facilities.
Recently, however, some researchers have started to evaluate speech synthesis systems using a common speech dataset.

IV. SYSTEM DESIGN AND COMPONENTS

Occupancy sensors for lighting control use infrared (IR) or acoustic technology, or a combination of the two. The field of view of the sensor must be carefully selected/adjusted so that it responds only to motion in the space served by the controlled lighting. For example, an occupancy sensor controlling lights in an office should not detect motion in the corridor outside the office. Sensors and their placement are never perfect, therefore most systems incorporate a delay time before switching. This delay time is often user-selectable, but a typical default value is 15 minutes. This means that the sensor must detect no motion for the entire delay time before the lights are switched. Most systems switch lights off at the end of the delay time, but more sophisticated systems with dimming technology reduce lighting slowly to a minimum level (or zero) over several minutes, to minimize the potential disruption in adjacent spaces. If lights are off and an occupant re-enters a space, most current systems switch lights back on when motion is detected. However, systems designed to switch lights off automatically with no occupancy, and that require the occupant to switch lights on when they re-enter are gaining in popularity due to their potential for increased energy savings. These savings accrue because in a space with access to daylight the occupant may decide on their return that they no longer require supplemental electric lighting.

V. IMPLEMENTATION

After having the user acceptance of the new system developed, the implementation phase begins. Implementation is the stage of a project during which theory is turned into practice. During this phase, all the programs of the system are loaded onto the user's computer. After loading the system, training of the users starts. Main topics of such type of training are:

- How to execute the package
- How to enter the data
- How to process the data (processing details)
- How to take out the reports

After the users are trained about the computerised system, manual working has to shift from manual to computerised working.
The following two strategies are followed for running the system:

A. Parallel run:

In such run for a certain defined period, both the systems i.e. computerised and manual are executed in parallel. This strategy is helpful because of the following:
Manual results can be compared with the results of the computerised system.
Failure of the computerised system at the early stage, does not affect the working of the organisation, because the manual system continues to work, as it used to do.

B. Pilot run:

In this type of run, the new system is installed in parts. Some part of the new system is installed first and executed successfully for considerable time period. When the results are found satisfactory then only other parts are implemented. This strategy builds the confidence and the errors are traced easily.

1) Application Interface

Login Screen
Input: Username and Password

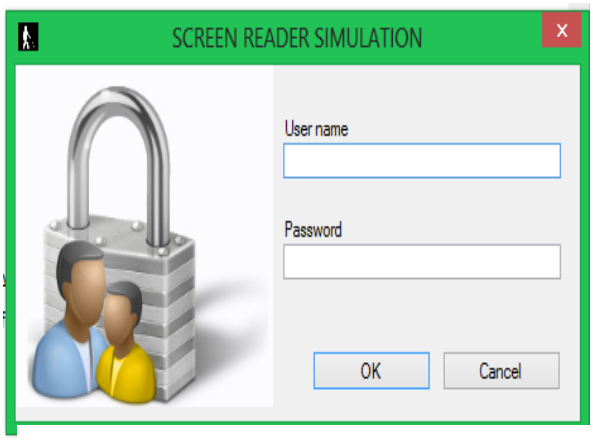


Figure 1: SCREEN READER SIMULATION

In figure 1 above, the administrator will ask an authorized user to enter the user name and password.

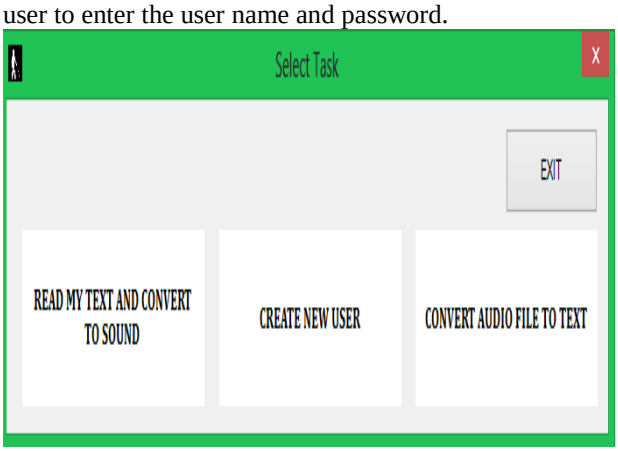


Figure 2: SELECT TASK

In figure 2 above, a user is required to select a task he/she want to perform from the three available options: Read my text and convert to sound, create new user and the conversion of Audio files to text.

Read my text and convert to sound

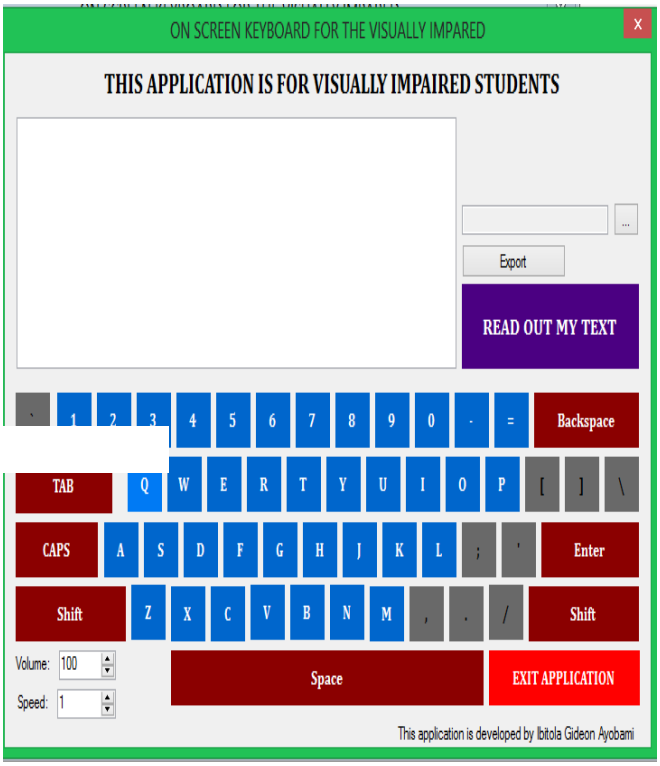


FIGURE 3: ON SCREEN KEYBOARD FOR THE VISUALLY IMPAIRED

Figure 3 above shows the keyboard for the visually impaired students. This reads everything on the screen as the mouse hovers. After the text has been inputted, the user can use the export button to export the text as an audio file (.wav format)

Add new user

FIGURE 4: ADD USER

Figure 4 allows a new user to create an account and the new user will be prompted for username and password.

Convert text file to speech

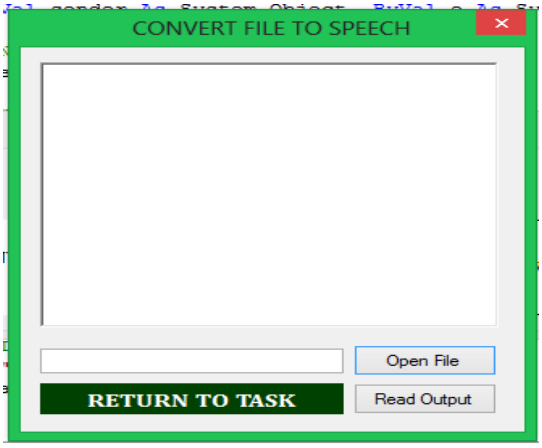


FIGURE 5: CONVERSION OF FILE TO SPEECH

Figure 5 shows the interface for conversion of file to speech. The user can select any .wav audio file then convert to text and can also read the outputted text.

2) Operating Procedure

A standard operating procedure is a set of instructions having the force of a directive, covering those features of operations that lend themselves to a definite or standardized procedure without loss of effectiveness. Standard Operating Policies and Procedures can be effective catalysts to drive performance improvement and improving organizational results. Every good quality system is based on its standard operating procedures (SOPs).

This project work is guarded strictly by some procedures. The information technology industry uses the terms 'Standard Operating Procedure' and SOPs interchangeably to describe a best practice approach to executing tasks related to the production and maintenance of hardware and software, as well as incident and change management.

This proposed software runs effectively on a computer system which has the minimum requirements. The requirements are split into two categories, namely:

Software Requirements: the minimum software requirements to run the program are listed below Microsoft Windows XP (Home and Professional Editions), Microsoft Access 2000, and Microsoft .NET 3.5 framework.

Hardware Requirements: the minimum hardware required to run the program are: Windows XP, Pentium 4, Hard disk of 40 Giga bytes and System memory (RAM) of 512 Mega bytes.

VI. CONCLUSION

This research work is developed to allow the visually impaired people to make use of the system with ease. It allowed the user to type using an on screen keyboard, type text that can be converted to audio files and then the read text from audio file allows the user to output an audio file to text for usage. The scope is that as the mouse moves, the user hears where the cursor is located and have an idea on what to do.

REFERENCES

[1] This article incorporates public domain material from the General Services Administration document "Federal Standard 1037C"

[2] Department of Defence Dictionary of military and associated terms

[3] Ulrich & Eppinger (2000). Product Design & Development. Irwin McGraw-Hill. ISSN 0-07-229647-X.

[4] History and Development of Speech Synthesis, Helsinki University of Technology, Retrieved on November 4, 2006.

[5] Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine ("Mechanism of the human speech with description of its speaking machine," J. B. Degen, Wien).

[6] Mattingly, Ignatius G. Speech synthesis for phonetic and phonological models. In Thomas A. Sebeok (Ed.), Current Trends in Linguistics, Volume 12, Mouton, The Hague, pp. 2451-2487, 1974.

[7] Kurzweil, Raymond "The Singularity is Near". Penguin Books. (2005).

[8] Lambert, Bruce (1992-03-21).. New York Times. http://query.nytimes.com/search/query?ppds=per&v1=GERSTMAN%2C%20LOUIS&sort=newest. Retrieved 2010-02-17.

[9] Arthur C. Clarke online Biography

[10] "Where "HAL" First Spoke (Bell Labs Speech Synthesis website)" Bell Labs. Retrieved 2010-02-17.

Trends in Application Areas of Artificial Intelligence: A Survey

Ayoade Akintayo M.
Computer Science Department,
Lead City University, Ibadan
Ibadan, Nigeria
ayoadeakins2002@yahoo.com

Okesola Kikelomo I,
Computer Science Department,
Lead City University, Ibadan
Ibadan, Nigeria
okesolakikelomo@gmail.com

Deji-Akinpelu Omokehinde O
Computer Science Department,
Lead City University, Ibadan
Ibadan, Nigeria
omokehindeakinpelu@gmail.com

Abstract— Artificial Intelligence (AI) as a field in computer science aims to produce intelligent machine that is, the creation of computer systems that can perceive, reason and act. This paper brings to the fore, some of the trends in the application area of Artificial Intelligence from its inception, through the present and to its future.

***Keywords*—Trends, Artificial Intelligence, Application area**

I. INTRODUCTION

The field of AI research was founded at a conference on the campus of Dartmouth College in the summer of 1956. This was where John McCarthy, Marvin Minsky, N. Rochester and C.E. Shannon, proposed a study of artificial intelligence with the these goals: to make machines use language, form abstractions and concepts, solve kinds of problems reserved for humans and make computers improve themselves [1].

John McCarthy defined the term as "the science and engineering of making intelligent machines". Consequently, traditional characteristics of AI emphasized human-level performance with the broad goal of simulating intelligence. Every intelligent system is expected to display certain capabilities, such as reasoning, problem solving traits, representation of knowledge, planning, learning and natural language processing.

AI can be evaluated from different views based on existing approaches according to [2]. Classical approach also referred to as Computational or Cognitive Approach: viewed AI as a computational process which mainly involves computer science, psychology, philosophy, and linguistics. While computer and engineering applications approach also referred to as embodied approach, viewed this field as a paradigm that employs the synthetic methodology with the following goals: understanding biological systems, abstracting general principles of intelligent behavior, and the application of this knowledge to build artificial systems such as robots or intelligent devices, which involve computer science, philosophy, engineering, robotics, biology, and neuroscience. This triggered the movement of AI out of computer science laboratories into robotics and engineering or biology laboratories.

While researches linked up to any of the subfield of these approaches are focused broadly on: Systems that think like humans, Systems that act like humans, Systems that think rationally (intelligently) and Systems that act rationally (intelligently) [3].

1961 marked the era of industrial robots where the first industrial robot worked on a general motors' assembly line. In the 70's, AI realized its first useful projects with real applications of intelligent systems. While the first successes provoked many projects in the early 80's, promising important role of artificial intelligence in computer technology. The end of the 80's to the beginning of 90's, AI projects failed to realize promised goals, and were criticized for inefficiency. AI achieved its greatest successes in 1990s and early 21st century application areas of AI as at then include: logistics, data mining, medical diagnosis and many other areas throughout the technology industry.

II. ARTIFICIAL INTELLIGENCE PRESENT TRENDS

The field of AI comprises of many areas that are not perfectly communicating as can be seen in the diverse dimension of its growth. Some of the current growth to mention a few are: Probabilistic graphical models with the use of Bayesian network, Machine learning algorithms with the use of decision trees, Support vector machines (SVMs), Boosting and Gaussian processes. Other areas are: Probabilistic relational models, Deep networks, Application of Bayesian network for tracking a person using GPS and robot learning by imitation, Controls for appliances using fuzzy logic. Decision making under uncertainty is a key aspect of natural intelligence that AI seeks to emulate.

A. Deep Learning

North Carolina State University researchers, in their work,, "Deep Learning-Based Goal Recognition in Open-Ended Digital Games", [4] used a family of machine learning techniques that can extrapolate patterns from large collections of data to make predictions known as deep learning; to develop an artificial intelligence (AI) software that can make accurate prediction about what goal a player is trying to achieve in a video game.

According to Dr. James Lester, this software was developed for use in educational gaming but it is also applicable for all video developers and it is a key step in developing player-adaptive games that can respond to player

actions to improve the gaming experience for both entertainment and education.

In order to evaluate the performance of this AI software, the researchers extracted the logs of the behaviours of 137 different players using "Crystal Island" educational game and fed the data into the software. Its predictive power in goal recognition is used to infer what an individual player is trying to accomplish, based on his or her actions at any given point in the game. Its performance increases with time as more data is made available to it. By checking the AI's response against the player log, the researchers found out that its prediction is satisfactory.

B. Fingertip Sensor

A sensor known as GelSight which was developed in 2009 by the lab of Edward Adelson, the John and Dorothy Wilson Professor of Vision Science at MIT, was modified by researcher in MIT and North-eastern University to extend the capability of the robot [5]. The enhanced GelSight sensor referred to as "Fingertip sensor" is tactile and it allows the robot to grasp a USB cable and insert it into a USB port. It is also smaller and has a faster processing algorithm than the original GelSight even though it is not as sensitive.

The researchers which includes a MIT team consists of Adelson, Rui Li, Wenzhen Yuan, and Mandayam Srinivasan, designed and built the sensor while Platt's team at Northeastern, is comprised of Andreas ten Pas and Nathan Roscup, who developed the robotic controller and conducted the experiments.

The Fingertip sensor consists of a slab of synthetic rubber coated on one side with a metallic paint that is transparent. The rubber adapts to any object it is pressed against, and the metallic paint smoothens out the light-reflective properties of diverse materials. This makes precise optical measurements possible. It is mounted in a cubic plastic housing, with an exposed paint-covered face. The four walls of the cube adjacent to the sensor face are translucent, and each conducts a different colour of light red, green, blue, or white emitted by light-emitting diodes at the opposite end of the cube.

The algorithms developed by Adelson's team can infer the three-dimensional structure of ridges or depressions of the surface against which the sensor is pressed from the different intensities of the different-colour light.

In the experiments, a Baxter robot from MIT spinout Rethink Robotics was equipped with a two-pincer gripper, one of whose pincers had a Fingertip sensor on its tip. Using conventional computer-vision algorithms, the robot identified the dangling USB plug and attempted to grasp it. It then determined the position of the USB plug relative to its gripper from an embossed USB symbol. Although there was a 3-millimetre variation, in each of two dimensions, in where the robot grasped the plug, it was still able to insert it into a USB port that tolerated only about a millimetre's error. [5]

C. Computerized Emotion Detector

A research published in the International Journal of Computational Vision and Robotics by [6] takes face recognition a step further by recognizing the emotion portrayed by a face. There is three-stage approach to a software emotion detector. The first has to do with the development of an algorithm for precise identification and definition of features on the human face. The second stage involves the analysis of the particular positions and shapes of the face. The third stage associates those features with a person's emotional state to decide whether they are happy, sad, angry, surprised, fearful or disgusted. Preliminary tests gave a 94 per cent success rate the team reports.

Some of the application areas of this emotion detector could be apprehension of would-be criminal through the recognition of negative emotions (like fear, anger etc.) at check points which might serve as a pro-active measure against terrorists while those with no negative facial expression can be physically checked and an automatic change of background or music based on the facial expression of the users.

D. Robo Brain

Researchers from Cornell University [7] developed a large scale artificial intelligent computational system called "Robo-brain" using a structural deep learning where information is stored in many levels of abstraction. It can learn from publicly available Internet resources can download and process thousands of YouTube videos, millions of how-to documents and appliance manuals and more than a billion images with the information being translated and stored in a robot-friendly format that robots will be able to draw on when they need it.

According to the researchers, a robot's computer brain stores information about the world in a form of a Markov model, which can be represented graphically as a set of points (nodes) connected by lines (edges). The nodes could represent objects, actions or parts of an image, and each one is assigned a probability i.e. how much it can be varied and still be correct.

In searching for knowledge, a robot's brain makes its own chain and looks for one in the knowledge base that matches within those limits. It is a huge, branching graph with abilities for multi-dimensional queries, according to Aditya Jami, a visiting researcher at Cornell, who designed the large-scale database for the brain. It will process images to pick out the objects in them, and by connecting images and video with text, it will learn to recognize objects and how they are used, along with human language and behaviour.

E. Robotic Cheetah

MIT researchers implemented in a robotic cheetah an algorithm for bounding they developed. The robotic cheetah is a sleek, four-legged assemblage of gears, batteries, and electric motors that weighs about as much as its real counterpart. In experiments performed on Killian Court track in MIT, it garnered up to 10 mph while sprinting along the lawn.

According to [8], the key to the bounding algorithm is in programming each of the robot's legs to exert a certain amount

of force in the split second during which it hits the ground, in order to maintain a given speed. In general, the faster the desired speed, the more force must be applied to propel the robot forward. This is in tune to the real world scenario in which sprinters increase their stride by pushing harder on the ground in other to gain more speed running forward. In experiments also carried out on treadmill terrain, it was discovered that the cheetah-bot was able to even adapt to slight bumps in its path.

Robots are known to be slow and heavy because of their weight and as a result are not capable of force control but the researcher claimed these shortcomings were catered for by a custom-design, high-torque-density electric motors that are controlled by amplifiers all of which together with a bio-inspired legs allowed it to have a force control on the ground without relying on delicate force sensors on the feet.

F. New Bracelet Strengthens Computer Security

Researchers from Dartmouth College [9], invented a new computer security approach that authenticates users continuously while using a terminal and automatically log them out when they leave or when someone else steps in to use their terminal.

This new approach called, Zero-Effort Bilateral Recurring Authentication (ZEBRA) was developed to address security risks inherent in the widely used authentication methods based on passwords, fingerprints or tokens which perform one-time authentication and rely on users to log out from their terminal when they leave. This is risky because users most of the times do not remember to log out.

One of the existing solutions proffered for this problem (inactivity timeouts) is inadequate as it suffers from being too long or short. Other solutions like user proximity and behavioural biometric are not reliable too as they provide insufficient information (user may be nearby the terminal without using it).

According to [9], in ZEBRA, a user wears a bracelet with a built-in accelerometer, gyroscope and radio on his or her dominant wrist which captures, records, processes and sends the wrist movement to a computer terminal. Also as the user interacts with the terminal, the terminal compares the wrist movement inputs it receives from the user via keyboard and mouse and confirms the continued presence of the user only if they correlate with those from the bracelet as this will imply that source of the two inputs are the same.

Experiments performed using ZEBRA as continuous authentication method yielded 85 per cent accuracy in verifying the correct user and identified all adversaries within 11 seconds while it produced 90 per cent of accuracy for correct users and identified all adversaries within 50 seconds when security is traded for usability. Overall, ZEBRA can recognize (in under a minute) when an unauthorized person steps in to use the terminal if the original user has stepped away from the terminal.

In the future, ZEBRA can be extended to mobile devices, such as smartphones or tablet computers, TV remotes, game controllers and medical devices where the user provides frequent inputs with his or her hand.

III. ARTIFICIAL INTELLIGENCE FUTURE

AI is growing to what is known as sensor-rich world, in which intelligence will be seen in every aspect of our day to day activities [10]. The future of AI will be based on Massive amounts of data, sophisticated probabilistic reasoning and effective machine learning algorithms, backed up with Massive computing power.

Artificial Intelligence is paramount to the futurist agenda. Having discussed the past and the current trends of Artificial Intelligence, it is imperative to briefly discuss its future trend. Research has shown that the U.S. Department of Defense wants future generations of fighter aircraft to come with copilots already installed. According to the U.S. Naval Institute, both the Navy and the Air Force are interested in having their next generation air superiority fighters built with Artificial Intelligence [11]. They are planning next-generation fighters that don't have just a human pilot. Future fighter jets may have an Artificial Intelligence co-pilot on board that can help with sensory data in addition to autonomously landing the plane on an aircraft carrier.

Although the exact purpose of AI in aircraft is not yet certain, it is very possible it has a number of uses from acting as a co-pilot to carrying out autonomous landings. Popular Science stated that the robotic co-pilots would be especially helpful for aircraft carriers, which are difficult to land on. It was reported recently that the Navy's X-47B experimental drone landed autonomously on an aircraft, demonstrating the usefulness of AI in such a situation.

Artificial Intelligence could also be useful in advanced sensor fusion. This involves combining data from various sensors to get a more accurate reading of a situation or location. The Navy and Air Force are hopeful to get assistance from industry experts in Silicon Valley to facilitate the design of such sixth-generation fighters.

In the next decade, AI will be able to communicate with humans in unstructured English using text or voice, navigate (not perfectly) in an unprepared environment and will have some rudimentary common sense (and domain-specific intelligence). The number of practical applications based on digitally recreated aspects human intelligence, such as cognition, perception, rehearsal learning, or learning by repetitive practice; will be on the increase.

As part of advances in computer vision, and aspect of AI; AI will be required to make the vision of Augmented Reality (AR) mature. AR, as exemplified by Google Glass, uses computer glasses to overlay graphics on the real world. It needs to quickly analyze what the viewer sees and generates graphics that provide useful information. To be useful, the glasses have to be able to identify complex objects from any direction, under any lighting conditions, no matter the weather. To be useful to a driver, for instance, the glasses would need to identify roads and landmarks faster and more

effectively than is enabled by any current technology. AR is still far from here, but should be achievable within the next ten years.

There are many other interesting areas in AI, but machine vision, cognitive prostheses, and robotic scientists are among the most exciting, and relevant to futurist goals.

REFERENCES

[1] McCarty John, Rochester Nathaniel, and Shannon E. Claude (1955): A proposal for the Dartmouth summer research project on artificial intelligent", AI magazine.

[2] Rolf Pfeifer and Fumiya Iida (2003): Embodied artificial intelligence: trends and challenges, Volume 2865 of Lecture Notes in Computer Science, page 1-26. Springer

[3] Joost N. Kok, Egbert J.W.Boers, Walter A. Kosters, and Peter Van der Putten, and Mannes Poel (2002). Artificial intelligence: Definition, trends, techniques and cases. Encyclopedia of life support systems (EOLSS) Volume 1 pp 1-20

[4] Wookhee Min, Eun Young Ha, Jonathan Rowe, Bradford Mott and James Lester, (2014): Deep Learning-Based Goal Recognition in Open-Ended Digital Games. Association for the Advancement of Artificial Intelligence.

[5] Massachusetts Institute of Technology. (2014, September 19): Fingertip sensor gives robot unprecedented dexterity. *ScienceDaily*. Retrieved September 22, 2014 from www.sciencedaily.com/releases/2014/09/140919122218.htm

[6] Agrawal, D.D., Dubey S.R., and Jalal A.S (2014): Emotion recognition from facial expressions based on multi-level classification. International Journal of Computational Vision and Robotics, Vol 4, No 4, pp 365

[7] Cornell University. (2014, August 25): 'Robo Brain' will teach robots everything from the Internet. *ScienceDaily*. Retrieved September 22, 2014 from www.sciencedaily.com/releases/2014/08/140825084931.htm

[8] Massachusetts Institute of Technology. (2014, September 15): Run, cheetah, run: New algorithm enables cheetah robot to run and jump, untethered, across grass. *ScienceDaily*. Retrieved September 25, 2014 from www.sciencedaily.com/releases/2014/09/140915114524.htm

[9] [9] Dartmouth College. (2014, September 22): New bracelet strengthens computer security. *ScienceDaily*. Retrieved September 25, 2014 from www.sciencedaily.com/releases/2014/09/140922113010.htm

[10] Klicek Bozidar, (1997): "Survey of the state and future trends of intelligent systems", original scientific paper.

[11] Popular Science. (2014, February 9): Pentagon wants Artificial Intelligence in Future fighters. Retrieved September 27, 2014 from http://www.popsci.com/article/technology/pentagon-wants-artificial-intelligence-future-fighters

An Adaptive Modeling and Execution Framework for a Knowledge-Based Intelligent Clinical Decision Support System to Predict Schizophrenia

Adekunle Y.A
Computer Science Department
Babcock University,
Ilishan-Remo, Ogun State, Nigeria.
adekunleya@gmail.com

Ebiesuwa Seun
Computer Science Department,
Babcock University,
Ilishan-Remo, Ogun State, Nigeria.
seunebi@gmail.com

Abstract— This paper proposes an adaptive framework for a Knowledge Based Intelligent Clinical Decision Support System for the prediction of schizophrenia which is one of the most deadly illnesses that has a monumental effect on the health of people afflicted with it and has for long remained a perennial health problem affecting a significant number of people the world over. In the framework the patient information is fed into the system; the Knowledge base stores all the information to be used by the Clinical Decision Support System and the classification/prediction algorithm chosen after a thorough evaluation of relevant classification algorithms for this work is the C4.5 Decision Tree Algorithm with its percentage of correctly classified instances given as 61.0734%; it searches the Knowledge base recursively and matches the patient information with the pertinent rules that suit each case and thereafter gives the most precise prediction as to whether the patient is prone to schizophrenia or not. This approach to the prediction of schizophrenia provides a very potent solution to the problem of determining if a person has the likelihood of developing this dreaded illness or is almost not susceptible to the ailment.

Keywords: **Schizophrenia, Clinical Decision Support System (CDSS), Medical Decision Support System (MDSS), Support Vector Machine (SVM) and Sequential Minimal Optimization (SMO).**

I. INTRODUCTION

Our world is fast evolving and in order to cope with the insatiable demand of the human race for the kind of living that can be described as top-notch in which people have all they need at their beck and call, there is the need to develop intelligent decision making applications that will drive systems or devices to carry out tasks that require human intelligence. This concept is known as Artificial Intelligence (AI). Artificial Intelligence has a variety of sub-fields which include expert systems, machine vision, machine learning and natural language processing amongst others.

A Decision Support System is an interactive computer-based system intended to help decision makers utilize data and models in order to identify and solve problems and make decisions [1]. According to the Clinical Decision Support (CDS) Roadmap project, CDS is "providing clinicians, patients, or individuals with knowledge and person-specific or population information, intelligently filtered or present at appropriate times, to foster better health processes, better individual patient care, and better population health."

A Clinical Decision Support System (CDSS) is an active knowledge system, where two or more items of patient data are used to generate case-specific recommendation(s) [2]. This implies that a CDSS is a decision support system (DSS) that uses knowledge management to achieve clinical advice for patient care based on some number of items of patient data. This helps to ease the job of healthcare practitioners, especially in areas where the number of patients is overwhelming.

According to the National Alliance on Mental Illness, mental illnesses are medical conditions that disrupt a person's clear thinking, feeling, mood, ability to relate to others, decision making ability and daily functioning. Mental illnesses are brain-based conditions that often result in a variety of symptoms that affect daily life. Some symptoms of mental illness include feeling sad, confused thinking, reduced ability to concentrate, excessive fears or worries, extreme mood changes of highs and lows, social withdrawal, sleeping problems, delusions, paranoia, hallucinations, extreme feelings of guilt, major changes in eating habits and suicidal thinking amongst others.

Mental illnesses include schizophrenia, depression, bipolar disorder, obsessive-compulsive disorder (OCD), posttraumatic stress disorder (PTSD), borderline personality disorder, anxiety disorder and others. Translated from Greek, the word schizophrenia means "split-brain". The so-called split is between beliefs, thoughts, perceptions, emotions and other brain functions.

The word schizophrenia is less than 100 years old. However, the disease was first identified as a mental disorder by Dr. Emil Kraepelin in 1887 and the illness is generally believed to have accompanied mankind throughout its history. Written documents that identify schizophrenia can be traced to the old Pharaonic Egypt, as far back as the second millennium before Christ. One of the first to classify mental disorders into different categories was the German physician, Emil Kraepelin who used the term "dementia praecox" for individuals who had symptoms we now associate with schizophrenia. The

nonspecific concept of madness has been around for many thousands of years and Kraepelin was the first to make a distinction in the psychotic disorders between what he called dementia praecox and manic depression. Kraepelin believed that dementia praecox was primarily a disease of the brain, and particularly a form of dementia. He named the disorder "dementia praecox" (early dementia) to distinguish it from other forms of dementia such as Alzheimer's disease which typically occurs late in life. The Swiss psychiatrist Eugen Bleuler, coined the term "schizophrenia" in 1911. He was the first to describe the symptoms as positive or negative.

Schizophrenia is defined as a chronic, more or less debilitating illness characterized by perturbations in cognition, affect and behavior, all of which have a bizarre aspect (Andreasen, 1982). Major symptoms of schizophrenia include incoherent thinking, delusions, hallucinations, apathy, alogia, avolition, disturbance of affect or emotional experience and bizarre behavior. Schizophrenia can be caused by a variety of factors which include the imbalance of two brain chemicals namely dopamine and serotonin, intake of narcotics, heredity, viral infections and highly stressful situations. Schizophrenia is classified into five sub-types which are paranoid, catatonic, hebephrenic, schizo-affective disorder and undifferentiated.

Like most other illnesses, the detection of schizophrenia is based on the symptoms shown by the individual in question, however this conventional method is not accurate enough because sometimes a person might show symptoms that suggest schizophrenia meanwhile he or she might be suffering from another mental ailment. The tendency of developing schizophrenia is traceable to the genes. The adoption studies of Kety et al and Heston lend credence to the fact that some biologically transmitted genes increase the risk of an individual becoming schizophrenic. Thus there is the need for decision support systems that can predict schizophrenia based on genetic information.

II. RELATED WORKS

A. Decision Support System for Heart Disease Based on sequential Minimal Optimization in Support Vector Machine

Here, Vadicherla & Sonawane (2013), the proponents of this system claim that computer based Medical Decision Support System (MDSS) is useful for the physicians with its fast and accurate decision making process. They opined that predicting the existence of heart disease accurately, results in saving the lives of patients followed by proper treatment. Their objective was to present a MDSS for heart disease classification based on sequential minimal optimization (SMO) technique (which incorporated its features like high accuracy and high speed) in support vector machine (SVM). In using this method, they illustrated the UCI (University College Irvine) machine learning repository data of Cleveland heart disease database and consequently trained the SVM by using SMO technique. Hence, they also claim that given the ease of use and better scaling with the training set size, SMO is a strong candidate for becoming the standard SVM training algorithm. Training a SVM requires the solution of a very large QP (Quantum Platform) optimization problem. SMO algorithm breaks this large optimization problem into small sub-problems. Both the training and testing phases give the accuracy on each record. The results proved that the MDSS is able to carry out heart disease diagnosis accurately in a fast way and it was reported to show good ability of prediction on a large dataset.

B. Data Mining in Clinical Decision Support Systems for Diagnosis and Treatment of Heart Disease.

According to Amin, Agarwal & Beg (2013) medical errors are both costly and harmful. Medical errors cause thousands of deaths worldwide each year. Hence, a clinical decision support system (CDSS) would offer opportunities to reduce medical errors as well as to improve patient safety. They affirm that one of the most important applications of such systems is in diagnosis and treatment of heart diseases (HD). This is because statistics have shown that heart disease is one of the leading causes of deaths all over the world (CDC Report). Data mining techniques have been very effective in designing clinical support systems because of its ability to discover hidden patterns and relationships in medical data. Here, the proponents also undertook a comparative analysis of the performance and working of six CDSS systems which use different data mining techniques for heart disease diagnosis. They conclude by asserting based on their findings that there is no system to identify treatment options for Heart disease patients. They further claimed that in spite of having a large amount of medical data, it lacked in the quality and the completeness of data thereby creating the need for highly sophisticated data mining techniques to build up an efficient decision support system. They claim that even after doing this, the overall reliability and generalization capability might still be questionable. Hence, the need to build systems which will be accurate, reliable as well as reduce the cost of treatment and improve patient care. More so, the building of systems which are understandable and which could enhance human decisions are very germane.

C. An Intelligent Decision Support System for the Operating Theater

In 2013, Sperandio, Gomes, Borges, Brito and Almada-Lobo asserted that decision processes inherent in operating theatre organization are often subjected to experimentation, which sometimes lead to far from optimal results. They further affirm that the waiting lists for surgery had always been a societal problem, with governments seeking redress with different management and operational stimulus plans partly due to the fact that the current hospital information systems available in Portuguese public hospitals, lack a decision support system component that could help achieve better planning solutions. As such they developed an intelligent decision support system that allows the centralization and standardization of planning processes which improves the efficiency of the operating theater and tackles the fragile situation of waiting lists for surgery. The intelligence of the system is derived from data

mining and optimization techniques, which enhance surgery duration predictions and operating rooms surgery schedules.

D. Decision Support System for the Diagnosis of Schizophrenia Spectrum Disorders

In 2013, Kahn, Perkins and Lieberman developed a decision support system for the diagnosis of schizophrenia spectrum disorders.

The development of this system is described in four-stages: knowledge acquisition, knowledge organization, the development of a computer-assisted model, and the evaluation of the system's performance. The knowledge is extracted from an expert through open interviews. These interviews aimed at exploring the expert's diagnostic decision making process for the diagnosis of schizophrenia. A graph methodology was employed to identify the elements involved in the reasoning process. Knowledge was first organized and modeled by means of algorithms and then transferred to a computational model created by the covering approach. The performance assessment involved the comparison of the diagnosis of 38 clinical vignettes between an expert and the decision support system. The results showed a relatively low rate of misclassification (18-34%) and a good performance by the decision support system in the diagnosis of schizophrenia, with an accuracy of 66-82%.

E. HIROFILOS: A Medical Expert System for Prostate Diseases (Constantinos Koutsojannis, Maria Tsimara & Eman Nabil, 2008)

In this study a fuzzy expert system for diagnosis and learning purpose of prostate diseases was described. HIROFILOS is a fuzzy expert system for diagnosis and treatment of prostate diseases according to symptoms that are realized in one patient and usually recorded through his clinical examination as well as specific test results. The user-friendly proposed intelligent system is accommodated on a hospital web page for use as a decision support system for resident doctors, as an educational tool for medical students, as well as, an introductory advisory tool for interested patients. It is based on knowledge representation provided from urology experts in combination with rich bibliographic search and study ratified with statistical results from clinical practice. Preliminary experimental results on a real patient hospital database provide an acceptable performance that can be improved using more than one computational intelligence approach in the future.

III. BRIEF LITERATURE REVIEW

Prior to the inception and subsequent adoption of intelligent systems for use in the health care domain, the vast majority of the world's population were at the mercy of the healthcare professionals who to a large degree decided the fate of the teeming patients. This era was marked with so many deaths that would otherwise have been avoided if there were intelligent systems that could replace the medical experts at times when they are not available. In view of the foregoing, endless research has been conducted over the years to come up with expert systems, decision support systems and other kinds of intelligent systems to provide a readily available alternative should the medical experts be unavailable.

Onuiri, Awodele & Idowu in (2014) proposed a framework for a Knowledge-Based Intelligent Clinical Decision Support System to predict the cormobidity between cataract and glaucoma disorders. The proposed framework was named Comorbidity Ontological Framework for Intelligent Prediction (COFIP). COFIP consists of the following sections:

i) Knowledge Representation/Preprocessing ii) Knowledge-Based Warehouse iii) Results Classification/Prediction Algorithm iv) Intelligent Clinical Pathway Generator System

The two algorithm used in the COFIP framework are Artificial Neural Networks (ANN) and Decision Trees (DT). The two are used to complement each other's strengths and weaknesses as the case may be. The result from ANN and DT are compared for optimization and the best values are chosen and used to generate the rules.

The rules generated are made from the result of the optimization and they are used to update the precondition sets A and B (cataract and glaucoma) contents in the knowledge base warehouse of the COFIP. Once the condition is ascertained, the patient-specific care plan is generated (Onuiri, Awodele & Idowu, 2014).

The Comorbidity Ontological Framework for Intelligent Prediction (COFIP) proposed by (Onuiri, Awodele & Idowu, 2014) served as the standard framework that was adopted in the design of the Adaptive Schizophrenia Modeling and Execution (SME) framework.

D. Knowledge-based Systems

An example of a knowledge based system is the expert system. It consists of components which include the knowledge-base (rules), inference engine, working memory (facts) and a user interface. The knowledge-base can be the production rules that are in form of "IF condition THEN action". The condition portion can be called antecedent and it represents a particular fact. If some facts are in the database then the action is performed. The action portion is called consequent.

The knowledge-base is responsible for storing factual and heuristic knowledge using one or more knowledge representation schemes to express knowledge about the application domain of interest. The inference engine is the mechanism that is used to manipulate symbolic information and knowledge so as to solve problems through reasoning. The inference engine determines which rule antecedent are satisfied by the fact. The working memory houses the data that is specific to the problem being solved. The individuals who interact with the system include the domain expert (the human expert whose domain knowledge is being sought), knowledge engineer (the one who encodes the expert's knowledge in a declarative form that can be used by the expert system), system

engineer (the one who builds the user interface, designs the declarative format of the knowledge-base and implements the inference engine) and the user (the one who will consult the system to get advice as would have been provided by the expert) [8].

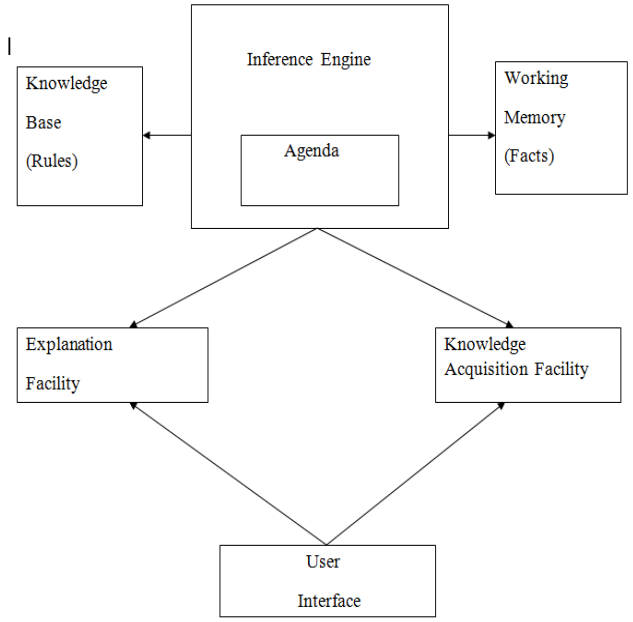


Fig. 1. Components of an Expert System *(Source: Miller, 2010)*

IV. CLINICAL DECISION SUPPORT SYSTEM (CDSS)

The clinical decision support system is another example of a knowledge based system. A clinical decision support system is an active knowledge system where two or more items of patient data are used to generate case specific recommendations [2].

A. Target Area of care

CDSSs assist doctors in assessing various clinical issues from accurate diagnosis of a particular disease to the treatment of the disease. The general target areas of care for CDSS are:

Preventive care which has to do with screening and disease management
Diagnosis which is done based on the patients' signs and symptoms
Follow-up management which has to do with frequent checkups
Hospital Provider Efficiency [9].

B. System Design

The system design for CDSS will usually include the following subsystems:

Communication which handles notification and alerts
Knowledge discovery which deals with rules and regulations
Knowledge repository which contains problem solving knowledge [10].

C. Factors leading to successful CDSS implementation

The following under listed factors lead to the successful implementation of CDSS:

Simple, user friendly interface
Automated decision support
Timely result
Workflow integration
Continuous Knowledge-base and update support [11].

V. PATTERN CLASSIFICATION METHODS

E. Pattern classification refers to the theory and algorithms of assigning abstract objects into distinct categories, where these categories are typically known in advance. For this research, the pattern classification methods considered are Decision Trees (DTs), K-Nearest Neighbor (KNN), Naïve bayes Classifier and Support Vector Machine (SVM).

A. Decision Trees

A decision tree consists of a root node, branch nodes and leaf nodes. The tree begins with a root node, then further splits into branch nodes and each node represents a choice among various alternatives. The tree then terminates with leaf nodes which are un-split nodes that represent a decision [12]. The classification of decision trees are carried out in two phases:

Tree Building or top down: This is computationally intensive and requires the tree to be recursively partitioned until all the data items belong to the same class.

Tree pruning or bottom top: It is conducted to improve the prediction and classification of the algorithm and minimize the effects of over-fitting which may lead to misclassification of errors [13].

Some notable decision tree algorithms include Classification and Regression Trees (CART), Iterative Dichotomiser 3 (ID3), C4.5 and C5.0.

The advantages of decision trees include:

- They are easy to interpret and comprehend
- They can handle both metric and non-metric data as well as missing values which are frequently encountered in clinical studies.
- Little data preparation is required since data does not need to be normalized.
- They can handle data in a short time frame.
- They can be developed using common statistical techniques.

The disadvantages associated with decision trees include:

- They can over fit the data and create complex trees that may not generalize well.
- A small change in the size of a dataset could result in a completely different tree

B. K-Nearest Neighbor

K-Nearest Neighbor (k-NN) is instance based learning for classifying objects based on closest training examples in the feature space. It is a type of lazy learning where the function is only approximated locally and all computations are deferred until classification. The k-nearest neighbor algorithm is amongst the simplest of all machine learning algorithms: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors. If k=1, then the object is simply assigned to the class of its nearest neighbor. The k-NN algorithm uses all labeled training instances as a model of the target function. During the classification phase, k-NN uses a similarity-based search strategy to determine a locally optimal hypothesis function. Test instances are compared to the stored instances and are assigned the same class label as the k most similar stored instances.

C. Bayes Classifier

A Bayesian network is a model that encodes probabilistic relationships among variables of interest. This technique is generally used for intrusion detection in combination with statistical schemes, a procedure that yields several advantages, including the capability of encoding interdependencies between variables and of predicting events, as well as the ability to incorporate both prior knowledge and data. However, a serious disadvantage of using Bayesian networks is that their results are similar to those derived from threshold-based systems, while considerably higher computational effort is required.

D. Support Vector Machine

Support Vector Machines have been proposed as a novel technique for intrusion detection. An SVM maps input (real-valued) feature vectors into a higher-dimensional feature space through some nonlinear mapping. SVMs are developed on the principle of structural risk minimization. Structural risk minimization seeks to find a hypothesis (h) for which one can find lowest probability of error whereas the traditional learning techniques for pattern recognition are based on the minimization of the empirical risk, which attempt to optimize the performance of the learning set. Computing the hyper plane to separate the data points i.e. training an SVM leads to a quadratic optimization problem. The implementation of SVM intrusion detection system has two phases which are training and testing. SVMs can learn a larger set of patterns and be able to scale better, because the classification complexity does not depend on the dimensionality of the feature space. SVMs also have the ability to update the training patterns dynamically whenever there is a new pattern during classification.

VI. METHODOLOGY

A very comprehensive dataset (Duke, 2007) consisting of 100,000 instances compiled from the NIMH (National Institute of Mental Health) data repository was used. This dataset was translated into the Attribute Relational File Format (ARFF) which is one of the file formats recognized by the WEKA (Waikato Environment for Knowledge Analysis) software in which the five distinct attributes used for this work were highlighted. These attributes are Genetic Mutation, Genome Build, Chromosome, Map and Log P. Hence, the dataset consists of five major columns, each representing the respective attribute.

The dataset was then induced with Classification algorithms namely C4.5 decision trees, Support Vector Machine (SVM), K-Nearest neighbor algorithm and Bayes Classifier Algorithm. The Classification algorithms were evaluated using the Waikato Environment for Knowledge Analysis software version 3.6.7 based on the percentage of correctly classified instances with the C4.5 decision trees having 61.0734%, the Support Vector Machine (SVM) algorithm had 50.0515%, the Bayes Classifier Algorithm had 50.2045% and the K-Nearest Neighbor algorithm had 50.1235%. Sequel to the result obtained from this evaluation, the C4.5 decision trees turn out as the Classification algorithm with the highest accuracy for this research. Thereafter, a decision tree program was written in Java with 35 lines of code for the core program to implement the C4.5 decision tree algorithm that will provide the requisite intelligence for this Clinical decision support system and help it make the right decisions promptly when supplied with patient information. The C4.5 decision tree algorithm will be embedded in the classification/prediction algorithm section of the clinical decision support system.

TABLE 1. COMPARISON OF SOME PATTERN CLASSIFICATION ALGORITHMS *(Source: Patel et al, 2012)*

Classifier	**Method**	**Parameters**	**Advantages**	**Disadvantages**
Support Vector Machine	A support vector machine constructs a hyper plane or set of hyper planes in a high or infinite dimensional space, which can be used for classification, regression or other tasks.	The effectiveness of SVM lies in the selection of kernel and soft margin parameters. For kernels, different pairs of (C, γ) values are tried and the one with the best cross-validation accuracy is picked. Trying exponentially growing sequences of C is a practical method to identify good parameters.	1. Highly Accurate 2. Able to model complex nonlinear decision boundaries 3. Less prone to over fitting than other methods	1. High algorithmic complexity and extensive memory requirements of the required quadratic programming in large-scale tasks. 2. The choice of the kernel is difficult 3. The speed both in training and testing is slow.
K Nearest Neighbour	An object is classified by a majority vote of its neighbours, with the object being assigned to the class most common amongst its *k* nearest neighbours (*k* is a positive integer). If $k = 1$, then the object is simply assigned to the class of its nearest neighbour.	Two parameters are considered to optimize the performance of the kNN, the number k of nearest neighbour and the feature space transformation.	1. Analytically tractable. 2. Simple in implementation 3. Uses local information, which can yield highly adaptive behaviour 4. Lends itself very easily to parallel implementations	1. Large storage requirements. 2. Highly susceptible to the curse of dimensionality. 3. Slow in classifying test tuples.
Bayesian Method	Based on the rule, using the joint probabilities of sample observations and classes, the algorithm attempts to estimate the conditional probabilities of classes given an observation.	In Bayes, all model parameters (*i.e.*, class priors and feature probability distributions) can be approximated with relative frequencies from the training set.	1. Naïve Bayesian classifier simplifies the computations. 2. Exhibit high accuracy and speed when applied to large databases.	1 The assumptions made in class conditional independence. 2. Lack of available probability data.
Decision Tree	Decision tree builds a binary classification tree. Each node corresponds to a binary predicate on one attribute; one branch corresponds to the positive instances of the predicate and the other to the negative instances.	Decision Tree Induction uses parameters like a set of candidate attributes and an attribute selection method.	1. Construction does not require any domain knowledge. 2. Can handle high dimensional data. 3. Representation is easy to understand. 4. Able to process both numerical and categorical data.	1. Output attribute must be categorical. 2. Limited to one output attribute. 3. Decision tree algorithms are unstable. 4. Trees created from numeric datasets can be complex.

VII. ADOPTED FRAMEWORK

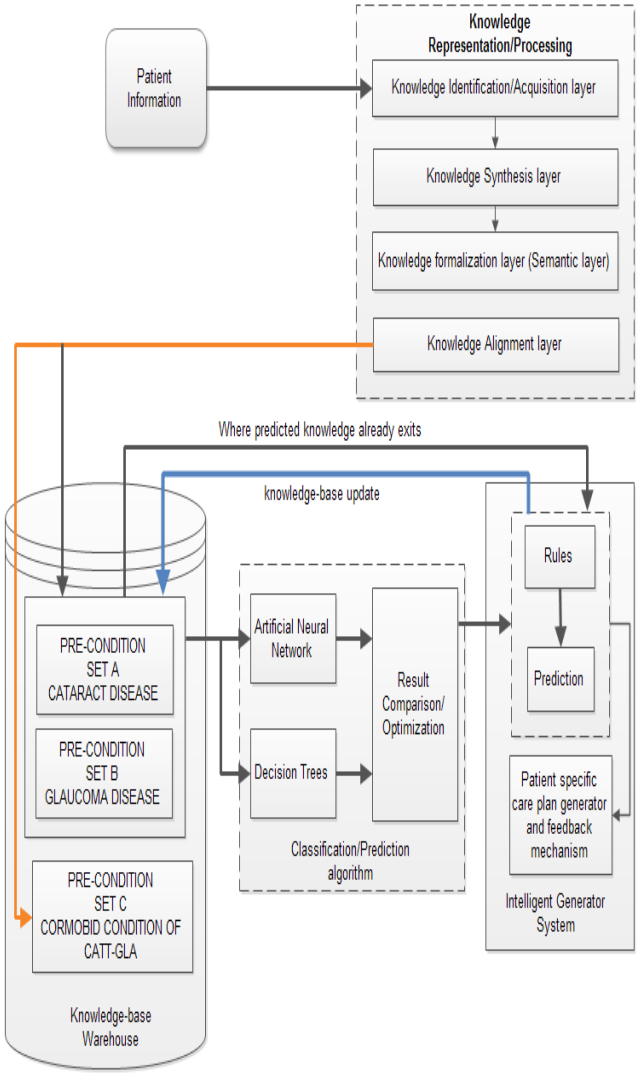


Fig. 2. Comorbidity Ontological Framework for Intelligent Prediction (COFIP) *(Source: Onuiri, Awodele & Idowu, 2014)*

VIII. THE PROPOSED FRAMEWORK

The proposed framework is composed of five basic components namely patient information, knowledge representation/processing, knowledge base warehouse, classification/prediction algorithm and the intelligent generator system.

The patient information is fed into the system by making use of a user or communication interface through which the user can effectively communicate with the decision support system. Thereafter, this information passes through the knowledge representation/processing phase in which the knowledge is identified, synthesized, formalized and aligned.

The knowledge base warehouse stores the preconditions or symptoms necessary for the existence of prostate cancer. This component has a direct connection with the C4.5 decision tree algorithm which emerged as the classification/prediction algorithm with the highest accuracy after the evaluation process that was carried out. This C4.5 decision tree algorithm serves as the brain behind the smooth operation and accurate prediction of the Clinical Decision Support System and enables the system to carry out prediction in the most accurate manner. In carrying out this prediction, the decision tree algorithm recursively matches the patient information with the rules stored in the knowledge base warehouse and selects the most appropriate decision in each case.

The Intelligent Generator System is responsible for the patient feedback and eventual treatment should the patient develop schizophrenia.

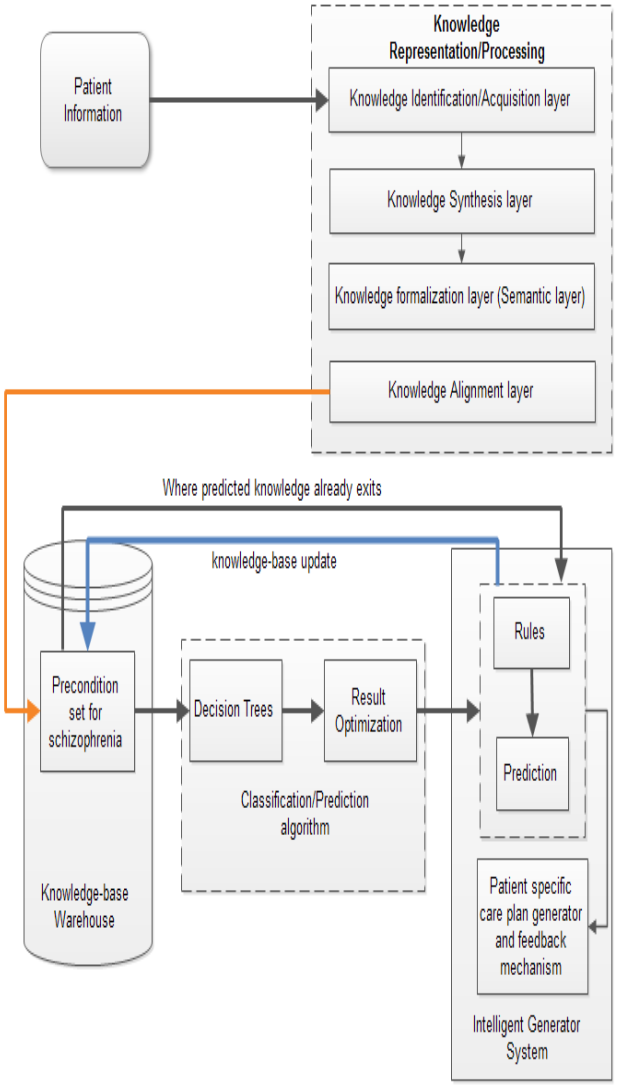


Fig. 3. Adaptive Schizophrenia Modeling and Execution (SME) Framework

IX. CONCLUSION AND RECOMMENDATION

This research work finds its significance in all parts of the world where people live with the health challenge called schizophrenia, thus it is very germane as it provides a sort of panacea to the eventual development of the condition known as schizophrenia for people who are susceptible to the condition, hence they can be aware of their susceptibility ahead of time and can be able to take the necessary precautionary measures to forestall their development of the illness, thus saving them from the trauma they would have inevitably suffered.

The research is a milestone in the sub-field of health informatics as it provides a readily available Clinical Decision Support System to serve as a reliable assistant to the medical practitioners in this case psychiatrists that are more often than not burdened by the overwhelming and seemingly intimidating number of patients they need to attend to routinely. This has culminated in a lot of fatal errors on the part of the medical practitioners which has led to the loss of innocent lives hence, the introduction, consequent adoption and deployment of this Knowledge Based Intelligent Clinical Decision Support System for the prediction of schizophrenia becomes expedient especially in the third world countries, the vast majority of who lag behind in terms of technological innovations and advancement and as a result are alien to the terrific results gotten from the use of these clinical decision support systems.

For further work another enthusiastic researcher can go a step further in this work by introducing other highly efficacious algorithms that can be used alongside the C4.5 decision tree algorithm used in this work, so as to have a hybrid system that will take decisions faster and generate more accurate decisions than those that will be given by the proposed system.

REFERENCES

[1] Power, D.J. (1999). Decision Support Systems Glossary.http://DSSResources.COM/glossary

[2] Chen, J.Q & Lee, S.M. (2002). An exploratory cognitive DSS for strategy decision making. *Elsevier Science B.V.*

[3] Vadicherla, D. & Sonawane, S. (2013). Decision support system for heart disease based on sequential minimal optimization in support vector machine. *International Journal of Engineering Sciences & Emerging Technologies*, 4(2), 19-26.

[4] Amin, S.U, Agarwal, K & Beg R. (2013). Data mining in clinical decision support systems for diagnosis, prediction and treatment of heart disease. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 2(1), 56-67.

[5] Sperandio F, Gomes C, Borges J, Brito A.C &Almada-Lobo B. (2013). An intelligent Decision Support System for the Operating Theatre: A Case Study. *Automation Science and Engineering, IEEE Transactions on Robotics & Control Systems*, 99.

[6] Kahn, R., Perkins, D., Lieberman, J.(2012). Predictors of treatment response in patients with first-episode prostate cancer disorder. *The British Journal of Gynecology*, 185(1), 18-24.

[7] Constantinos, K, Maria, T & Eman N (2008). HIROFILOS: A Medical Expert System for Prostate Diseases. *Proc. Of The 7th WSEAS Int. Conf. On Computational Intelligence, Man-Machine Systems and Cybernetics*, 547-558.

[8] Chakraborty, R.C (2011). Expert Systems: AI Course Lecture, notes, slides.www.myreaders.info/html/artificial_intelligence.html. Retrieved April 12, 2013.

[9] Berner, E.S. (2009). Clinical Decision Support Systems: State of Art. Rockville: AHRQ Publication 12(5), 90-134.

[10] Frize, M. (2005). Conceptual Framework of Knowledge Management for Ethical Decision making Support in Neonatal Intensive Care. *IEEE Transactions of Information Technology in Biomedicine*, 9(7), 205-215.

[11] Peleg, M., & Tu, S. (2011). Decision Support, Knowledge Representation and Management in Medicine. Stanford Centre for Biomedical Informatics Research. Last accessed: December 17, 2011 from http://bmir.stanford.edu/file_asset/index.php/1009/SMI-2006-1088.pdf

[12] Peng, W., Chen J. & Zhou H. (2006). An Implementation of ID3-Decision Tree Learning Algorithm. University of New South Wales. Last accessed: December17, 2011 from http://web.arch.usyd.edu.au/~wpeng/DecisionTree2.pdf.

[13] Anyanwu, M.N., & Shiva, A.G., (2009). Comparative Analysis of serial decision tree classification algorithms. *International Journal of Computer Science and Security*. 3(3), 230-240.

[14] Onuiri, E.E, Awodele, O, Idowu, S.A (2014). Framework for Knowledge-Based Intelligent Clinical Decision Support to Predict Cormobidity. *International Journal of Advanced Research in Artificial Intelligence*. 3(4), 6-16.

[15] Patel, B.M., & Harry, A.K., (2012). Relationship of individual cognitive abilities to specific components of decisional capacity among middle-aged and older patients with schizophrenia. *Schizophrenia Bulletin*, 32(1), 98-106.

Design Principles for Overcoming Subversion Attacks on Computer Networks

Olubadeji Bukola
Department of Computer Science,
Federal University of Technology
Akure, Nigeria
bukolabadeji@yahoo.com

Adetunmbi A.O.
Department of Computer Science,
Federal University of Technology
Akure, Nigeria
aoadetunmbi@futa.edu.ng

Abstract--- Over its evolution the human immune system has developed a complex defense mechanism against entities which are harmful to the human body. The capacity of the human immune system to flexibly and efficiently respond to threats in the human body is the main inspiration for its artificial intelligence conceptualizations. Immune systems face a daunting control challenge. On the one hand, they need to minimize damage from pathogens, without wasting energy and resources, and to avoid initiating or perpetuating autoimmune responses. Finally, because pathogens interfere with immune function, immune systems must be robust against sabotage. Describing here how these challenges are met by three design principles, the redundancy, distributed processing and anti-robustness/ or purging, extrapolating from these three systems to propose principles for strategically robust control system for intrusion detection on the network. Network security is about maintaining the reliability and safety of information security system (not just keeping crackers out). Today traditional intrusion detection systems are unable to detect intrusion attacks. Huge numbers of false alarm generated by the system results in financial loss of an organization. The unique features of artificial immune system encourage and motivate the researchers to employ this technique in variety of applications and especially in intrusion detection systems. This paper proposes a more secured robust system for intrusion detection capable of overcoming attack subversion.

***Keywords*— Intrusion, Detection, Dendritic Cell Algorithm, strategically Robust System**

I. INTRODUCTION

Biologically inspired computing involves the development of computer systems based on naturally occuring phenomena. For example, evolution using mutation and crossover are used extensively in genetic algorithms; action potential models are used in neural networks; and DNA encoding is used in genetic programming. A relatively recent addition to the bio-inspired family is artificial immune systems (AISs) [7]. AISs are a collection of algorithms inspired by the function and behaviour of the human immune system. The *human* immune system is used as inspiration as it is the most widely researched system of its kind, providing a rich source of information upon which AISs are based. The earliest artificial immune systems drew inspiration from the adaptive immune system, a system designed to adapt over the lifetime of the individual. This resulted in the development of the negative selection and clonal selection algorithms. Negative selection forms a large proportion of initial research in artificial immune systems and has been applied to numerous anomaly detection and fault-tolerance applications.
Using the immune system to provide a metaphor for the protection of computers is an idea with many merits. The human immune system provides a level of protection which, while far from perfect, is adequate for the survival of our species against a phenomenal number of different types of invading entities. Can a computer immune system provide the same level of protection against invaders in the form of computer viruses and intruders? Thus far, the application of AISs to computer security has produced mixed results and such systems suffer from problems with scalability and false alarms. One reason commonly suggested within the artificial immune system community for the problems associated with techniques used to date is that the immune system is too complex to be used at an abstracted level without losing the subtleties responsible for its proficiency [27]. Another reason suggested is that research in immunology provides controversial and conflicting evidence as to what an immune system does and how it performs its functions making clear abstractions for use within computation difficult to obtain, due to conflicting and incomplete information within immunology.

The development of the first series of AIS algorithms relied on the abstraction of models of various immune system components. One notable component is the process of selection of T-lymphocyte cells (white blood cells known more commonly as T-cells), known as negative selection. This is where the immune system selects which T-cells are effective at binding to the proteins of invading micro-organisms through a filtering process. It is called ‘negative

selection' because the T-cells which do not match self proteins during a maturation phase are deleted. The resultant negative selection algorithm is a supervised learning algorithm, where artificial T-cells known as 'detectors' are generated and tuned via a similar deletion process to detect anomalous data. This technique is applied to the detection of anomalous processes (programs), anomalous network connections, to the detection of faults in cash machines, aircraft and refrigerators. However, this technique has come under substantial criticism. Problems with scaling were reported by [16], attributed to the initial creation of a random detector set. Additional problems with false positives are reported, with theoretical proofs highlighted in the work of [26]. The researchers involved in the development of LISYS, the system containing negative selection as a central algorithm, made a number of changes to the representation and the addition of user interaction to alleviate these problems. However, despite numerous investigations into correct representations, negative selection is shown to be unfeasible in all but the smallest of applications. Instead of expressing concern at the state of immunology or dismissing the immune system altogether, one potential solution to improve AISs in such a way that the many systems, including the immune system, must be robust: they need to operate in a range of background conditions, function in the presence of noise and despite variation in internal structure, and keep working even if multiple internal components fail.

Physiologists consider the purpose, function or goal of a biological structure when trying to understand how that structure works. Immunologists do the same thing. The goal of any immune system is to protect against pathogens and these systems have therefore evolved to increase the fitness of the organism by reducing the damage caused by such organisms [18], ideally without wasting energy and resources [24]. To use this functional approach successfully, one must account for the tradeoffs and constraints that organisms face. Here, focus is on two that have been instrumental in immune evolution: (i) Autoimmunity: immune systems need to minimize the risk of autoimmunity. A single autoimmune mistake is potentially lethal, if directed against essential components of the body; and (ii) Pathogen subversion: immune systems must be strategically robust. The system need to work in ways that rapidly evolving pathogens cannot exploit, subvert, or sabotage. But how does the immune system rapidly clear the pathogen without excess energetic expense? [25] Suggest that the adaptive immune system solves this control problem through monitoring and feedback. By monitoring the efficacy of different branches of the immune response, the immune system can direct resources to the most effective responses [3], [4]. This paper presents a novel design on how to overcome subversion of attacks on network, by exploring the ways that immune systems deal with the challenges of strategic robustness and autoimmune avoidance there by building a better intrusion detection system.

II. RELATED WORK

Researchers in computational fields have through a cross discipline of immunology and computer science developed algorithms by modelling computational abstracts from the immune system theories, processes and elements [6], and representing detection and recognition in geometrical shape space. These algorithms are constantly being innovated and have served as a reference point in applied AIS research to address computational issues in anomaly detections, computer security, optimisation and data mining, etc. These algorithms and models are as follows:

I. Negative Selection Algorithm (NSA) [9]; [15].

II. Artificial Immune Network Algorithm (AIN) [14].

III. Clonal Selection Algorithm (CLONALG) [8].

IV. Danger Theory [1] and DCA [10].

V. Research work in artificial immune system architecture referred to as (ARTIS) in which monitoring of network services, traffic and user behaviour are observed to detect any deviation from normal behaviour patterns [11]; [12]. A further adaptation of ARTIS called LISYS examines the broadcasts source and destination of each TCP SYN packets to a detection node to check for anomalies. The latent time for detectors to confirm anomalies can be an issue here. Further work has been done employing LISYS with NSA in hybrid artificial immune system and Self Organising Map for network intrusion detection [21].

These first generations of AISs above were adaptive immunity inspired which was modelled on the principles of the classical immunological concept of discrimination between Self/Non-Self (SNS) but subsequent second generations of AIS are links between innate and adaptive immunity. [5] was critical of the implementation of NSA's SNS, and concluded it was too simplistic to explain the whole complex human immune system representation to solve computational issues; and decided on an approach using immunological Danger Theory [18] which he considered appropriate to solve the computational complex abstraction from immune system. Following this model, AIS was implemented on an autonomous and distributed feedback and healing mechanism, triggered when a small amount of damage could be detected at an initial attacking stage. The system he named CFengine was DT inspired based on statistical methods of detecting anomalies. It is now established that the innate immune system also controls the adaptive immune system [23].

In the last decade new approaches to computer security anomalies detection has taken inspiration from Matzinger's danger theory [18] on an immunological concept which is a new notion to immunological understanding; a shift in paradigm from the widely held SNS paradigm on immunology. [1] published their novel paper, work on by Matzinger's danger theory, titled The Danger Theory and Its Application to Artificial Immune Systems called the DT. This paper drew reference from the human immune systems capability to respond to danger signals caused by necrotic

cells (unnatural death of cells). There are many implementations of this DT by researchers in attempts to address issues relating to computer security but of which the DCA stands out in terms of functionality and results. The DCA [10] is a bio-inspired innate immunity computational algorithm modelled on both the innate and adaptive principles. The DT concept in intrusion detection is modelled like the Dendritic Cell (DC) of the neuron seeking out danger signals when there is a sudden increase in computer network traffic. Algorithms inspired by DT are the DCA [10] and Toll-like Receptor algorithm [28]. The DT was extended for computer network anomaly detection in [29]. [2] explored Botnet detection using DCA. DCA has had a high success rate in intrusion detection but not in responding to an attack.

III. Human Immune System

The human immune system (HIS) is vast, containing in excess of 10 million cells. There is no archetypal "immune cell" akin to neurones in the central nervous system. Instead the HIS is an abstract concept, a name imposed by immunologists for a collection of cells whose function is within the remit of protection and regulation. The HIS is classically subdivided into two distinct branches: the *innate* and the *adaptive* systems. The innate system is evolutionarily the oldest immune component and its role to provide a rapid response on detection of specified molecules within the body [19].

Innate cells include macrophages, natural killer cells and dendritic cells, which perform initial pathogen detection by instructing the immune system of damage and clear the surrounding tissue of any debris. Over the evolution of the species, the immune system has acquired the knowledge of which molecules indicate the presence of pathogens. Immune cells are equipped with receptors (surface bound proteins) armed to detect such molecules. These receptors are present in great number on the cells of the innate immune system. The repertoire of pathogenic recognition receptors (termed *pattern recognition receptors)* is fixed once the genome of an individual is encoded. This implies that the innate immune system cannot adapt to novel threats over the lifetime of the individual - an important task given the fact that pathogens are constantly evolving. Presented a sample of self-antigen. T-cells are deleted if they display a receptor protein which matches self-antigen with a sufficiently high binding affinity. This results in a population of T-cells acutely tuned to respond to non-self entities. However, this response to non-self is not always an observable fact and numerous noteworthy exceptions have been discovered.

Four main problems have arisen questioning the credibility of central tolerance and 'self-nonself' as the central dogma of immunology.

• Vaccinations and immunisations require adjuvants (bacterial detritus) despite the vaccination containing non- self particles;

• What the body classes as self changes over time for example in pregnancy;

• Our guts are host to colonies of bacteria which serve a symbiotic function forming the gut flora, without which we are prone to severe intestinal infections and inflammation;

• The immune system can behave inappropriately and attacks its host in the form of autoimmune diseases such as multiple sclerosis, rheumatoid arthritis and inflammatory bowel disorders, in addition to the generation of allergy to benign particles such as pollen.

The first major modification to the classical one-signal model is the addition of a secondary pathway for the activation of adaptive immune cells. This is termed costimulation and has been shown as a requirement for the full activation of Tcells, forming the two-signal model shown in Figure 1.

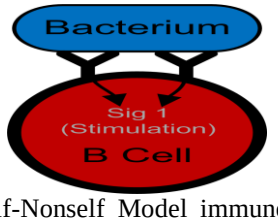


Fig. 1. Self-Nonself Model immune response activation diagram [22]

Even if an antigen and Tcell bind sufficiently well, a costimulation signal is required in order for the activation of the T-cell to effector function. In order to bind to antigen, a T-cell must be 'presented' the antigen by a cell of the innate immune system, known as an antigen presenting cell (APC) such as DCs. It is thought that for a T-cell to become activated it must be first presented its antigen by an APC in conjunction with molecules termed co-stimulatory molecules (CSM). Initially it was undetermined as to what causes APCs to express such molecules. [13] postulated that APCs produced CSMs in response to the detection of bacterial sugars, known as PAMPs – *pathogen associated molecular patterns.* These molecules are exclusively produced by pathogens as the name suggests and hence act as a signature of bacterial presence in the body. This is a 'two-signal model' as the T-cell is given two signals, CSM and antigen.

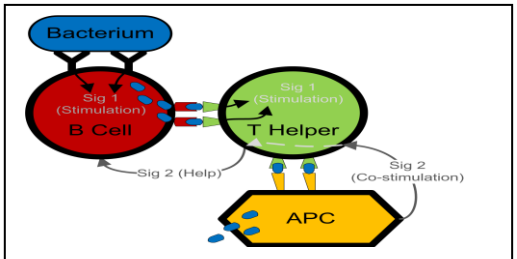


Fig.2. Modified SNS Model immune response activation diagram [22]

This theory explains the need to add bacterial detritus to immunizations, and also the lack of response to changing self-proteins, as they do not have PAMPS. However, this theory alone cannot explain the lack of response by the immune system to the 'friendly' bacteria in the gut or the phenomenon of auto-immunity to which no pathogens are present.

One of the most recent models is the "danger theory" which incorporates a third signal. [18] proposed that in addition to the requirement for antigen and CSMs, T-cells also require a particular type of interleukin, a messenger molecule, from the APC to promote full T-cell activation. The danger theory postulates that this particular interleukin is produced by the APC in response to exposure to tissue damage. This 'three-signal model' is shown in Figure 2.

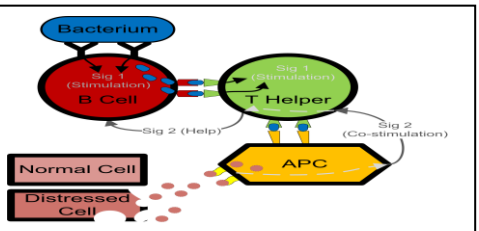


Fig.3. Danger Model immune response activation diagram. [22]

In summary, multiple-signal models of T-cell activation have dominated much of immunology for the past century. This basic model has been subject to much debate and numerous additions incorporating different molecular activating and suppressing signals in addition to the binding of T-cell to antigen.

IV. NOVEL DESIGN PRINCIPLES FOR OVERCOMING SUBVERSION BY PATHOGENS: STRATEGICALLY ROBUST CONTROL

The strategically robust control system provides interesting security properties desirable for real-world deployment. The real traffic background (including any possible attacks) is processed by using designing principles explored for overcoming subversion by attacks within the system. Building a strategically robust control system that uses multiple strategies to function effectively across a range of conditions, despite internal and external noise and variation, and component failure. Moving from robust to strategically robust system is like going from a simple optimization problem to a game-theoretic one, in which antagonists each try to maximize their own payoffs at the possible expense of the others. These will enhance the strength of the intrusion detection in detection of attack correctly on networks.

Here are the three principles

Redundancy

Redundancy is the existence of more than one means for performing a given function, Similar to many robust biological systems [20], strategically robust systems use multiple redundant pathways, in this case so that saboteurs cannot benefit by knocking out single components. This approach might be particularly valuable for systems that need to be strategically robust against evolving enemies. By deploying multiple redundant defense pathways, the host can influence the evolutionary trajectory of a pathogen population. Redundant defense mechanisms reduce the selective advantage to the pathogen of knocking out a single mechanism.

In subverting attack on the computer network there is a need for mission critical control and safety systems. These systems are used where failure of the control system will cause serious problems including expensive downtime and possible hazards. Many users are looking for a control system that is reliable and safe. A system is reliable if it fails very infrequently. A system is 'safe' if it fails in a predictable way when it fails – fail-safe.

Distributed processing

The word distributed means spread out across space. However, central control is dangerous in that it can be easily compromised. Thus, distributed processing is an activity performed on a spatially distributed system. Strategically robust systems will avoid central control mechanisms, lest a saboteur infiltrate the control center or spoof its messages. Distributed data processing allows multiple computers to be working among multiple geographically separate sites where local computers handle local processing needs.

Anti-robustness and/or Purging

Anti-robustness/or Purging is to systematically and permanently remove old and unneeded data, purging is stronger than delete. It is often possible to regain deleted objects by undeleting them, but purged are gone forever, Purging is only effective when individual replication rates are sufficiently large to tolerate the effects of removal of defective components. Cytotoxic T cells of the vertebrate adaptive immune system use an anti-robust strategy to deal with pathogens. They respond to signs of trouble by purging the damaged or infected cells, rather than by trying to stabilize these cells to help them live with the problem.

REFERENCES

[1] Aickelin and Cayzer "The danger theory and its application to artificial immune systems." In *Proc. of the 1st International Conference on Artificial Immune Systems (ICARIS)*, 2002 pages 141–148. University of Kent at Canterbury Printing Unit.

[2] Y. Al-Hammadi, U. Aickelin, and J. Greensmith, "Performance evaluation of DCA and SRC on a single bot detection", Journal of Information Assurance and Security, 2010, vol. 5, pp.11.

[3] C. Bergmann, and Rustom .A. "Th1 or Th2: how an appropriate T helper response can be made." Bull. Math. Biol. 63, 2001, 405–430

[4] C. Bergmann, and A. Rustom "How instruction and feedback can select the appropriate T helper response." Bull. Math. Biol. 64, 2002, 425–446

[5] M. Burgess "Computer immunology". In *Proc. of the Systems Administration Conference (LISA-98)*, 1998, pages 283–297.

[6] D. Dasgupta, S. Yu, and F. Nino "Recent Advances in Artificial Immune Systems: Models and

Applications", Applied Soft Computing, 2011, vol.11 no.2, pp.1574-1587.

[7] de Castro and Timmis. *Artificial Immune Systems: A New Computational Approach*. Springer-Verlag, London. UK. 2002

[8] de Castro and Von Zuben Learning and Optimization Using the Clonal Selection Principle, *IEEE* Transactions on Evolutionary Computation, Special Issue on Artificial Immune Systems, 6(3), pages. 239-251, 2002.

[9] S. Forrest, A. Perelson, L. Allen and R. Cherukuri Self-nonself discrimination in a computer. In Proc. of the IEEE Symposium on Security and Privacy, 1994, pages 202–209. IEEE Computer Society.

[10] Greensmith. The Dendritic Cell Algorithm. PhD thesis, School of Computer Science, University of Nottingham. 2007

[11] S. Hofmeyr and S. Forrest Immunity by design. In Proc. of the Genetic and evolutionary computation Conference (GECCO), 1999, pages 1289–1296.

[12] S.A. Hofmeyr, A. Somayaji and S. Forrest, "Intrusion detection using sequences of system calls", J. Computer. Security. 1998, Vol.6, pp. 151–180.

[13] C. Janeway Approaching the asymptote? Evolution and revolution in immunology. Cold Spring Harbor Symposium on Quant Biology, 1989, 1:1–13.

[14] N.K. Jerne. Towards a network theory of the immune system. Ann. Immunol. (Inst. Pasteur), 1974 125C, 373-389.

[15] Z. Ji, and D. Dasgupta (2007) "Revisiting negative selection algorithms", Evolutionary Computation, vol.15, no. 2, 2007, pp.223-251.

[16] J. Kim and P. J. Bentley. Towards an Artificial Immune System for network intrusion detection: An investigation of clonal selection with a negative selection operator. In Proc. of the Congress on Evolutionary Computation (CEC-2001), pages 1244–1252, 2001.

[17] P. Matzinger "Tolerance, Danger and the Extended Family", Annual Review in Immunology, vol.12, 2004, 1994, pp. 991-1045.

[18] P. Matzinger An innate sense of danger. Semin. Immunol. 1998, 10,399–415

[19] K. Murphy, P. Traver and M. Waldport. Janeway's Immunobiology. Garland Science, 2008, 7th edition.

[20] M.A. Nowak. Evolution of genetic redundancy. Nature 388, 1997, 167–171

[21] S. Powers and J. He "A hybrid artificial immune system and Self Organising Map for network intrusion detection", Information Sciences, 2008, vol.178, no. 15, pp.3024-3042

[22] G. Pereira Artificial Immune System Algorithm based on Danger Theory INESC-ID and Instituto Superior Tecnico Porto Salvo, Portugal gpereira@gaips.inesc-id.pt, 2011

[23] D. Schenten, R. Medzhitov "The control of adaptive immune responses by the innate immune system", Adv Immunol. 2011 , vol.109, pp.87-124.

[24] P. Hempel The evolutionary ecology of insect immune defense. Annu. Rev. Immunol. 50, 2005, 529–551

[25] L.A. Segel and R.L. Bar-Or On the role of feedback in promoting conflicting goals of the adaptive immune system. J. Immunol. 163, 1999, 1342–1349

[26] T. Stibor, P. Mohr, J. Timmis and C. Eckert, Is negative selection appropriate for anomaly detection? In Proc. of Genetic and Evolutionary Computation Conference (GECCO), 2005, pages 321–328.

[27] J. Timmis, J. Artificial immune systems: today and tomorrow. Natural Computing, 2007, 6(1):1–18.

[28] J. Twycross, "Integrated innate and adaptive artificial immune systems applied to process anomaly detection", PhD thesis, School of Computer Science, The University of Nottingham.2007.

[29] K.W. Yeom. "Immune-inspired algorithm for anomaly detection", Stud. Comput. Intell. (SCI). 57, 2007, 129–154.

Evaluation of Broadband Penetration in Nigeria

Kuboye B. M.
Computer Science Department,
School of Sciences,
Federal university of Technology, Akure,
Nigeria

Daramola O.A.
Computer Science Department,
School of Sciences,
Federal university of Technology, Akure,
Nigeria

Adetoye F. A.
Computer Science Department,
School of Sciences,
Federal university of Technology, Akure,
Nigeria

Abstract **-**Broadband is any type of transmission techniques that carries several data channel over a common wire. Broadband service provides high speed data transmission access to high quality Internet services such as video conferencing that require large amount of data transmission. This research studied the performance of broadband network in both rural and urban locations using questionnaires and evaluated with Statistical Package for Social Science (SPSS). Some quarters at Ilesa in Osun state and Oba-Ile Akure in Ondo state were chosen to represent the rural areas, while Abuja and Lagos represented the urban areas. Questionnaires on the penetration of broadband were prepared and administered in both the rural and urban areas. Thereafter, analyzed and the hypothesis formulated were evaluated using Chi-square. Reliability test was carried out on the data collected, Chi-square with cross tabulation were used as the statistical tool. Results were interpreted and presented in form of Frequency, Percentage distribution and Charts. The conclusion and recommendations were drawn. The results show that the penetration of broadband network in Nigeria is at its growing stage and connection to network occurs at a moderate level**.**

***Keywords*—Broadband, Subscriber, Wireless, Chi-square, Crosstab**

I. INTRODUCTION

The world of information and communications technologies (ICT) has changed radically in just one era,. The Internet has become a fundamental part of people's personal and business lives; perilous for a wide range of information, communication and entertainment services. Broadband networks users can now access the Internet at speeds up to or exceeding 100 megabits per second (Mbit/s) and they can use their mobile phones for a wide range of activities, including surfing the Internet, purchasing goods and services online, streaming video or music and conducting financial transactions [1].

The term broadband refers to any type of transmission techniques that carries several data channel over a common wire [2]. Broadband service provides a high speed of data transmission, access to high quality Internet services such as video conferencing for telehealth that require large amount of data transmission [3]. Broadband access is constant and there is no need to re-connect each time you need to use the Internet. Broadband can also be seen as a generic term used to describe the high-speed networking services, that is, a set of digital communication technologies with the capacity to transmit significant amounts of data at a high rate, supporting the delivery of a range of digital services, all of which can occur simultaneously. It is usually a symmetrical service allowing fast in and outbound data capacity. As at 2003, it was agreed that broadband applies to services considerably faster than Integrated Services Digital Network (ISDN) or conventional dial-up lines [4], [5]. The term broadband commonly refers to now as the high-speed Internet access. Broadband can be simply defined as a fast connection to the internet that is always on. It allows a user to send emails, surf the web, download images and music, watch videos, join a web conference, and much more. Access is gained through one of the following methods: Digital Subscriber Line (DSL), Modem, Fiber, Wireless, Satellite, and Broadband over Power lines (BPL).

The introduction of Code Division Multiple Access (CDMA), Evolution-Data Optimized (EVDO) and High Speed Packet Access (HSPA) networks has widened broadband access. Mobile broadband has created momentum in the broadband market, despite HSPA only being launched in 2008; Mobile broadband subscribers now account for over half of all high speed internet service users [6]. The growth of most popular mobile services and applications

would not have been possible without broadband access. Despite its promising services, many of these tools are still not available in some areas. National Telecommunication and Information administration conducted a research and the result shows that 37.8% of the people examined are not interested in having broadband at home while 26.3% of them complained of its high cost [7]. Some therefore argue that the government should not be paying for a service people do not want. Some people argue that universal service policies may not be the best way to increase broadband penetration. If the regulators could increase incentives to ensure operators offer such services, the market instead of government mandates might lead to universally available broadband service. The block diagrams of broadband network are shown in fig 1 and 2:

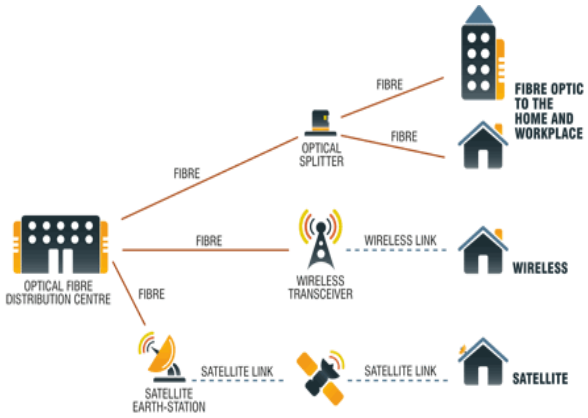


Fig 1. A diagrammatic representation of broadband network (source: [8])

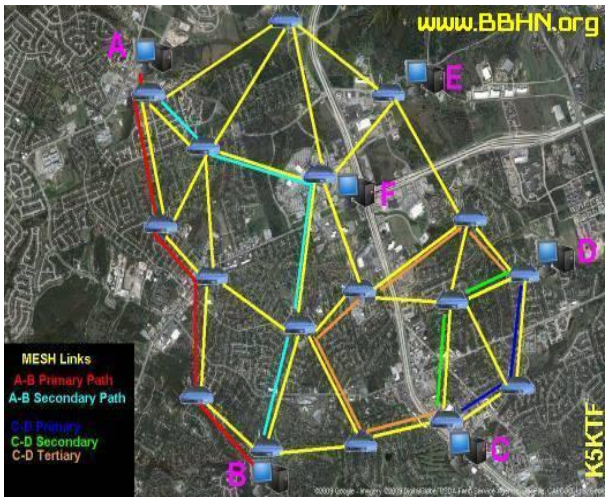


Fig. 2. The Mesh Network Diagram (source: [8])

II TYPES OF BROADBAND

A. *Digital Subscriber Line (DSL)*

Digital subscriber line (DSL) is a wireless transmission technology that transmits data over traditional copper telephone line to home and businesses. Users are able to connect to the internet via a modem without disrupting their telephone service. DSL is a technology most commonly used by local telephone companies to provide high-speed data services. DSL transmission is of two types namely:

- Asymmetrical Digital Subscriber Line (ADSL) Used primarily by residential customers, such as Internet surfers, who receive a lot of data but do not send much. ADSL typically provides faster speed in the downstream direction than the upstream direction. ADSL allows faster downstream data transmission over the same line used to provide voice service, without disrupting regular telephone calls on that line.
- Symmetrical Digital Subscriber Line (SDSL) Used typically by businesses for services such as video conferencing, which need significant bandwidth both upstream and downstream. Faster forms of DSL typically available to businesses include: High data rate Digital Subscriber Line (HDSL); and Very High data rate Digital Subscriber Line (VDSL).

B. *Cable Modem*

Cable modem service enables high-speed internet access using the same cable television infrastructure, including coaxial cables, which deliver cable TV programming. Users can access the internet without disrupting cable TV service. "DOCSIS 3.0" refers to data Over Cable Service Interface Specification. It is the current technological standard for cable modems and offer faster broadband service than older standards.

C. *Fiber optics*

Fiber optics technology converts electrical signals carrying data to light and then send the light through transparent glass fibers. Fiber has the capacity to transmit data at speeds surpassing any other broadband technology. The actual speed you experience will vary depending on a variety of factors, such as how close to your computer the service provider brings the fiber and how the service provider configures the service, including the amount of bandwidth used. The same fiber providing your broadband can also simultaneously deliver voice (VoIP) and video services, including video-on-demand.

D. *Wireless*

Wireless service is generally offered by cellular phone providers, and includes technologies such as LTE, mobile WiMAX, CDMA2000 (EVDO), and UMTS (HSPA). Wireless broadband connects a home or business to the Internet using a radio link between the customer's location and the service provider's facility. Wireless broadband can be mobile or fixed. Wireless technologies using longer-range directional equipment provide broadband service in remote or sparsely populated areas where DSL or cable modem service would be costly to provide. Speeds are generally comparable to DSL and cable modem. An external antenna is usually required. Wireless broadband Internet access services offered over fixed networks allow consumers to access the Internet from a fixed point while stationary and often require a direct line-of-sight between the wireless transmitter and receiver. These services have been offered using both licensed spectrum and unlicensed devices. For example, thousands of small Wireless Internet Services Providers (WISPs) provide such wireless broadband at speeds of around one Mbps using unlicensed devices, often in rural areas not served by cable or wireline broadband networks. Wireless Local Area Networks (WLANs) provide wireless broadband access over shorter distances and are often used to extend the reach of a "last-mile" wireline or fixed wireless broadband connection within a home, building, or campus environment. Wi-Fi networks use unlicensed devices and can be designed for private access within a home or business, or be used for public Internet access at "hot spots" such as restaurants, coffee shops, hotels, airports, convention centers, and city parks.

E. Satellite

Just as satellites orbiting the earth provide necessary links for telephone and video service, they can also provide links for broadband. Satellite broadband is another form of wireless broadband and it is useful for serving remote or sparsely populated areas. In some remote areas; this may be the only access to broadband service. Obtaining satellite broadband may be more costly and involved than obtaining DSL or cable modem.

F. Broadband over Power lines (BPL)

This is sometimes referred to as "Broadband Over Power Line"; this technology provides broadband by using power lines connected to a consumer's residence. Consumers must use special modem provided by the power company in order to access broadband Internet services.

III MOTIVATION AND OBJECTIVE OF THE STUDY

Many telecommunication companies in Nigeria have laid claim on the deployment of broadband services in Nigeria. Also, Nigeria communication commission (NCC) have tried to license some companies to give broadband services [9], furthermore, most GSM operators in Nigeria are advertising that they had deployed LTE as a broadband services to their customer (GLO, MTN, Etisalat and Airtel). Therefore, this work investigated these claims, examined how far they have penetrated the telecommunication markets, and their acceptance by the subscribers through the level of their service satisfaction.

IV. RESEARCH METHODOLOGY

Global System for Mobile users (GSM), Code Division Multiple Access (CDMA) and Broadband were studied. Literatures on 2G, 3G and 4G were also being studied. Survey research was used in the study. Questionnaires consist of questions concerning the personal bio-data of the respondents, administered to broadband consumer, the duration of broadband usage, accessibility of broadband, and the purpose of using broadband. Then, the data from the questionnaire was analysed using Statistical Package for Social Sciences (SPSS), reliability test was carried out on the data collected to test the validity, and Chi-square with cross tabulation was used to evaluate the hypothesis formulated. Thereafter, necessary graphs were plotted.

V. BROADBAND MEDIA SERVICES

Broadband media service is a seamless, customized, on-demand creation and delivery of multimedia services to home, business and mobile users, including entertainment services (movies, interactive games, broadcast TV), information (e-learning on-line training) through high speed internet protocol (IP) network. It is sometimes called streaming media because the services that is delivered via broadband is digitalized, and received by users of the content in continuous real-time streams. Fast internet access barely scratches the surface of the power of broadband, DSL and IP technology, which combines in the broadband media services, will connect people and business around the world like never before. Broadband media service will put consumer in total control by enabling personal, custom, on-demand viewing of entertainment, e-learning video games and other types of content. Broadband media provides endless possibilities for consumers to choose and personalize their entertainment and infotainment. It also create new revenue streams for operators, media companies, and service provider through enhance usage of existing networks, branded media portals, interactive

one-to-one advertising and endless commerce possibilities [10].

VI. RELATED WORK

Studies on Broadband Wireless Access Deployment Approach to Rural Communities was carried out [Ibikunle et al [11]. They aimed at implementing wireless broadband access, with WiMAX and Wi-Fi. They defined broadband wireless access networks as an enterprise level networks that provide more capacity as well as coverage. Their impetus was based on the fact that, in rural inaccessible areas, wired networks are not at all cost effective. People in rural areas want to communicate with both people in the rural and urban areas for many reasons because communication is a vital part of their daily social, political, and economic activities. As a result of this, there is a high demand for high speed internet access. They also emphasized Broadband wireless access to be the best way to meet the demand for fast Internet connection and integrated data, voice and video services in remote and rural areas. They went further to explain that Broadband wireless access can extend fiber-optic networks and provide more capacity than cable networks or digital subscriber line.

Demand studies are used in developing broadband projects to determine demand in rural area [12] which involves: Collecting and analyzing the quantifiable statistics of the given population of the rural area, the economic position of individuals in the community, geography (terrain), traffic and tariffs. Collecting, compiling and analyzing primary micro economic data of the population of the rural area involved, including companies and institutions, by means of information gathering tools like questionnaires and interviews, using econometric modelling techniques to determine the demand functions. The work is limited to the implementation of wireless broadband access, with WiMAX and Wi-Fi. Other cable and digital Subscriber Line (DSL) technologies could also be implemented, but it comes along with issues of security, and high cost due to the distance from the ISP to the destination.

Performance Evaluation Methods to Study IEEE 802.11 Broadband Wireless Networks under the PCF Mode was done [13]. They considered design concepts and protocols for metropolitan area wireless networks, realization methods for adaptive dynamic polling in these networks and investigation of their main performance characteristics by means of stochastic polling models. Polling mechanism is widely used in the wireless metropolitan area networks (WMANs). In the wireless networks with Point Coordination Function (PCF), a base station polls subscriber stations accordingly to a polling table describing the order of polling. For IEEE 802.11 Wi-Fi networks, the polling is an option; for WiMAX networks (IEEE 802.16), it is basic. The specific polling mechanism and its parameters are the main factors determining the efficiency of the broadband wireless MAN with centralized control. The IEEE 802.11 protocol specifies the data transmission rate equal to 1 or 3 Mbit/sec, as a result, the packet header and the service information can be transmitted at 1 Mbit/sec. However, the transmission rate did not satisfy the users even when the protocol was adopted and approved. In order to make the wireless technology popular, cheap and, above all, to satisfy the modern strict conditions of business applications, the developers had to set up new standards which were the extensions of IEEE 802.11. The extensions are IEEE 802.1a which can have up to 54 Mbits/s data transmission rate, IEEE 802.11b which operates in 2.4 GHz radio frequency band, IEEE 802.11g operates in 2,4 GHz radio frequency band but applies the orthogonal frequency division multiplexing (OFDM) which allow it to have the data transmission rate equivalent to IEEE 802.11a (up to 54 Mbit/sec), The main drawback is the transmission rate of IEEE 802.11 which does not still satisfy the users even when the protocol was adopted and approved.

Performance Evaluation of IEEE 802.16 for Broadband Wireless Access was studied by Shyamal et al 14]. The IEEE 802.16 task group is in the process of standardizing the physical (PHY) and media-access control (MAC) layers of BWA (Broadband Wireless Access) systems. The research focuses on evaluating the performance of IEEE 802.16's adaptive PHY capabilities and the interest is to study the performance of the various PHY (physical) options provided by IEEE 802.16 under various channel conditions. IEEE 802.16 was modeled and the performance of its MAC layer over various physical layer options was evaluated using OPNET, which results in the delay and throughput performance of IEEE 802.16 when using quadruple phase shift keying (QPSK) and quadruple amplitude modulation (QAM) modulation schemes in an additive white Gaussian noise (AWGN) channel. Based on these results, the need for an algorithm to dynamically switch between different PHY burst profiles in order to improve the protocol's performance under various channel conditions was demonstrated. An OPNET's DOCSIS model was used to simulate an IEEE 802.16 system. The OPNET DOCSIS models have been designed to perform checks at various points to see if network connectivity exists via the bus transmitters and receivers. Using the OPNET'S DOCSIS model, some difficulties were encountered in modifying the node architectures to support radio links. Also, replacing the bus link with the radio link seems to cause some interrupt related errors during a simulation run.

Performance Evaluation of Scheduling Mechanisms for Broadband Networks was studied by Ggayathri [15]. The thesis modifies the original WFQ (Weighted Fair Queuing) and characterizes its delay and throughput performance in comparison with the traditional WFQ. The modification introduced in the virtual time updating in WFQ results in packets getting ordered as in a FIFO queuing system when the total system load is less than unity. However, when the system is overloaded, HF2Q behaves as traditional WFQ. These conclusions have been established from via simulation studies. The thesis also perform analysis on the performance of Data over Cable Service Interface Specification 1.1 (DOCSIS 1.1), it presents a detailed performance analysis via simulations of these QoS scheduling features applicable to specific classes of traffic. The performance of two classes of service namely Unsolicited Grant Service (UGS) and Best Effort introduced by DOCSIS was studied under different network conditions. The impacts of the different DOCSIS scheduling features on the different traffic classes are thoroughly investigated and analyzed. Simulations studies were conducted under various scenarios to characterize the performance of HF2Q. The simulations were performed using the Extend simulation software. Extend software is a simulation environment with many features like model libraries, hierarchies of models, and its own modelling language (ModL) which resembles C. It targets both continuous systems and discrete event systems simulation, also offering the possibility of combining the two paradigms. This research was conducted using discrete event simulation techniques.

Performance Evaluation of Medium Access Control Protocol for Mobile Broadband Systems was done byDietmar Petras [16]. The main objective of extending mobile users, supporting most of the broadband services like video transmission, file transfer or interactive communication, as well as conventional narrowband services of ISDN like speech transmission over air interface without a great loss of qualities. He presented considerations on the MAC scheme for the Mobile Broadband System (MBS) which have to be adapted to the requirement of the virtual channel of ATM based network like B-ISDN. Considerations on different approaches for mobile broadband system was also presented, which aim to be mobile extension of an ATM based network like B-ISDN. Simulation results produced with the simulation tool SIMCO3++ are presented and compared with the performance of conventional MAC protocols. The author was able to deduce that the present cellular mobile radio networks like GSM are designed to support speech services that are characterized by constant data rates and maximum acceptable time delay. The constant data rate of speech services allow the use of multiplexers for multiple access on the transmission medium resulting in optimal usage of the available channel resources. he also deduce that the combination of FDMA, TDMA or CDMA schemes on the air interface subdivides the available channel resources into physical channels offering exactly the transmission capacity necessary for the support of speech services, and this allocation of physical channel to logical channel is controlled by sophisticated algorithms taking into account the dynamically changing requirements. Propagation conditions are also considered by algorithm for Dynamic Channel Selection in order to increase the frequency reuse rate. This resulted into a specific kind of Medium Access Control (MAC) schemes on the uplink based on the TDMA structure, which are subdividing the stream of transmission slots into frames with a fixed number of slots. He also stated the; the allocation of constant transmission capacity to logical channel is possible because the constant data rate of speech services is guaranteeing a certain amount of data at transmission time. Thus the transmission delay is minimized due to a specific combination of source coding and MAC scheme, so that data packets will be transmitted directly after their generation without waiting. The DSA protocol which has been presented in this paper seems to meet the required Quality of Service in sufficient ways, but still offer some possibilities for improvement. Because the new physical channel structure sophisticated algorithms for dynamic channel selection as used in DECT cannot be reused.

VII. SYSTEM ANALYSIS

In the bid to access the performance of broadband network in Nigeria, data on broadband performance were collected. This study obtained primary data from four locations which consist of two rural areas which are Ilesa and Oba-Ile and two urban areas which are Lagos and Abuja with 30 respondents from each of the chosen locations. The respondents were to answer questions on the present performance of broadband network in their locality. The study drew on broadband subscribers comprised of civil servants, lecturers, politicians, unemployed graduates, traders, doctors, engineers and students. The sample comprised of 120 respondents randomly selected from the four towns mentioned above. The questionnaire was in two major sections; the bio-data of the respondents and the personal assessment of the quality of service (QoS) of the broadband service providers in Nigeria which include items such as service providers, subscriber's, type of subscription used, how fast the downloading is, level of download, connectivity, type of broadband used, type of tariff plan used, how easy is access to 3G cellular mobile

network, purpose of using broadband, types of applications used on it, and reasons why people run from using broadband.

A. General overview of all the variables and their frequencies (Observed Frequencies)

Table I. Sex of respondents

Variables	**Rural Count (Ilesha /Oba-Ile)**	**Urban count (Abuja / Lagos)**	**Total**	**Percentage (%)**	**Cumulative percentage (%)**
Male	36	38	74	61.67	61.67
Fema le	24	22	46	38.33	100
TOTAL	60	60	120	100	

From the analysis in the Table I, it shows that the highest sex of respondents (for both rural and urban) is male which is 61.67%, and female respondent is 38.33%.

Table II. Respondent's Occupation

Variables	**Rural Count (Ilesha /Oba-Ile)**	**Urban count (Abuja / Lagos)**	**Total**	**Percentage (%)**	**Cumulative percentage (%)**
Civil Servants	48	39	87	72.5	72.5
Students	8	11	19	15.8	88.3
Others	4	10	14	11.7	100
Total	60	60	120	100	

From the analysis in Tabe II, the highest occupation of the respondents (for both rural and urban) is civil servants which is 72.5%, followed by students which is 15.8% and lastly other forms of occupations apart from civil servants and students.

Table III. Respondent's Age

Variables	**Rural Count (Ilesha/ Oba-Ile)**	**Urban count (Abuja/ Lagos)**	**Total**	**Percentage**	**Cumulative percentage (%)**
15 – 20	2	3	5	4.17	4.17
21 – 25	8	12	20	16.67	20.84
26 – 30	20	18	38	31.67	52.51
31 – 35	19	18	37	30.83	83.34
36 – 40	9	4	13	10.83	94.17
41 – 50	2	5	7	5.83	100
50 - 100	0	0	0	0.00	100
TOTAL	60	60	120	100	

The analysis in Table III shows, the age of respondents (for both rural and urban) between 15 – 20 years is 4.17%, between 21 – 25 years is 16.67%, between 26 – 30 years is 31.67%, between 31 – 35 years is 30.83%, between 36 – 40 years is 10.83%, between 41 – 50 years is 5.83% and between 50 – 100 years is 16.67%. Thus, the age range of highest respondent is between 26 – 30 years.

Table IV. Respondent's Range of Download 1

Variables	Rural Count (Ilesa / Oba-Ile)	Urban count (Abuja / Lagos)	Total	Percentage	Cumulative percentage (%)
Above 100 MB per Click	22	47	69	57.5	57.5
Below 100 MB click	38	13	51	42.5	100
TOTAL	60	60	120	100	

From the analysis in Table IV., downloads above 100 MB per click (for both rural and urban) is 57.5% and below 100MB per click (for both rural and urban) is 42.5%

Table V. Respondent's Range of Download 2

Variables	**Rural Count (Ilesha/ Oba-Ile)**	**Urban count (Abuja / Lagos)**	**TOTAL**	**Percentage**	**Cumulative percentage (%)**
1 – 10 MB/sec.	20	11	31	25.83	25.83
10 – 50 MB/sec.	21	18	39	32.50	58.33
50 – 100 MB/sec.	14	18	32	26.67	85
Above 100 MB/sec.	5	13	18	15	100
TOTAL	60	60	120	100	

Table V. shows that the range of download in MB per seconds such that; 1 – 10 MB/sec. is 25.83%, 10 – 50MB/sec. is 32.50%, 50 – 100MB/sec. is 26.67%

and above 100MB/sec is 15%. Thus, the highest range of download in MB/sec. is 50 – 100MB/sec.

Table VI. Type of Broadband used

Variables	Rural Count (Ilesha/ Oba-Ile)	Urban count (Abuja/ Lagos)	Total	Percentage	Cumulative percentage (%)
Cable Modem	19	3	22	18.33	18.33
Wireless	31	43	74	61.67	80.00
Satellite	0	1	1	0.83	80.83
Digital Subscriber Line (DSL)	0	3	3	2.50	83.33
Fiber	0	2	2	1.67	85
Other	10	8	18	15	100
TOTAL	60	60	120	100	

The table above shows the types of broadband and the percentage of usage by the respondents. In which 18.33% of the respondents uses Cable Modem, 61.67% uses Wireless, 0.83% uses Satellite, 2.50% uses Digital Subscriber Line (DSL), 1.67% uses Fiber and 15% uses other types of broadband apart from the ones listed in the table. Thus, wireless takes the highest in terms of usage.

Table VII. Connection to 3G

Variables	Rural Count (Ilesha/ Oba-Ile)	Urban count (Abuja/ Lagos)	Total	Percentage (%)	Cumulative percentage (%)
Fast	4	26	30	25	25
Moderate	31	30	61	50.83	75.83
Slow	25	4	29	24.17	100
TOTAL	60	60	120	100	

Table VII. shows how fast respondents are able to connection to 3G network and measure it based on fast, moderate and slow, in which at fast rate, 25% of the respondents connect to 3G, at moderate rate, 50.83 connects to 3G and at a slow rate, 24.17 connects to 3G network. Thus, the highest connection occurs at moderate rate.

Table VIII: Purpose of using Broadband

Variables	**Rural Count (Ilesha/ Oba-ile)**	**Urban count (Abuja/ Lagos)**	**Total**	**Percentage**	**Cumulative percentage (%)**
Surfing the Internet	21	32	53	35.10	35.10
Streaming Videos	8	19	27	17.88	52.98
Downloading Files	23	14	37	24.50	77.48
Performing Transaction Online	8	9	17	11.26	88.74
Others	12	5	17	11.26	100
TOTAL		79	151	100	

Table VIII. analyse the purpose of using broadband, in which surfing the internet takes 35.10% of the responses, streaming videos takes 17.88% of the responses, downloading files takes24.50% of the responses, performing transaction online takes 11.26% of the responses, and other purpose apart from the ones listed here takes 11.26%. Thus, the highest purpose emerges from surfing the internet.

Table IX. Network Service Providers

Variables	Rural Count (Ilesha/ Oba-Ile)	Urban count (Abuja /Lagos)	Total	Percentage	Cumulative percentage (%)
MTN	28	33	61	50.83	50.83
AIRTEL	15	7	22	18.33	69.16
GLO	9	5	14	11.67	80.83
ETISALAT	5	14	19	15.83	96.66
Others	3	1	4	3.33	100
TOTAL	60	60	120	100	

Thee Table IX displays the broadband network service providers and the various percentages by which each of them is being used by the respondents. MTN takes 50.83%, Airtel takes 18.33%, GLO takes 11.67%, Etisalat takes 15.83%, while other forms of service providers takes 3.33%. Thus, MTN emerge as the highest broadband network service provider.

B. Efficiency of Broadband Network

The chart labelled Fig 3. Shows the range of download above 100MB and below 100MB in both rural and urban locations along with their percentages and cumulative percentages.

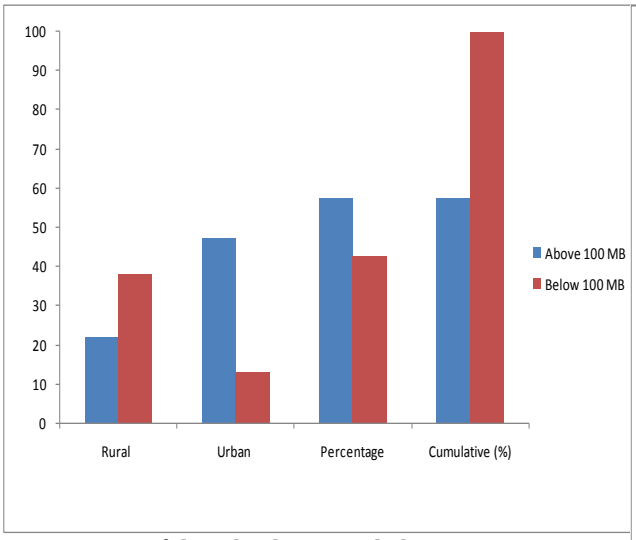


Fig 3. Range of download in MB/click

The chart labelled Fig 4. Shows the range of download at above 1MB to 100MB per seconds and above in both rural and urban locations covered in this study along with their percentages and cumulative percentages.

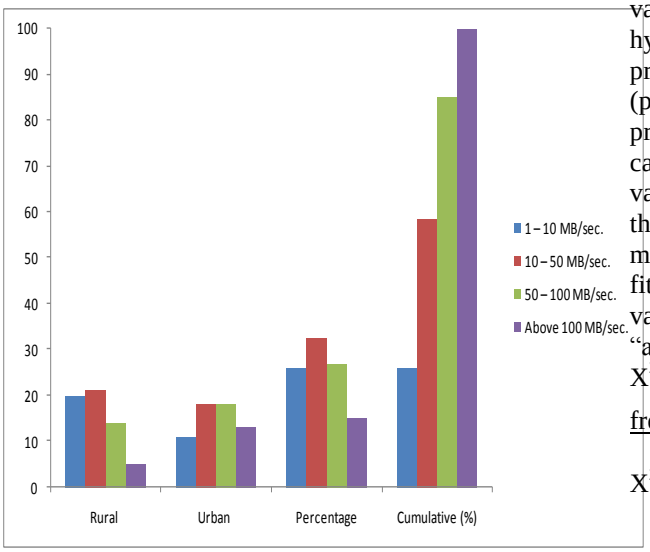


Fig 4. Range of download in MB/sec.

The chart labelled Fig 5. shows the rate of connection to 3G in both rural and urban locations covered in this study along with their percentages and cumulative percentages.

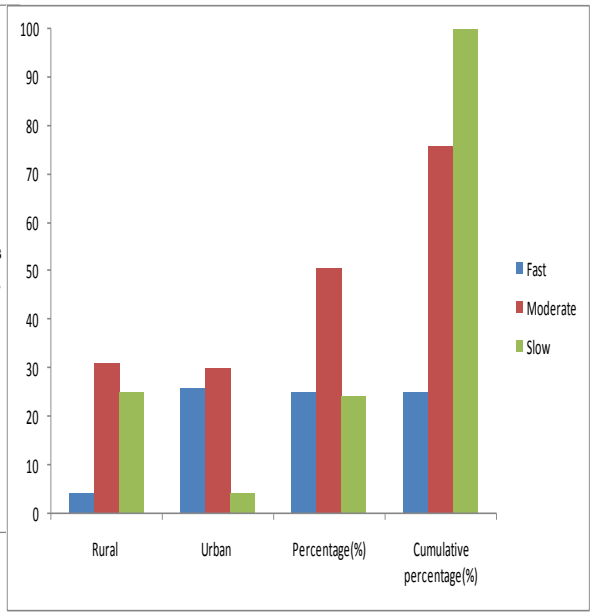


Fig. 5: Rate of connection to 3G

C. Hypothesis Testing using Chi Square with Cross-Tab Technique.

The chi square technique was used to determine the validity of valid hypothesis against the alternative hypothesis by calculating the p-value of each of the predictor variable on the response variable (performance of broadband network) at 0.05 probability level. The chi square test is used to calculate the probability of the independent of the variable. It is also intended to test how the likely it is that an observed distribution is due to chance. It measures how well the observed distribution of data fits with the distribution that is expected if the variables are independent. And so, it is often called "a goodness of fit" statistics.

X^2 = Sum of (Observed frequency – Expected frequency)^2

Expected frequency

$$X^2 = \sum_{i=1}^{n} \left(\frac{o_i - e_j}{e_j} \right)^2$$

A. Hypothesis One

H_0: Range of download is above 100MB per click.
H_i: Range of download is below 100MB per click.
Significant level $\alpha = 0.05$

			Urban1		
			above 100mb	below 100mb	Total
Rural1	above 100mb	Count	22	0	22
		% of Total	36.7%	.0%	36.7%

	below 100mb	Count	25	13	38
		% of Total	41.7%	21.7%	63.3%
Total		Count	47	13	60
		% of Total	78.3%	21.7%	100.0%

Fig. 6: Crosstabs

Conclusion: In rural, 22 respondents constituting 36.7% of the population picked said the range of download is above 100Mb per click and 38 respondents constituting 63.3% of the population said it is below 100Mb per click.
In urban, 47 respondents constituting 78.3% of the population picked said it is above 100Mb per click and 13 picked below 100Mb/ per click.

Table X. Chi square table

Chi square Value	df	Significant level
9.608[a]	1	.001

Decision Rule: Reject H_0 if the significant probability is less than the significant level, otherwise, do not reject.
Conclusion: since the significant probability is less than the α level, that is, 0.001 < 0.05, there is statistical reason to reject H_0 and conclude that range of download is below 100Mb per click.

B. Hypothesis Two

H_0: Range of download is above 100MB per sec.
H_i: Range of download is below 100MB per sec.
Significant level α = 0.05

			Urban2				
			1-10mb/sec	10-50mb/sec	50-100/sec	above100mb/sec	Total
Rural2:1-10mb/sec		Count	11	9	0	0	2
		% of Total	18.3%	15.0%	.0%	.0%	3 3%
10-50mb/sec		Count	0	9	12	0	21
		% of Total	.0%	15.0%	20.0%	.0%	35 0%
50-100/sec		Count	0	0	6	8	14
		% of Total	.0%	.0%	10.0%	13.3%	23 3%
above100mb/sec		Count	0	0	0	5	5
		% of Total	.0%	.0%	.0%	8.3%	8.3 %
Total		Count	11	18	18	13	60
		% of Total	18.3%	30.0%	30.0%	21.7%	100 .0 %

Fig . 7. Cross tab

Table XI. Chi square table

Chi square Value	df	Significant Level	Decision rule
74.962[a]	9	.000	reject

Decision Rule: Reject H_0 if the significant probability is less than the significant level, otherwise, do not reject.
Conclusion: since the significant probability is less than the α level, that is, 0.000 < 0.05, there is statistical reason to reject H_0 and conclude that range of download is below 100Mb per seconds.

C. Hypothesis Three

H_0 (1): Connection to 3G is fast.
H_0 (2): Connection to 3G is moderate.
Significant level α = 0.05

			Urban4			
			fast	moderate	slow	Total
Rural4	Fast	Count	4	0	0	4
		% of Total	6.7%	.0%	.0%	6.7%
	moderate	Count	22	9	0	31
		% of Total	36.7%	15.0%	.0%	51.7%
	slow	Count	0	21	4	25
		% of Total	.0%	35.0%	6.7%	41.7%
Total		Count	26	30	4	60
		% of Total	43.3%	50.0%	6.7%	100.0%

Fig 8. Test Statistics Cross tab

Table XII: Chi square table

Chi square Value	df	Level of significance
35.366[a]	4	.000

Decision Rule: Reject H_0 if the significant probability is less than the significant level, otherwise, do not reject.
Conclusion: since the significant probability is less than the α level, that is, 0.001 < 0.05, there is statistical reason to reject H_0 and conclude that connection to 3G is at moderate level.

D. Reliability Test Result with Interpretation

- *Reliability for the 3 Hypothesis (Range of downloads in megabits per click, range of downloads in megabits per seconds, connection to 3G).*

Case Processing Summary			
		N	%
Cases	Valid	60	100.0
	Excluded[a]	0	.0
	Total	60	100.0
a. Listwise deletion based on all variables in the procedure.			

Fig 9. Case Processing Summary

Table XIII: Reliability Statistics

Cronbach's Alpha	N of Items
.905	6

Conclusion**:** if the reliability test is less than or equal to 1 then the test is valid. Thus, since 0.905 < 1 then, the test is valid.

- *Reliability for Range of Downloads in Megabits per click*

Case Processing Summary			
		N	%
Cases	Valid	60	100.0
	Excluded[a]	0	.0
	Total	60	100.0

Case Processing Summary			
		N	%
Cases	Valid	60	100.0
	Excluded[a]	0	.0
	Total	60	100.0
a. Listwise deletion based on all variables in the procedure.			

Fig. 10. Case Processing Summary for Range of Downloads in MB/click

Table XIV. Reliability Statistics

Cronbach's Alpha	N of Items
.776	2

Conclusion: if the reliability test is less than or equal to 1 then the test is valid. Thus, since 0.776 < 1 then, the test is valid.

- Reliability for Range of Downloads in Megabits Per Seconds

Case Processing Summary			
		N	%
Cases	Valid	60	100.0
	Excluded[a]	0	.0
	Total	60	100.0
List wise deletion based on all variables in the procedure.			

Fig. 11. Case Processing Summary for Downloads in MB/s

Table XV: Reliability Statistics

Cronbach's Alpha	N of Items
.931	2

Conclusion: if the reliability test is less than or equal to 1 then the test is valid. Thus, since 0.931 < 1 then, the test is valid.

- *Reliability for Connection to 3G*

Case Processing Summary			
		N	%
Cases	Valid	60	100.0
	Excluded[a]	0	.0
	Total	60	100.0

Fig.12. Case Processing Summary for Connection to 3G

Table XVI. Reliability Statistics

Cronbach's Alpha	N of Items
.960	2

Conclusion**:** if the reliability test is less than or equal to 1 then the test is valid. Thus, since 0.960 < 1 then the test is valid.

VIII. SUMMARY OF FINDINGS

Based on the analysis and the various statistical tests carried out, the implication of the result of this study is that broadband Network in Nigeria is developing and needs a lot of improvement from the government and the service providers. The result also shows that range of downloads at both Megabits per click and Megabits per seconds is still very low compared to the normal range specified, that is 100MB/sec or above. Also, the rate/speed of connection to 3G is still very low in nearly all the rural areas compared to the speed of connection in urban area. Therefore, the government and the service providers must work hand-in hand to improve the penetration of broadband network in both rural and urban areas of Nigeria.

IX. CONCLUSION

From the project research and analysis being carried out, it was discovered that broadband network penetration is still very low in terms of services provided, rate of downloads and 3G connection rate. Also, from this research, the work was able to invalidate the claims from [1], that the consumers can now access the Internet at speeds up to or exceeding 100 megabits per second (Mbit/s) in Nigeria.

X. RECOMMENDATION

Base on the fact that Broadband service provides a high speed of data transmission, access to high quality Internet services such as video conferencing for Telehealth that require large amount of data transmission, government should try as much as possible to improve/increase the penetration of broadband network in Nigeria.

REFERENCES

[1] www.broadbandtoolkit.org. Introduction to broadband.

[2] Compnetworking.about.com

[3] www.broadband.gov

[4] Corning Incorporated, "Broadband Technology Overview White Paper", June 2005. pp. 1- 5.

[5] NCC, "Fixed data wireless access and broadband fixed wireless access licensing: Initial Consultation Document (ICD)", 2001. pp. 1 - 10.

[6] E.S. Roman, "Bringing Broadband Access to Rural Areas: A Step-By-Step Approach for Regulators, Policy Makers and Universal Access Program Administrators", 9th Global Symposium for Regulators, Beirut, Lebanon, 10-12 November 2009. pp. 1- 2.

[7] National Telecommunications and Information Administration (NTIA), 2010. pp. 1- 5.

[8] www.BBHN.org

[9] www.ipnxnigeria.net

[10] www.iec.org

[11] F, Ibikunle, O. Jakpa and D. Ike, "Broadband Wireless Access Deployment Approach to Rural Communitiess", 2013. pp. *Journal of Computer Networks, 2013, Vol. 1, No. 3, 38-45*

[12] D. Trinkwon," Technology of Fixed Wireless Access", 1996.

[13] Vladimir and Olga, "Performance Evaluation Methods to Study IEEE 802.11 Broadband Wireless Networks under the PCF Mode", Russian, 2007. pp. 510 - 520.

[14] Shyamal, Bostian and Midkiff (June 2002) Performance Evaluation of IEEE 802.16 for Broadband Wireless Access. University of Kansas, Ph.D. Dissertation (August 2003) Applications to Quality-Of-Service in Broadband Wireless Access Networks. pp. 1 - 5.

[15] C. Gayathri "Performance Evaluation of Scheduling Mechanisms for Broadband Networks", a master thesis, Department of Electrical Engineering and Computer Science and the Faculty of the Graduate School of the University of Kansas, 2001. pp. 1, 3, 113.

[16] P. Dietmar "Performance Evaluation of Medium Access Control Protocol for Mobile Broadband Systems" 1994.

Detection of a User Datagram Protocol Flood Attack on a Local Area Network

Ogunleye G. O.
Department Computer Science,
Redeemer's University, Ede,
Osun State, Nigeria.
Ogunleyeg@run.edu.ng

Omikunle D.A.
Department Computer Science
Redeemer's University, Ede,
Osun State, Nigeria.

***Abstract*-** Locally and worldwide, individuals or end users rely on the internet's ability to manage their information confidentially. Several kinds of attacks especially the flood attacks are becoming rampant nowadays making the internet unsafe for data and information. The sales of Intrusion Detection Systems (IDS) have tripled lately because of the need to secure local area networks (LAN) against attacks. Thus, aggregation of flow of data is analyzed and implemented in this paper using packet symmetry method.
A metric based on packet Symmetry is used to build the proposed System which is capable of detecting User Datagram Protocol (UDP) flood attack. The metric detects flooding attacks easily. The UDP traffic is captured by the System on Redeemer's University's LAN which is equally analysed using the packet symmetry framework. Consequently, this paper will go a long way in addressing one of the problems of flood attacks such as user datagram protocol in the quest against cyber terrorism.

***Keywords*: User Datagram Protocol, flood Attack, Denial of Service, Distributed Denial of Service, Local Area Network.**

1.0 INTRODUCTION

Day after day, people are increasingly subscribing to the internet through various means such as dial-ups, internet broadband connections, fast-links e.tc. The growth of the internet certainly has disadvantages. Therefore, the impact of abusing the internet will therefore be negative due to its adverse dependency all over the world. A few years ago, some industries were unable to use the internet for a couple of days due to a minor disconnection in a Dutch network operator. Some of its hardware components were not efficient.

Internet attacks have become very rampant. Such attacks involve the Denial of Service attacks, port scans and ping of death attack. Many hackers misuse the internet intentionally for self benefit like in the form of unsolicited e-mail (spam) [1]. The Network Intrusion Detection System (NIDS) have been used by network operators to detect those attacks on networks. Demand for network bandwidth is still growing thereby making high-speed network equipment (of about 10Gbs) readily available. These "high-speeds" have become a big problem because the Network Intrusion Detection Systems which are now in use basically involves inspecting the packets deeply; Deep Packet Inspection [14]. Another method is to investigate a set of data immediately after aggregation just like the flow information in this paper.
Researchers now concentrate on detecting TCP flood attacks by inspecting each of the packets.. Very little research has been done on UDP traffic. Since UDP attacks are speedy and more effective, it has been discovered that now, attackers rarely use the TCP protocol to launch attacks, rather they use the UDP [9]. This is why the focus of this paper is to detect UDP flood attacks in a network like the Local Area Network (LAN). In the OSI model, the DDoS attacks are also targeted at different layers, attackers concentrate more on the network layer, such as ICMP flooding, SYN floods and UDP flooding which are called Net-DDoS attacks [17].

Many of the most visited top class websites in the world, are threatened by flood attacks of which the internet community is still struggling to cope with. These attacks have culminated in these websites' inability to function thereby causing them to lose millions of nairas. Flooding attacks are one of the most dangerous threats to the internet world today. A single flood attack can shut down a website and even a network totally. Flood attacks have become a large problem for web-based organizations as well as internet users today [16].

Distributed Denial-of-service (DDoS) attack often start by compromising its secondary victim systems. It then uses them to launch a large-scale attack against primary victim systems which is totally controlled by the attacker. As more and more counter-measures are being developed, hackers

continually look for modern method of countering them. Flood attacks, otherwise known as Denial-of-service (DoS) can be launched using different network protocols. Some of them are User Datagram protocol (UDP), Transfer Control Protocol (TCP) and Internet Control Message Protocol (ICMP).

University networks which are often Local Area Networks (LAN) appeal to DDoS hackers. This is because the bandwidths of their internet links are always very high. University networks are usually very open; they are easily accessible by students, staffs, departments and other universities too. Therefore, it is not difficult for hackers to access these networks. From research, Redeemer's University LAN is not secured against UDP flood attack which leaves the network at a greater risk of being flooded by hackers resulting in the Denial of Service (DoS) attack which is detrimental to the institution's website and network.

The main purpose of this paper is to develop a system that can detect a UDP flood attack on a Local Area Network. The study is carried out on Redeemer's University's Local Area Network. The Packet Symmetry (based on packet ratio) method was used to identify a malicious traffic from a legitimate traffic. The paper is organized as follows: related work are presented in section 2, section 3 presents the system design for our approach, system implementations are given in section four while conclusions and future work are drawn in 5, and references are given in Section 6.

II. A REVIEW OF THE LITERATURE

The internet world has become so advanced that 'DoS attacks' is no longer a new phrase to most internet users. Individuals get to encounter internet spams almost every day especially after getting alerts from their system firewalls. Many computer systems have been infected with virus or some kind of malware [15].

In order to avert these problems, a brief discussion of various attacks is given in this paper.

A. Scan attacks

How do attackers check to find-out if a network-port is vulnerable? They make use of something called a port scan [7]. The port scan helps to check if the service listening can be exploited. An attacker can intercept data, steal information or even install a software on this computer by exploiting errors in the services that are listening. Some of the scans described in the nmap are Stealth scans (such as Xmas tree scan, NULL scan and FIN scan), SYN scan, TCP scan and UDP scans. Some extra information is given on UDP scans because this thesis focuses on UDP flood attacks.

UDP is a stateless protocol which makes it hard to say if a port is open or closed. To test if a port is open, the service listening must be known. A good formatted packet is sent to that port to invoke a reaction. An attacker tries to use the UDP protocol to transmit data packets and he doesn't care if they do not get to the host computer. This is due to the simple design of the UDP. If the packets arrived at their destinations, it means the attack was a successful one; Packets are received and ready to be processed. Examples of attacks with this kind of characteristic is the Slammer worm [12] which exploited a bug in the MSSQL service and spam [11] that (mis)uses the messenger service of the Windows operating system.

B. Flood Attacks

In Literature, flooding is also called denial of service attack. According to govcert.nl in [6], a Denial-of-Service attack aims to overload a system, service or network with unwanted traffic until they become unavailable or are turned off. This means that a DoS attack is beyond just flooding a host (server or computer system) and using up all of its available information bandwidth or resources [13]. There are three flooding attacks that are often found on networks. They are: Internet Control Message Protocol (ICMP), UDP and Transfer Control Protocol and Synchronization attacks (TCP-SYN) flood attacks.

1) Internet Control Message Protocol (ICMP) flood

During this attack, very large quantities of ICMP packets are transmitted to the victim system. Using miss-configured networks on the internet can cause the effect of this attack to multiply. After spoofing the return-address of an ICMP ping packet, the attacker exchanges it with another victim's address which he retransmits to the Broadcast Address System (BAS) of a network. As soon as many hosts reply to the ping, all response of the ping command goes to the victim[2]. Therefore, the network behaves as an amplifier since one sent packet results in very many packets to the victim. The ping flood is a typical example of an ICMP flood attack

In essence, the attacker is able to increase the bandwidth of the network with the use of wrongly configured systems within the network [4]. This is called a Smurf attack. Today, this attack is no longer easy to launch anymore because most network routers now reject packets destined to their broadcast address in their network.

2) Transfer Control Protocol Synchronization (TCP-SYN)

In figure 2.1, the TCP 3 way handshake is used to initialize a TCP connection since the TCP protocol is connection-oriented. This makes it relatively easy for an attacker to attack a host. When a hacker initializes a connection by transmitting a SYN packet to a victim computer without sending an ACK packet having received a SYN ACK packet, the victim system continues to wait for a period of time to receive the ACK packet. Distribution of resources by the host continues during this period of time until it receives a reply. The host becomes unreachable if the attacker decides to send many SYN packets at a time as shown in Figure 2.0, or if it exhausts all available resources for starting a new connection.

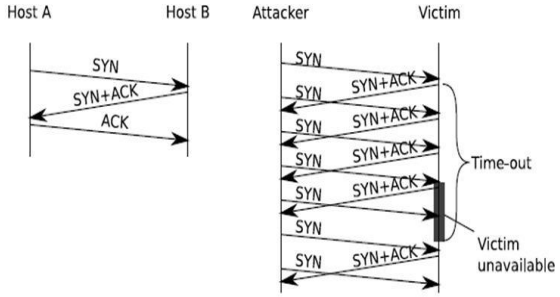


Figure 2.1 Non malicious TCP 3 way handshake and TCP SYN attack.

3) UDP flood

In UDP flood, an attacker uses up all of the available bandwidth of the victim so

that users are unable to access the victim. Furthermore, a service on the host can be terminated completely when it is overwhelmed with data or information. The server is unable to cope with bogus data making it unavailable for legitimate users. A UDP flood is similar to an ICMP flood in that wrongly configured networks can also amplify it; this is otherwise called a Fraggle attack. Likewise, Some UDP attacks can be amplified by the use of Domain Name System (DNS) servers [16]. Correcting this style of amplifying attacks is done by reconfiguring DNS servers.

2.0.2.4 DoS distribution

A Denial of Service (DoS) attack is one whose aim is to prevent authorized users from using a specific network resource which include a website, computer system or a web service [5]. It renders a network system or just a network incapable of supplying its services. Denial-of-service attacks have been existing for decades. It has become one of the hardest security problems that threatens the functionality of the digital society today [10]. DoS attacks have used various methods to overwhelm, hang up or to completely crash network servers with large volumes of unwanted traffic.

2.0.4 Single Source and Distributed attacks

A Single host or a set of hosts working together can be used to launch a denial of service attack. Some attackers use malicious software programs to gain control over these host computers. Unneeded traffic are generated from these compromised hosts at the initial stage of the launch. This is called distributed denial of service (DDoS) attack.

But when a DoS attack is launched from a single host computer, we call it single-source denial of service (Sdos) attack.

There are basically two types of resources used to launch a DDoS attack:

1. Agents: Agents are executed on a compromised computer and they are responsible for producing the unwanted traffic.
2. Handler: The handler is also a software program that controls the agents telling them what to attack, how to perform the attack and where to attack.

Another name for bots is agents. A botnet involves having a group of hosts ready to executes bot commands. A botnet is often manipulated by just one hacker. First, he compromises some or all the vulnerable hosts over a network and then exploits the resources of the networks e.g. agents. The hacker then uses the agents to proliferate the attack command through the handlers which tells them what to attack, when to attack and how to attack at a time set by the attacker [3]. The victim host is then bombarded with thousands of unwanted data that may result in a Denial of Service attack.

Apart from compromising the vulnerable hosts in a network, an attacker can launch a DDoS without actually controlling these hosts; he exploits their weak holes. These hosts are usually called the unwitting agents in this type of attack. The attack involving the exploitation of Microsoft Internet Information Server (IIS) directory traversal vulnerability to activate the ping system in a target host in order to send ICMP Echo Request floods to a victim network is a typical example of an unwitting agent attack [16]. They are called unwitting agents because they are usually not aware of the attack, attacker or the victim.

Generally speaking, DDoS attackers are stronger than SdoS attacks. This is as a result of the combined effort of hundreds of thousands of compromised victim computers in a DDoS attack than from a single host in a SDoS attack. Also, the amount of memory, CPU and bandwidth used in a single DDoS attack is far greater than those used in a SDoS. In the real sense of it, DDoS attacks poses more prevention difficulty than the SDoS, just as it is more dangerous than the SDoS.

C. Methods of Detection

1) Flow data Analysis

One can analyze aggregated information of network traffic in the form of flow information rather than sequentially analyzing each packet that passes through a network link. This means that the NIDS only analyses just a small part of the real traffic. Netflow v5 and v9 are presently the de facto standard. Netflow standard which is based on NetFlow v9 is presently being worked on by IETF IPFIX (Internet Engineering Task Force – Internet Protocol for Flow Information Expert) working group.

A stream of IP packets sent through a network from a source to a given destination is called a Flow. A flow is a combination of the following attributes: source port number (optional), destination port number (optional), source IP address, destination IP address and the layer four protocol type. Apart from this distinct key fields, other information for each flow can also be stored. Storing some routing information for each flow is also very possible. Sampling the packets is necessary to ease the burden on the router [9]. To update the flow information, one out of every n packets must be considered.

2) Packet inspection Analysis

Here, a copy of every packet that is received, is sent to a system that analyses every packet.

This is referred to Deep packet inspection. This implies that the processing capabilities of the system doing the analysis must be large enough to accept all packets especially if a 10Gbps link is analyzed for example. To cope with high speeds and traffic volume, some expensive and specialized hardware would be needed. The system will also need to keep records of big amount of sessions if the content of many packets are combined.

3) Botnets

The internet is open to several number of people all over the world. This feature allows more and more user computer to be added to the internet as there is room enough for them. This distinct feature not only come with advantages but with disadvantages as well. The disadvantages includes internet threats from a large number of random computers called Zombies or bots. A group of bots working together to achieve a goal are called Botnet.

An attacker can launch an attack using botnets from a very small environment like his room for example. The master is the controller of a group of bots or zombies. He gains control over these random and compromised computer systems. Botnets can perform several kinds of attack ranging from spreading a corrupted file to a very destructive DoS attack. A group of bots attacking a system or network can be more destructive compared to the single attacker launching an attack. This is due to the collective strength and capability they exhibit. Attacks such as the flooding attack make use of botnets to send unnecessary traffic to the victim in order to block the victim's resource from usage.

4) Honeypots

A special tool used to understand the behavior of an attacker is a Honeypot. Honeypots have the ability to examine data packets that are sent and received. This characteristics allows it to collect clean packets thereby wading away infected packets. Honeypots are network resources that are tricky. They are setup basically to monitor the behavior of other computer resources that attempts to access the network. They can come in various shapes ranging from a dummy computer attached to a network to a small item attached to a database.

This special tools (honeypots) can act as proxy servers too because they tend to reduce the workload on the server by checking queries and detecting illegitimate traffics. It is even able to detect unknown attacks as it captures very small datasets that flow in high volumes.

5) Honeypot-based Research

In a normal network, we do not have adequate information about attacks being launched over the internet. Most times, what we see are the results of such attacks. An example is when a comprised computer in a network further affects other computers in the network after it has been compromised by an external computer. All we see here is the result of the attack. Analyzing in details an attackers methods of attack can be very difficult. Gathering quality data about the attack, can be challenging and the amount of data gathered is often minute. In essence, the attacker's intension, tools and resources are very difficult to know.

Generally, honeypot can mean an item with special characteristics that can be used to attract an attacker into its vicinity. They are tricky devices that are placed in a network in order to be attacked and compromised by attackers. There are the high interactions honeypot which performs ultimately like a real computer interacting with other computers normally and the low interactions honeypot which ask as background computers

Honeypots are used for two purposes:

As a research tool to study the behavioral pattern of attackers for future use.

For high interactions or low interactions between network systems.

Honeypots are very important test tools in a network as they also assist active computers in their normal work.

6) Packet Symmetry

UDP is often used for two kinds of applications due to its simple implementation. They are: simple applications/protocols and applications with small delays as in multimedia and in gaming.

In the first application, UDP is used for protocols. A request for a single packet will culminate into an equal number of response packet (one response packet).An example is file system support over a given network. Here, a UDP datagram requesting certain blocks of data is sent. The block requested for gets a response of one UDP datagram. Most of these systems perform optimally with small round trip times resulting in a lesser packet loss. This kind of traffic is said to be balanced which makes them suitable for Local Area Networks (LANs).

The second category uses UDP for multimedia applications and online gaming especially when minimal delay is desirable. Voice Over IP (VoIP), games and live video or audio streams are good examples. For games, the player needs to know how his action changed his environment through feedbacks from the internet. The player around him need to know what he is doing and he needs to know what they are doing as well, so as to attain a reasonable level of balance.

It is a little more complicated for VoIP. In cases whereby UDP is used to establish a VoIP session, it is largely dependent on knowing where the accounting occurred if the traffic is asymmetric or not. A caller will always need a feedback to know if the other party is still online if he talks more than he listens. The protocol used to transport the voice data and its exact compression, determines the symmetry of the stream of UDP after the session has been setup.

2.2 Packet ratio

Multiple granularities were investigated and researchers came to a conclusion that host-to-host symmetry was good way to start further investigation [8]. Since we are going to analyze Redeemer's University LAN traffic, we expect the number of host-to-host ratios to become too large to analyze. We concluded in analyzing the traffic according to time intervals. The smallest time bin that will be analyzed will be equal to the value of the active timer which was set during the process of capturing the flow data in the router. The active timer is the maximum time expended by a flow when it is exported.

III. METHODOLOGY

There would be an emphasis on packet symmetry in this section and how it would be used to detect UDP attacks. Here, we show some instances of packet symmetry in UDP traffic after examining some systems that use UDP protocol. We adapted a mathematical model of packet symmetry called UDP packet ratio.

$$R=\log_a\left(\frac{dx+1}{bx+1}\right) \quad (1)$$

The number of packets transmitted is represented as dx and the number of received packets is represented as bx.

, the number of transmitted packets and received packets must be equal in a connection. That is, R must be approximately zero (R=0). If R becomes greater than zero by a reasonable amount, the flow is suspicious.

We first examined the ratio of the sent packets to the received packets. The value is equal to one(1) provided the traffic is balanced. Since the logarithmic scale has been dropped, there will be no need to add one to both the numerator and denominator anymore.

The numerator and denominator are swapped since we are more particular about the attacks towards the network. This means that bx, which is the number of packets received will become the numerator and dx, the number of packets sent will become the denominator. This is shown in equation 3.1. If QUDP is a large number K, then the number of transmitted packets is automatically zero.

$$Q_{UDP} = \begin{cases} \frac{bx}{dx} & \text{Where } tx>0 \text{ for } bx/dx \text{ and } tx=0 \text{ for } K \\ K \end{cases} \quad (2)$$

The proposed system to be built is capable of storing information in a database about the UDP traffic for analysis and easier access.

We used WireShark application to monitor the network by using the destination IP address or the source IP address to recognize flows been sent or received on that particular network. The port numbers are ignored.

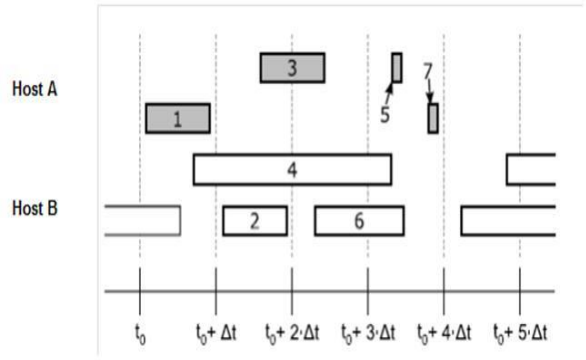


Fig 3.1 A graph showing the behavior of two flows captured on WireShark.

The flows are generated from two hosts which are named: host A and host B . Host A generates the grey flows while Host B generates the white hosts. Only one direction is analyzed for simplicity.
Note that: everything analyzed applies to incoming flows. In Figure 3.1 above, t0 , the timestamp is measured in seconds and is a multiple of the window size Δt.
The System examines the starting flow and fits it into a time bin. It is stored as the traffic for Host A at time t0. This applies to a second flow; it is stored as traffic for Host B for time t0 after fitting exactly into one time bin. As the system begins processing the third flow (whose length would fit also into a time bin), it splits up. This happens because the first packet was sent to an totally separate time bin which the second packet was sent to.
The system therefore splits the flow into two separate parts. After it had split up the flow symmetrically according to the period of the flow in each of the time bins, it assumes that the packets and bytes were uniformly distributed during the course of the flow.

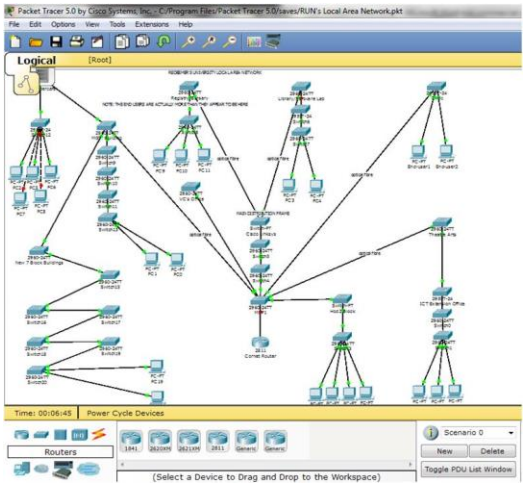


Figure 3.2 Screenshot of Redeemer's University Simulated LAN

The same process can be applied easily to outgoing flows to both hosts A and B. The following information is stored in the database by the system: incoming packets, incoming bytes, timestamp, host (IP address of host), outgoing packets and outgoing bytes.
Figure 3.2 is the simulated LAN network of Redeemer's University. It was simulated with the aid of packet tracer application due to the complexity of the real network. It shows the different host computers, switches and router on the LAN and how they are connected together. Internet Protocol (IP) addresses were assigned to each computers and actions that will generate UDP traffic was prompted on the network model in order to study the UDP flow data over the network.

Figure 3.3 is a flowchart summarising all the workflows of the proposed system.

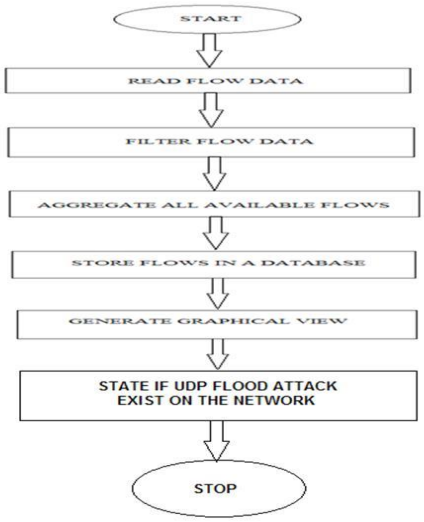


Figure 3.3 Flow Chart of the System

IV. SYSTEM IMPLEMENTATION

Packet Symmetry was used to analyze UDP traffic captured from Redeemer's University LAN for a period of ten minutes. The method was used to find out if a UDP flood attack exists on the network. This chapter discusses data analysis and plots the graph afterwards in order to study the flow pattern. JAVA programming Language was used to implement Packet Symmetry in this project.

4.1 The Testbed

In order to test the system, two networks ANet and BNet were setup within the University's LAN. Figure 4.1 shows how all these networks connect together. Note that ANet and BNet are smaller networks between the University's LAN which are connected directly to a router. The System (UDP Flood Detector) that has been built based on Packet Symmetry was installed on one of the computers in the mini network. The system collects only UDP flow over the network.
The System was set to accept a maximum amount of packets. The System alerts a UDP Flood attack if the maximum number of received packets exceeds the maximum number of transmitted packets.

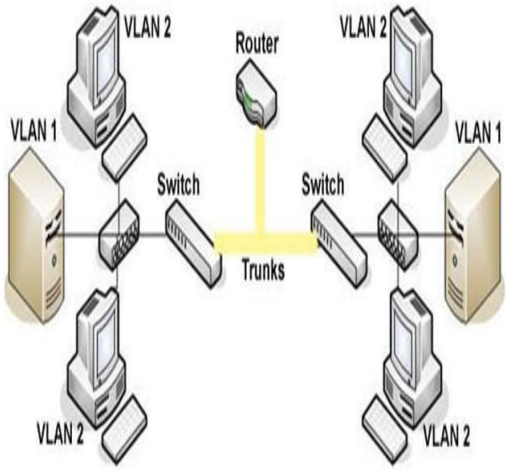


Figure 4.1 Setup that was used to test the System (UDP Flood Detector)

The maximum length a flow can have during exportation is the same as the time bins. That is, we make sure that we set the active timer of the router to be equal to the time bins. The active timer was set to half a minute (30 sec.).

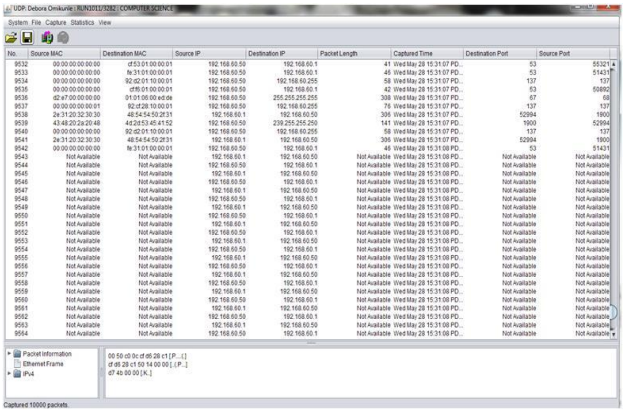


Figure 4.2 Screen shot of flow data captured by the System

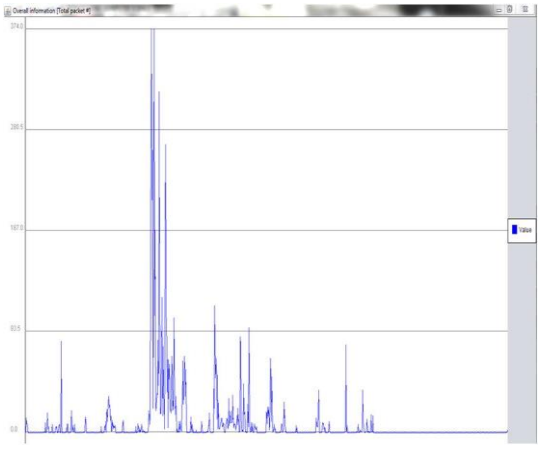


Figure 4.3 A graph showing the cumulative percentage of all packet ratios of data captured within a short period of time.

The time series analysis describes how the packet ratio varies in equal time interval. To get the resulting packet ratio for the entire network per time, the System sums up and divide all sent and received packets.

The System investigates the effects of different sizes of data after flow data has been captured from the network. The System analyzes all ratios and presents a cumulative percentage of them. To study the pattern of all the ratios that have been calculated with different granularities, the System also plot the cumulative percentage of all the ratios.

Figure 4.4 displays a graphical view of the flow on. It is clear that something suspicious was going. It appeared to be a UDP flood towards a host in the RUN network. It is also very clear that the number of received packets and sent packets of RUN hosts are balanced, that is, close to one.

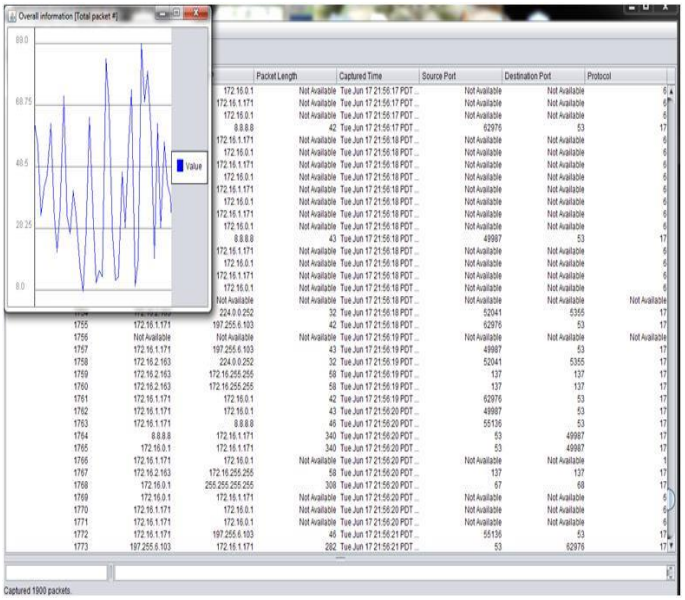


Figure 4.4 Flow data plus graph capture

Figure 4.5 shows an alert that popped-up on detection of a UDP Flood attack.

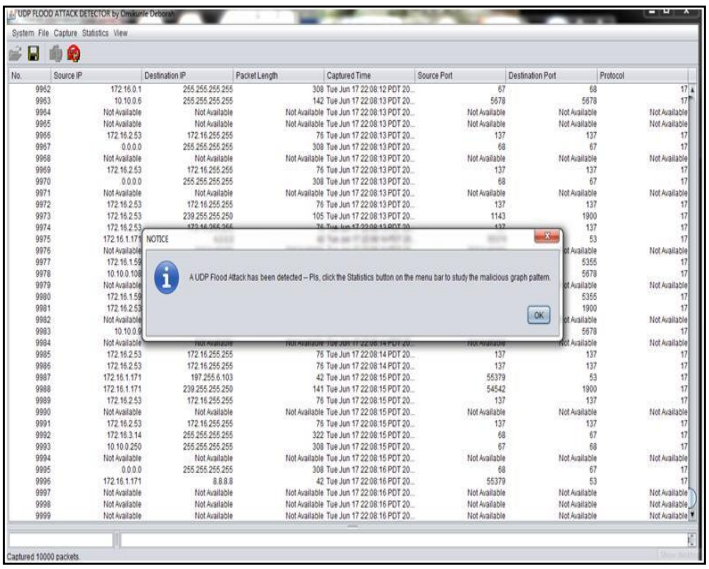


Figure 4.5: AUDP Flood Attack Detection.

Figure 4.4 displays a graphical view of the flow on. It is clear and evident that something suspicious was going. It was confirmed to be a UDP flood towards a host in the RUN network. It is also very clear that the number of received packets and sent packets of RUN hosts are balanced, that is, close to one. Viewing the time series representation of the BNet data, we expect the attacks to be visible. Figure 4.5 shows an alert that popped-up on detection of a UDP Flood attack at those interval.

V. CONCLUSION AND FUTURE WORKS

A brief introduction on the various kinds of attacks that can corrupt the internet was given in the paper. Several systems that could protect the network against these attacks were also discussed. We discovered those systems can rely on flow data analysis or deep packet inspection to analyze the data. Furthermore, some other metrics that can be used to detect attacks were discussed. The study focused on packet symmetry based on packet ratio. This method assumes that the packets sent via a UDP are close or equal to the received packets. If this doesn't happen, then an attack is suspected.

In order to detect the UDP attacks using only flow data, a method specified by one of the earlier researchers based on packet symmetry was used. The metric was capable of detecting almost all floods out of a reference data set.

We evaluated a data set from Redeemer's University's LAN and we were able to detect two flooding attacks; each lasted a few seconds. The main part of the attack(s) consisted of small UDP packets between 32 and 48 bytes. The attack(s) towards the university's LAN lasted a few seconds.

A. Future Works

A good knowledge of when the host is active or inactive is necessary for a better analysis of the granularity of the host. This can be done with the aid of DHCP server logs and the VPN access logs. An algorithm can also be developed to detect active hosts using the flow data. Asides UDP, other protocols can be used to see if a host is inactive or not.

In order to know if this is a good measure for other protocols, further research can be done. Result shows that there can be some symmetry flows from other protocols as presented by [8]. This means that a symmetry of 2:1 is expected between the number of sent packets and received packets in a TCP flow. But there are exceptions where the symmetric ratio turns out to be one (1). Most times, it ranges from between 0.5 and 2.0. This will also help to detect a TCP and other protocols Flood Attack using packet symmetry with only flow data.

Furthermore, future research should be able to differentiate between a UDP flood attack and when there is an upsurge in traffic in order not to categorize a minor upsurge in traffic as an attack.

REFERENCES

[1] M. Brunker, (2003). "Spam Block Lists Bombed to Oblivion".MSNBC.COM 24 .

[2] Cisco. (2007). Netflow solutions guide. Retrieved from http://www.cisco.com/univercd/cc/td/doc/cisintwk/intsolns/netflsol/nfwhite.pdf.

[3] C. Cowan,., P . Wagle, C. Pu, S. Beattie, and J. Walpole, (2000). "Buffer Overflows: Attacks and Defenses for the Vulnerability of the Decade", DARPA Information Survivability Conference and Exposition, 2, 119-129.

[4] P. Criscuolo, (2000). Distributed Denial of Service Trin00, Tribe Flood Network, Tribe Flood Network 2000, And Stacheldraht CIAC-2319". Department of Energy Computer

[5] K. David, and L. Ruby (2001)."Remote Denial of Service Attacks and Countermeasures," Princeton University Department of Electrical Engineering Technical Report CEL2001- 002.

[6] GOVCERT.NL. (2005) Aanbevelingen ter bescherming tegen denial-of-service-aanvallen. Retrieved from http://www.govcert.nl/download.html?f=59.

[7] IPFIEWG. (2008). Retrieved from http://www.ietf.org/html.charters/ipx-charter.html. ISCC 2004, 1, 339–344.

[8] C. Kreibich, A. Warfield, J. Crowcroft, S. Hand, & I. Pratt, (2005). Using packet symmetry to curtail malicious traffic. In Proceedings of the Fourth Workshop on Hot Topics in

[9] G. Lee, (2000). Denial-of-service attacks rip the internet-technology news. Retrieved from mail.im.tku.edu.tw/../00839316.pdf

[10] M. Martin, (2003). "Router Expert: Smurf/Fraggle Attack Defense Using SACLS", Networking Tips and Newsletters,www.searchnetwork.techtarget.com. Retrieved from http://searchnetworking.techtarget.com/tip/1,289483,sid7_gci856112,00.html

[11] Microsoft. (2007). Messenger service window that contains an internet advertisement appears. Retrieved from http://support.microsoft.com/kb/330904.

[12] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, & N. Weaver, (2003).

[13] K. Munivara, M. Rama, & M. Ganesh (2011). Flooding Attacks to internet threat monitors, doi:5121/ijcsit.2011.3612, 161-162.

[14] NRC Handelsblad. Storing bij KPN treft internet en pinverkeer.(2008). Retrieved from http://www.nrc.nl/achtergrond/article1086758.ece.

[15] D. Sanden, (2008). Detecting UDP attacks on high speed networks using packet symmetry with only flow data. Retrieved from www.utwente.nl/../2008-sande.pdf

[16] US-CERT. (2006). The continuing denial of service threat posed by DNS recursion (v2.0). www.phreak.org.www.phreak.org/archives/exploits/denial/fraggle.c

[17] A. Dhruv and P. Hasmukh (2014), Detection and Mitigation of DDOS Attack against Web Server, International Journal of Engineering Development and Research, Volume 2, Issue 2, ppg 2321-9939.

Integrating Enterprise Applications with Distributed Database

Akintola Kolawole Gabriel
Department of Computer Science,
Federal University of Technology, Akure.

Abstract- Enterprise Application Integration (EAI) is defined as the use of software and computer systems architectural principles to integrate a set of enterprise computer applications. With enterprises becoming more and more heterogeneous, distributed computing is one area that is extremely needed to achieve enterprise integration. This paper looks at the various areas in which distributed databases could be used to achieve enterprise integration.

Keywords: Integration, enterprise, software, architecture

I. INTRODUCTION

Enterprise Application Integration (EAI) is an integration framework composed of a collection of technologies and services which form a middleware to enable integration of systems and applications across the enterprise. In my organization, we have several applications such as payroll, course management system, Examination System, Supply chain management applications (for managing inventory and shipping), and Health care system.

These systems typically cannot communicate with one another in order to share data or business rules. This lack of communication leads to inefficiencies, wherein identical data are stored in multiple locations, or straightforward processes are unable to be automated. Enterprise application integration (EAI) is the process of linking such applications within a single organization together in order to simplify and automate business processes to the greatest extent possible, while at the same time avoiding having to make sweeping changes to the existing applications or data structures.

A vertical fragmentation of organizational units with each unit being structured within three architectural layers, is shown in Figure 1.

The *business architecture* layer defines the organizational structure and the workflows for business rules and processes. It is a conceptual level expressed in terms meaningful to actual users of application systems.

The *application architecture* layer defines the actual implementation of the business concepts in terms of enterprise applications. At this layer, it is the central goal to provide the "glue" between the application domain described in the business architecture and the technical solutions described in the technology architecture. Research in information systems aims at filling the gap between business and technology, which requires interdisciplinary cooperation between the application domain and information technology.

The *technology architecture* layer defines the information and communication infrastructure. At this layer, IT is challenged to achieve the business requirements.

The *technology architecture* layer defines the information and communication infrastructure. At this layer, IT is challenged to achieve the business requirements.

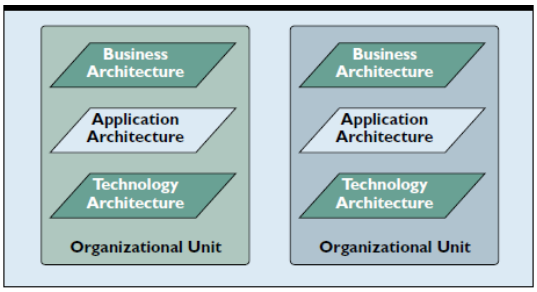


Fig. 1 Vertical fragmentation of Organization Units. (Hasselbring, 2000)

According to (Hasselbring, 2000) enterprise Integration can be achieved by the following:

Inter-organizational processes. At this layer, business engineering (Hasselbring, 2000) seeks to organize a commercial undertaking in a competitive way, whereby business processes cut horizontally through the traditional organization structure. Business process reengineering aims at continuously improving those processes. To support the intra-organizational business processes within organizations effectively, the existing information systems must be integrated.

Enterprise application integration. The goal is to integrate independent enterprise resource planning (ERP) systems at this layer. This is usually achieved by means of some kind of messaging services.

Middleware integration. At this layer, the techniques for building componentized information

systems with state-of-the-art infrastructures such as CORBA, database gateways, and transaction monitors, are employed. Middleware integration addresses the syntactical level ("plumbing" and "wiring") while Enterprise Application Integration also addresses a semantic level.

The borderline between Enterprise Application and middleware integration cannot always be pinpointed precisely. For instance, the Object Management Architecture of the OMG defines the Object Request Broker, which can be deployed for middleware integration, and also high-level services (such as business objects) that address Enterprise Application Integration.

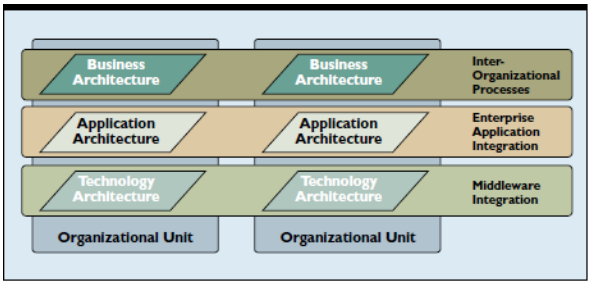


Fig 2. Horizontal integration to support the business processes. (Hasselbring, 2000)

The deployment of Enterprise resource planning (ERP) systems often requires reengineering the business processes to align with the ERP system. However, it is usually unacceptable to require the business to change to the applications' functionality; instead the information architecture should align with the business organization. SI and componentization aim at supporting the business processes, while preserving the investments in (legacy) systems [Hasselbring, 2000].

The Disciplines in System Integration [Hasselbring, 2000] **are:** Parallel and distributed computing, Database systems, Software Engineering, Artificial Intelligence, Multimedia Systems. In this paper we look at the database aspect of enterprise integration.

II. The Role of Database in System Integration

This paper looks at how enterprise can be integrated using database systems. it will cut across the role of distributed database systems and datawarehouse in system integration.

Problems of coupling and integrating heterogeneous database and information systems have been addressed in the database area for some time [Hasselbring, 2000]. While research on parallel and distributed systems emphasizes the integration of computational components, research on database systems is more concerned with the integration of data. Federated database systems, for instance, approach the integration of heterogeneous databases by means of schema integration. However, when it comes to managing transactions over multiple local systems, for instance for executing transactional workflows, the problems to be solved are often very similar to those that arise in parallel and distributed systems. These similarities sometimes cause the reinvention of techniques that have already been elaborated in other disciplines [Hasselbring, 2000].

Distributed Database

It is increasingly important for organizations to achieve additional coordination of diverse computerized operations. To do so, it is necessary to have database systems that can operate over a distributed network and can encompass a heterogeneous mix of computers, operating systems, communications links, and local database management systems. Distributed Database Management System (DDMS) is specially developed for Heterogeneous Database Platforms. 'Distributed Database Management System' focus mainly on heterogeneous distributed data management. It is a system for managing databases stored on multiple computers in a network [Ani etal. 2006].

A distributed database is a set of databases stored on multiple computers that typically appears to applications as a single database. Consequently, an application can simultaneously access and modify the data in several databases in a network. Distributed Database management System is a software system that manages a distributed database while making the distribution transparent to the user. An enterprise of (say) more than 100 computers that act as database clients. If there is only one database server, then all the queries sent by these clients have to be answered by the same single server. This will lead to severe traffic congestion at the server and hence on also in the network. It is common in an enterprise to have many departments. Each of these departments will need access to a common database used throughout in the Enterprise. However the majority of data requirements of each of these departments will be department specific.

So this paper will look at the technology of distributed database system. That is how we can integrate different departments or organizations with different databases using distributed computing. It will discuss distributed databases, architectures of distributed databases, query processing, and look at some applications that supports distributed computing.

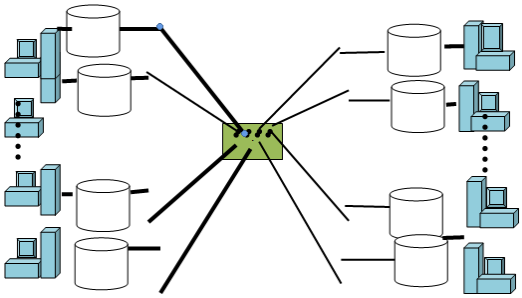


Fig. 3. Conceptual Diagram of the distributed database Integration System

III. DISTRIBUTED DATABASE SYSTEMS

A distributed database is a database in which storage devices are not all attached to a common CPU. It may be stored in multiple computers located in the same physical location, or may be dispersed over a network of interconnected computers. Collections of data (e.g. in a database) can be distributed across multiple physical locations. A distributed database can reside on network servers on the Internet, on corporate intranets or extranets, or on other company networks. The replication and distribution of databases improves database performance at end-user worksites [Wikipedia].

A distributed database system (DDB) is a collection of multiple, logically interrelated database distributed over a computer network. A distributed database management system (D-DBMS) is the software that manages the DDB and provides an access mechanism that makes this distribution transparent to the users. So distributed database system (DDBS)= DDB + D-DBMS.

Distributed Database environment

According to Craig Borysowich (2007), a distributed database is not stored in its entirety at a single physical location. Instead, it is spread across a network of computers that are geographically dispersed and connected via communications links. A distributed database allows faster local queries and can reduce network traffic. With these benefits comes the issue of maintaining data integrity. A key objective for a distributed system is that it looks like a centralized system to the user. The user should not need to know where a piece of data is stored physically.

Forms of Distributed Data

According to Craig Borysowich (2007), there are five categories of distributed data: replicated data, horizontally fragmented data, vertically fragmented data, reorganized data, separate-schema data.

Replicated Data

Replicated data means that copies of the same data are maintained in more than one location. Data may be replicated across multiple machines to avoid transmitting data between systems. Replicas can be read only or writable. Read only replicas have changes made to the original and then propagated outwards to the replicas. Writable replicas propagate changes back to the original using either a "write through" or a "write back" strategy. Write through implies a synchronous connection and a "real-time" update to the original. The write back strategy allows changes to be propagated when it is most appropriate (i.e., a "store-and-forward" or an asynchronous concept).

Consider the timeliness of the transactions against the data. How up-to-date or current does the data/replicas have to be? The biggest concern with replicated data is how to handle "collisions." Refer to Issues to Consider When Distributing Data for further information on handling collisions. Replicated data can simplify disaster recovery because data can be restored to the failed site from one of the replicated copies. Replicated data is most effective when data is not updated frequently. This tool suite maintains replicated databases for each installation.

Horizontally Fragmented Data

Horizontally fragmented data means that data is distributed across different sites based on one or more primary keys. This type of data distribution is typical where, for example, branch offices in an organization deal mostly with a set of local customers and the related customer data need not be accessed by other branch offices.

Vertically Fragmented Data

Vertically fragmented data is data that has been split by columns across multiple systems. The primary key is replicated at each site. For example, a district office may maintain client information such as name and address keyed on client number while head office maintains client account balance and credit information, also keyed on the same client number.

Reorganized Data

Reorganized data is data that has been derived, summarized, or otherwise manipulated in some way. This type of data organization is common where decision-support processing is performed. There may be some instances where the on-line transaction processing (OLTP) and decision-support database

management systems are different. Decision-support typically requires better query optimization and ad hoc SQL support than does OLTP. OLTP usually requires optimization for high-volume transaction processing.

Separate-Schema Data

Separate-schema data maintains separate databases and application programs for different systems. For example, one system may manage inventory and one may handle customer orders. There may be a certain amount of duplication with separate-schema data.

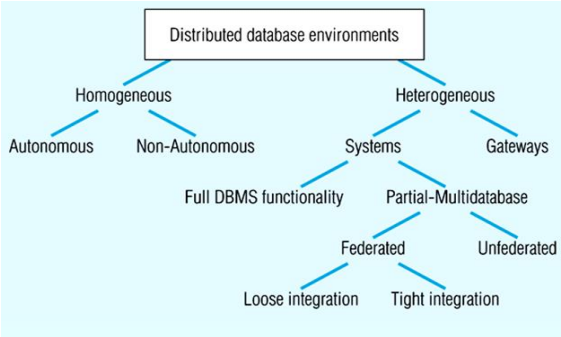


Fig.4. Distributed database environments

Types of distributed database systems

- Homogeneous - Same DBMS at each node
  - Autonomous - Independent DBMSs
  - Non-autonomous - Central, coordinating DBMS
  - Easy to manage, difficult to enforce

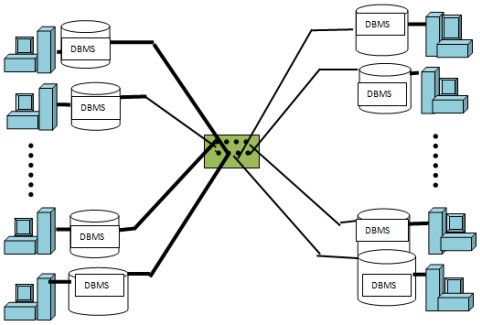


Fig.5. Conceptual Diagram of the Homogeneous Integration System

- Heterogeneous - Different DBMSs at different nodes
  - Systems – With full or partial DBMS functionality
  - Gateways - Simple paths are created to other databases without the benefits of one logical database
  - Difficult to manage, preferred by independent organizations

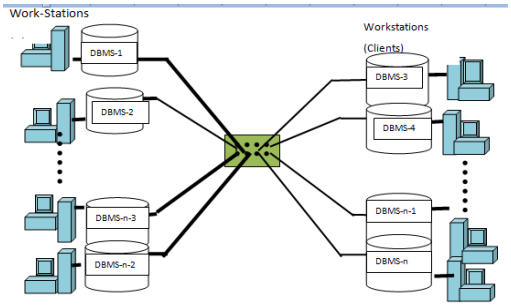


Fig.6. Conceptual Diagram of the Heterogeneous Integration System

IV. Enterprise Integration Method using Shared Database:

Databases play a very vital role in enterprise integration. Any application requires data access. Database systems are integrated with related applications and can be accessed only by the interfaces provided by the Application Server (AS) or middleware. There are different kinds of middleware's -- RPC (remote Procedure Clients), CORBA (Common Object request Broker Architecture), J2EE now JEE (Java Enterprise edition), MOM (Message Oriented Middleware) and JMS (Java Message Service). All these technologies are either design patterns or commercial products.

Database Middleware Systems:

Database Middleware can be defines as "systems used to integrate collections of data sources over computer network". Database Middleware Systems use the concept of 'data integration servers' [Manuel, 2000]. The Data integration Servers provide a uniform to the application viewing the data. There are two ways of deploying an 'integration server' – 1) Database Gateway and 2) Database Mediator [Manuel, 2000].

Database Gateway

In the Database Gateway approach the commercial database is configured to access a remote database through a 'gateway'. The gateway is responsible for providing access methods to the remote data.

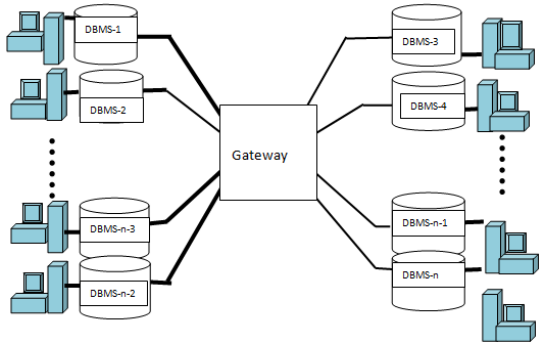


Fig.7. Conceptual Diagram of a gateway Integration System

Data Base Mediators

In this approach, the integration server used a mediator server for distributed query processing. The mediator uses the functionality of 'wrappers' to access and modify the data. In these two approaches the user defined types and the query operators are defined in global operators and contained in libraries. The user defines libraries must be linked to clients in the system.

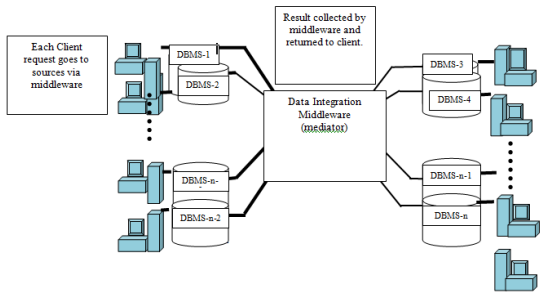


Fig.8. Diagram of the Homogeneous Integration System

Dataware House

In simple terms, a data warehouse (DW) is a pool of data produced to support decision making; It is also a repository of current and historical data of potential interest to managers throughout the organization. Data are usually structured to be available in a form ready for analytical processing activities (e.g. online analytical processing [OLAP], data mining, querying, reporting and other decision supporting applications). A data warehouse is a subject-oriented, integrated, time -variant, non-volatile collection of data in support of management's decision-making process. [Efraim T., et al. 2010]. The day-to-day operations of an organization are done by using the OLTP system. This system is good for normal operations and few decision making but the system is inadequate when it comes to strategic decision support. The lack of historical data in OLTP makes it unsuitable to provide a comprehensive information about the operations of the business. DW on the other hand, provides a central repository of historical data which provides an integrated platform for historical analysis of data. With a data warehouse and Online Analytical Processing (OLAP), users can perform better data analysis and gain better knowledge from the repository data.

V. Designing the Data Warehouse

The following discussion outlines the process of the data warehouse design. It involves the logical design, the OLAP design, and Data mining design.

The logical Design

The logical design of data-warehouse is defined by the dimensional data modeling approach Kimball and Ross (2002). Unlike the Entity Relationship (ER) and Unified Modeling Language (UML) data modeling processes, the logical design of Data Warehouse (DW) is defined by the dimensional data modeling approach. To minimize the join operations which slow down queries, normalization is not the guiding principle in DW design. A schema is a collection of database objects, including tables, views, indexes, and synonyms. There is a variety of ways of arranging schema objects in the schema models designed for data warehousing. The following are the two types of schemas commonly used in dimensional data modeling.

•Star schema

The star schema is perhaps the simplest data warehouse schema. It is called a star schema because the entity-relationship diagram of this schema resembles a star, with points radiating from a central table. The center of the star consists of a large fact table and the points of the star are the dimension tables.

A star schema is characterized by one or more very large **fact** tables that contain the primary information in the data warehouse, and a number of much smaller **dimension** tables (or lookup tables), each of which contains information about the entries for a particular attribute in the fact table.

Star schema facilitates quick response to queries. The core detailed values are stored in fact table.

The dimensional info and hierarchies are kept in dimension tables.

•Snowflake schema

The snowflake schema is a more complex data warehouse model than a star schema, and is a type of star schema. It is called a snowflake schema because the diagram of the schema resembles a snowflake.

Snowflake schemas normalize dimensions to eliminate redundancy. That is, the dimension data has been grouped into multiple tables instead of one large table. For example, a product dimension table in a star schema might be normalized into a products table, a product category table, and a product manufacturer table in a snowflake schema. While this saves space, it increases the number of dimension tables and requires more foreign key joins. The result is more complex queries and reduced query performance.

VI. Dimensional Data Modeling Approach

The dimensional approach is quite different from the normalization approach followed when designing a database for daily operations.

Data Hierarchies in Dimensional tables

Each of the dimensions contains at least one hierarchy. The hierarchies allow users to analyze data aggregations using the OLAP. This allows related items to be grouped and summarized for high level analysis while retaining the ability to drill down to more specific product detail.

The dimensional tables keep the data that are used for analyzing the data warehouse. For example we can look for Faculty who taught a particular course.

Fact Table

Fact table contains dimension attributes and measures. Dimension attributes are FKs or other attributes called degenerate dimension (<<dd>>). Measures are the values to be aggregated when queries group rows together. The Fact table is composed of two types of attributes: dimension attributes and measures.

Data Warehouse are commonly organized with one large central fact table, and many smaller dimensions tables. This configuration is termed a star schema. Below is the diagram of the relationship between the fact table and the Dimensional tables.

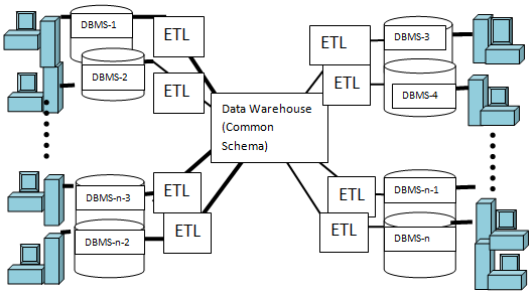


Fig. 9. Conceptual Diagram of a Data warehouse Environment

VII. Conclusion

Database systems play a pivotal role in enterprise integration today. As we see diverse databases in different organizations joining together through ICT for business enterprising, more technologies are evolving at meeting the standardization data Integration. XML is one of the current technologies today for data integration and EDI. More middle ware will continue to emerge for data integration .

References

[1] Ani joy ANIL JOY ARUN C P ARUN KUMAR E 2006 *Distributed Data Base* Management System http://www.seminarprojects.com/Thread-distributed-database-management-system-full-report#ixzz1U4Nuvy9R

[2] http://www.seminarprojects.com/Thread-distributed-database-management-system-full-report

[3] Craig Borysowich (2007) Overview of Distributed Databases : Enterprise Implementation Issues & Solutions http://it.toolbox.com/blogs/enterprise-solutions/overview-of-distributed-databases-16228

[4] Hasselbring, W.(2000) *Information System Integration.* Communications of the ACM, 2000. **43**(6): p. 32-38.

[5] Leon Chen Lecture Notes on distributed database.

[6] Manuel Rodr´ıguez-Mart´ınez, Nick Roussopoulos, (2000) MOCHA: A Self-Extensible Database Middleware System for Distributed Data Sources_, MOD 2000, Dallas, TX USA

[7] Efraim T., Jay E., Teng-Peng L., Ramesh S, (2010). Decision Support and Business Intelligence Systems (8th ed) Prentice Hall.

[8] Infogold Data Warehouse, Data Mart, Data Mining, and Decision Support Resources, http://infogoal.com/dmc/dmcdwh.htm

[9] Kimball, R. and Ross, M., (2002). The Data Warehouse Toolkit, second edition. John Wiley and Sons, Inc., USA

Performance Evaluation of Wired and Wireless LAN under Different Scenarios

Daramola O.A.
Computer Science Department
School of Sciences
Federal University of Technology, Akure
Nigeria

Kuboye B.M.
Computer Science Department
School of Sciences
Federal University of Technology, Akure
Nigeria

Ogunleye O.M.
Computer Science Department
School of Sciences
Federal University of Technology, Akure
Nigeria
hollubunmy.pcwizard@yahoo.com

***Abstract*—**Computer networks have grown rapidly since their inception over the past few decades as they provide the means for us to communicate, transfer files, and access network resources. Usage of the Internet is becoming more important due to the reason that network technologies are now adopted for purposes like education, business, banking, defence, communication and so on. Complex and rapid developments have been made by researchers to enhance Wired and Wireless Networks. The wired computer networks offers a secured and faster means of connectivity, however, the need for mobility is gradually shifting the ground in favour of wireless network technology. This work makes a presentation of current research works in the field of wired and wireless network technologies. Performance of wired and wireless LAN network is simulated under different scenario to determine the suitability of each network technology for certain situation. OPNET (Optimised Network Engineering Tool) network simulator is employed in the simulation. OPNET provides an easy way of predicting and estimating the performance of networks. Performance metrics like throughput, delay and retransmission attempts are measured to analyse the performance of wired and wireless computer networks in a LAN environments.

Keywords- Computer Networks, Wired Networks, Wireless Networks, OPNET, Evaluation

I. INTRODUCTION

The invention of computer has been one of the greatest forces in technological advancement over the ages as it has helped to shape the human world due to its application in virtually all areas of human lives. One of the after-effects of the computer invention is the computer network. Computer network can simply be referred to as the interconnection of different computers and other devices for the purpose of accessing and sharing common resources, which could be hardware resources such as printers or software resources such as files.

Computer network over the years has contributed immensely to data and information sharing because more than one computer can now connect and communicate with each other through a medium (wired and wireless) and some set of protocols for the purpose of transferring and sharing data. Over the past few decades, networks have grown tremendously as they provide means for making use of computer resources. For instance, the internet is becoming more popular day by day due to the wide usage of network technologies, where they serve several purposes in areas such as education, business, banking, government and even in defence [1].

Wired and wireless networks have advanced greatly over the years due to the fact that several network models have been developed by researchers with the aid of network simulators for the purpose of optimization and performance evaluation. Computer simulation has been extensively used in investigating these networks, the reason for this is that it the cost of setting up and implementing real network models are reduced.

Wired transmission of network data comes with the problem of mobility and extensive cabling, while on the other hand, wireless network transmission has overcome these constraints as it offers a flexible means of data communication, which is implemented together with the wired network or differently as an alternative means of data communication. Despite these differences, wired and wireless networks offers closely similar bandwidth and services to network users. [1].

As networks are developed and modified, network planning has become highly imperative for the purpose of determining their viability and validating their performance. This task of evaluation can be very expensive and tedious when making use of real life systems, but computer simulation offers an easy and simpler way of carrying out the tasks of evaluating and predicting the performance of a computer network design. Due to this reason, simulators have become powerful tools in the hands of network analysts in designing, modifying and evaluation of computer networks, [2].

Therefore, in this study, simulation technique will be adopted in studying and comparing the performances of wired and wireless LANs under different scenarios of varied number of users and packet sizes. These scenarios are of consideration because they have a telling effect on network performance. With more users and more packets being generated in a network, it is obvious that network traffic will increase, which in turn leads to more collision between the packets and this ultimately affects the network's throughput. But the question to be answered in this work is that; which of the two networks (wired LAN and wireless LAN) handles a particular scenario best?

II. RESEARCH MOTIVATION

The performance of network is very important in telecommunication and the effectiveness of other network applications, [2]. It is very important to have a clear understanding of how these networks perform by considering parameters, which include throughput, delay, collision count etc. The Local Area Network is the most common type of computer network available owing to its cheaper cost of setup, easy maintenance and others. It is the type of network that is found on campuses, offices and even homes.

This popularity has led to several research works. In the process of going through previous related works in this work area, it was discovered that some of the evaluations were limited to one type of network (either wired or wireless) because the basis for evaluation was not feasible for both networks. They evaluated four access methods to determine the effect of transmission errors on their performances by considering the Bit Error Ratio (BER). The access methods considered were Distributed Coordination Function (DCF), Slow Decrease Method, Asymptotically Optimal Backoff and the Idle Sense Mechanism. At the conclusion of their research, it was discovered that the performances of these access methods varies as the BER also varies. This could not be feasible of wired LAN because of different access methods.

Also in [4], evaluation of Ethernet LAN was carried out using the same packet size for different Ethernet wiring standards and the researchers admitted the fact that varying packet sizes will definitely produce different results.

Another reason for taking on this research is for the purpose of learning network simulation and make use of it in evaluating network performance. Some of the researches done in this field were done using mathematical analysis. For instance, [5], analysed the performance of CSMA/CA protocol in wireless LAN to verify its applicability and performance enhancements. He did this by making use of analytical method. [6], pointed out that analytical method are sometimes based on assumptions, which makes the result obtained not always accurate. Based on this, analytical method sometimes is inadequate in evaluation unlike computer simulation, which requires fewer assumptions.

Finally, it will be justifiable to compare wired and wireless LANs using a uniform benchmark and to see the effects of varying parameters on their performances as this will enable us to determine which of the two networks performs better under various conditions.

III. REVIEW OF RELATED WORKS

This section gives a brief review of related literatures concerning the subject matter of this research. In carrying out this research the following works were reviewed, however, the literatures reviewed are not limited to the ones discussed below.

i. Performance Analysis of Wireless LANs [5]. The objective of this work was to analyse the performance of the CSMA/CA protocol and to verify its applicability and performance enhancement in wireless LANs and the methodology adopted was Mathematical Analysis of CSMA/CA protocol. The major limitation with this work was the inadequacy of analytical models in evaluation and also the real-life operation of the network was not modeled it was only the protocol.

ii. Enhancements and Performance Evaluation of Wireless LAN [7]. Objective and Motivation for this work was the need to improve the loss performance of wireless LANs due to their limited bandwidth. In this work, three approaches of improving wireless LAN performance were simulated. These approaches are; (1) Tuning the physical layer related performance. (2)Tuning the IEEE 802.11 parameters and (3) Using enhanced link layer protocol.

iii. Analysis of LAN under different Ethernet Wiring Standards with variation in Time and Components [4].Motivation and objective for the research was the emergence of new standards in Ethernet and the need to determine the effects of varying standards on network performance. The methodology adopted was simulation of LAN under the 10BaseT and 100BaseT wiring standards with varying time of propagation. Performance metrics such as collision count, traffic received was measured. The limitation in the research was the restriction to only the 10BaseT and 100BaseT wiring standards and the usage of a constant frame size. The research was also limited to wired LAN.

IV. RESEARCH OBJECTIVES AND METHODOLOGY

This work was aimed at taking an in-depth study of wired and wireless local area networks, for the purpose of comparing how they perform under varying network conditions and also to determine which network will be suitable for use with respect to a particular situation. This work assists in (1) Determining what condition is suitable for each type of network. (2) Providing a means for network administrators to understand the reasons for the good or bad performance of their networks. (3) Providing adequate knowledge of the local area network and (4) Understanding the concept of simulation in networking.

The network simulator used is OPNET I.T Guru due to its popularity and availability of an academic version for free. The metrics measured are throughput, collision count and latency while the parameters for measuring them are the number of users and packet sizes sent through the network. Simulation environments for the two networks considered were created and the metrics were measured over each network for each parameter. A simulation model of wired and wireless LAN was also created for both high load and low load conditions. Packets were sent through these networks and data was collected for the performance metrics chosen. Also utilization of the network hubs was one of the metrics measured so as to determine the intensity of activities taking place in it for every simulation model developed.

The collected data was displayed in a graphical format for the sake of clarity and this data was analysed and compared for both wired and wireless LAN for the purpose of achieving the goal of this research, which is to determine the performances of the two networks.

V. EVALUATION TECHNIQUES

Simulation of wired and wireless LANs was done using OPNET (Optimized Network Engineering Tool) to evaluate the performance of the two networks. Simulation is preferred to other evaluation techniques in this work because it is able to provide an almost real life experience of the network environment. Simulation is defined as the imitation of the behavior of a real system over time.

OPNET was used because it provides a comprehensive development environment for the simulation and performance analysis of communication networks. It supports a wide range of communication systems from a single LAN to the internet. OPNET makes use of discrete event simulations (DES) to analyse system performance and their behavior. It can simulate a wide variety of different networks which are linked to each other. The following points are highlighted as reasons for making use of OPNET:

i. It has a high support level for network models.
ii. It provides a simple user interface to establish several networks.
iii. The problem of coding, which is involved in other simulators like NS2 is avoided in OPNET.
iv. It enables a drag and drop approach for simulation and this makes it easy to use.
v. In addition to the above points, OPNET provides simulation environment for latest network technologies such as the Mobile Adhoc networks, Zigbee networks and wireless mesh networks.
vi. There is availability of a free academic edition of the software for the purpose of academic usage.

In carrying out performance evaluation of a system, it is necessary to have some parameters that will be used in evaluating the system's performance. Parameters are those factors that define a system and determine its performance. In this work, two parameters were considered in evaluating the performance of wired and wireless local area networks. These parameters are number of users and packet sizes.

A. Packet size

A network packet refers to a formatted unit of data transferred by a packet switched network. In transferring data over the network, the data is broken down into smaller units called the frame or packet.

B. Number of Users This can also be referred to as nodes. A network node is a point of connection to a communication network. It can also be referred to as an active electronic device with the capability of sending and receiving information over a communication channel, which is connected to a communication network. These two parameters will be used in the evaluation experiment by varying them for each of the network to be evaluated.

VI. SIMULATION SCENARIOS

For the purpose of effective comparison of the two networks under consideration, the following four (4) different scenarios were considered in the simulation process.

- *Scenario 1*

This scenario contains both wired and wireless LAN each comprising of twenty (20) users with network packet generation of 100 byte packets at every 4 milliseconds. This scenario is called 20 user low load.

- *Scenario 2*

This scenario contains both wired and wireless LAN each comprising of twenty (20) users with network packet generation of 100 byte packets at every 1 millisecond. This scenario is called 20 user high load.

- *Scenario 3*

This scenario contains both wired and wireless LAN each comprising of fifty (50) users with network packet generation of 100 byte packets at every 4 milliseconds. This scenario is called 50 user low load.

- *Scenario 4*

This scenario contains both wired and wireless LAN each comprising of fifty (50) users with network packet generation of 100 byte packets at every 1 milliseconds. This scenario is called 50 user high load.

VII. METRICS USED IN EVALUATION

Metric simply refers to a system of related measures that facilitates the quantification of some particular characteristics. The metrics to be measured in this simulation are; throughput (traffic received), collision count, transmission attempt and latency.

A. Throughput

Network throughput refers to the rate of successful network data delivery over a communication channel. It is made up of the sum of data rates that are delivered to all terminals in a network.

B. Collision

Collision occurs in a network when there is interference between two or more devices on the same network attempting to transmit data at exactly the same time.

C. Transmission/Retransmission Attempt

This metric is a function of the collision that occurs within the network nodes. It is a measure of the rate at which the network attempted to retransmit collided packets until they are successfully transmitted or discarded due to expiration of the retry limit.

D. Latency

Network latency refers to any kind of delay that occurs in data transmission over a network.

VIII. SIMULATION SET-UP AND RESULTS

A. Simulation Set-up

In this simulation, four (4) network models were designed with two different scenarios each. The models are

i. Wired LAN with 20 user nodes: This model consists of 20 nodes connected to two (2) 16-port hubs at 10 nodes per hub. These hubs are then linked with each other through a network switch. For the low load scenario, the packet size is set to 100, while the interarrival time is set to 0.004 seconds, which means that the average traffic generation of each station is at the rate of 100-byte at every 4 milliseconds. The average traffic that each node generated is calculated as:

IX. SIMULATION SET-UP AND RESULTS

A. Simulation Set-up

In this simulation, four (4) network models were designed with two different scenarios each. The models are

i. Wired LAN with 20 user nodes: This model consists of 20 nodes connected to two (2) 16-port hubs at 10 nodes per hub. These hubs are then linked with each other through a network switch. For the low load scenario, the packet size is set to 100, while the interarrival time is set to 0.004 seconds, which means that the average traffic generation of each station is at the rate of 100-byte at every 4 milliseconds. The average traffic that each node generated is calculated as:

100 bytes/packet x 8 bits/byte x 1packet/0.004sec = 800bits x 1/0.004 = 200kbps

With these, the entire network is then expected to generate traffic at the rate of 4mbps (200kbps x 20 nodes).

For the high load scenario, packet size remains at 100 but the interarrival time is set to 0.001 seconds, which means each node generates 100 byte packet every 1 millisecond resulting into traffic generation at 800kbps and the entire traffic generated by the network stands at 16mbps.

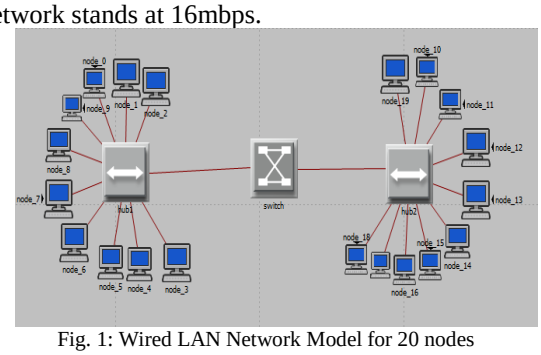


Fig. 1: Wired LAN Network Model for 20 nodes

ii. Wired LAN with 50 users

This model consists of 50 nodes connected to four (4) 16-port hubs at 12 nodes per hub. These hubs are then linked with each other through a network switch. For the low load scenario, each node generates traffic at the rate of 200kbps, while the entire network generates traffic at 10mbps. For the high load scenario, packet generation at each node is at 800kbps and for the entire network is 40mbps.

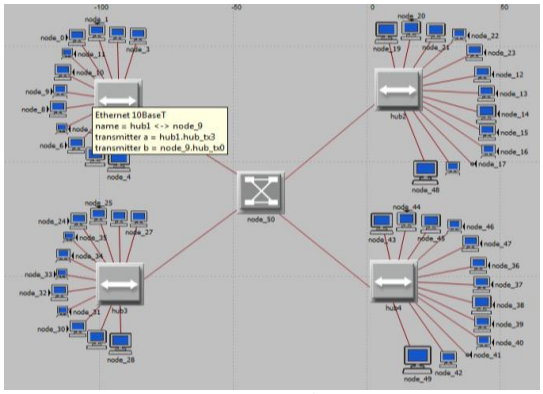


Fig. 2: Wired LAN Network Model for 50 nodes

iii. *Wireless LAN with 20 nodes*:

This model consists of 20 nodes connected to a central Access Point. For the low load scenario, each node generates traffic at the rate of 200kbps, while the entire network generates traffic at 4mbps. For the high load scenario, packet generation at each node is at 800kbps and for the entire network is 16mbps.

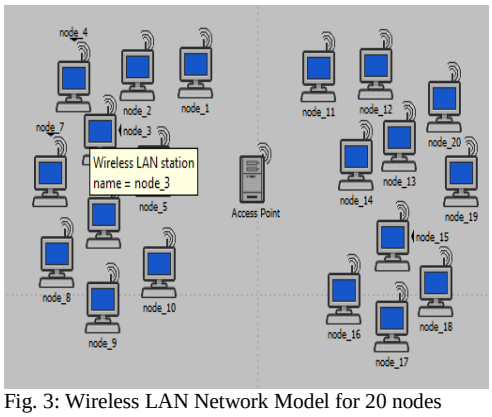


Fig. 3: Wireless LAN Network Model for 20 nodes

iv. Wireless LAN with 50 nodes

This model consists of 20 nodes connected to a central Access Point. For the low load scenario, each node generates traffic at the rate of 200kbps, while the entire network generates traffic at 16mbps. For the high load scenario, packet generation at each node is at 800kbps and for the entire network is 40mbps.

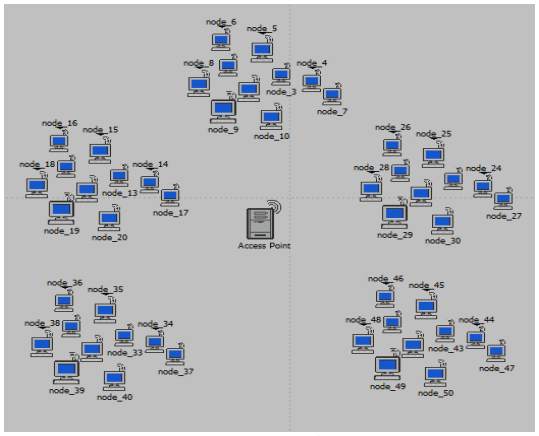


Fig. 4: Wireless LAN Network Model for 50 nodes

B. Simulation Results

From the Fig. 5 below, it is seen that the rate of traffic received in bits/sec is approximately equal to the sent bit rate for wired LAN with 20 nodes and each node generating traffic at the rate of 200kbps, while there is a noticeable decrease in traffic sent and received rate for wireless LAN with the same parameters as that of the wired LAN. For wired LAN, there is stability in the traffic generation and delivery rate. The network begins generation of traffic after about 5 seconds, generating and delivering packets at about 4mbps. For the wireless network, there is a massive decrease from traffic generation at 4mbps to delivery at an average of about 200kbps.

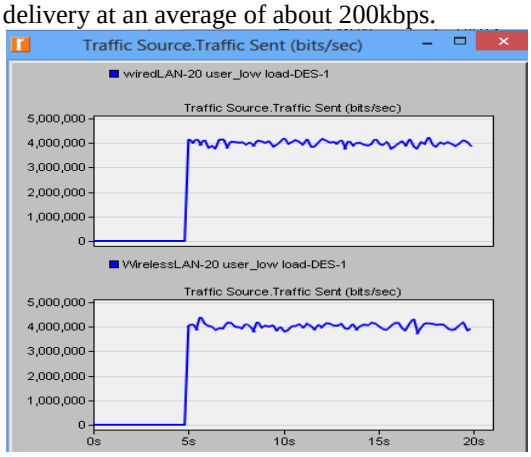


Fig. 5. Traffic sent (bits/sec) for 20 users under low load

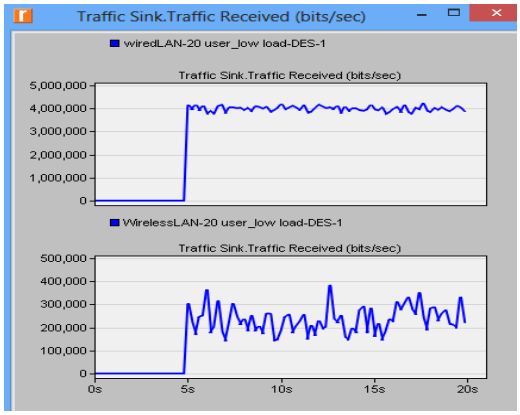


Fig. 6. Traffic received (bits/sec) for 20 users under low load

The differences between the sent and receive rates in both networks can be accounted for by inspecting the latency, which is the end to end delay of traffics which are successfully transmitted in the network. From the fig 6, it is seen that the for the wired LAN, the latency is almost negligible at a maximum of 0.00040 bit per second compared to wireless LAN with a delay of almost 0.5 bit per second. This in turn resulted into more retransmission attempt in wireless LAN than we have in wired LAN as it is seen in Fig. 7.

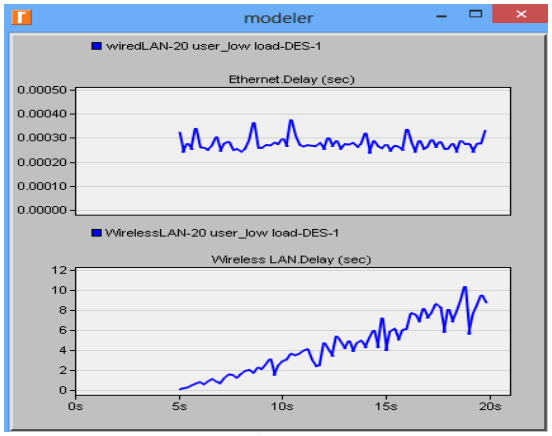


Fig. 7. Transmission delay for 20 users under low load

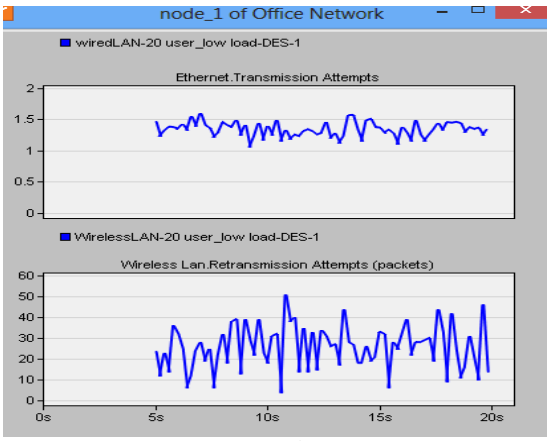


Fig. 8. Retransmission Attempt for 20 users under low load

From the Fig. 9, it is noted that keeping the number of users the same and increasing the number of packets generated by each node makes a noticeable effects on traffic generated and received ratio of wired LAN. It is seen that only half of the generated traffic was received. However, traffic received in wireless LAN is almost the same with that of low load.

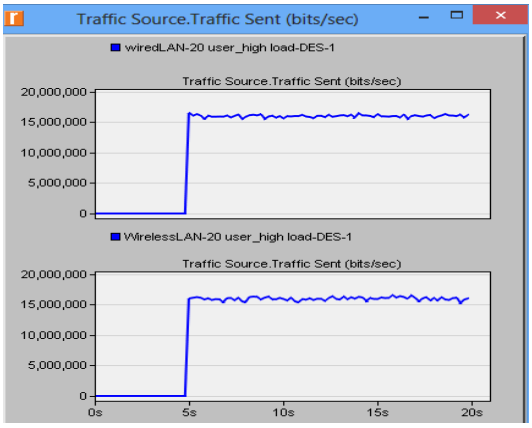


Fig. 9: Traffic sent (bits/sec) for 20 users under high load

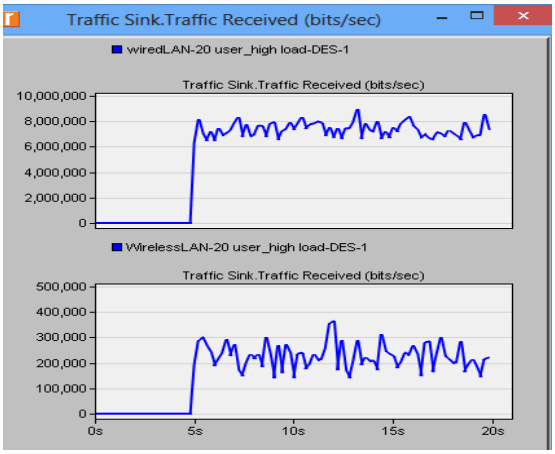


Fig. 10. Traffic received (bits/sec) for 20 users under high load

From Fig. 11, when the number of users is increased from twenty to fifty and traffic generated by each node is kept at 200kbps, it is observed that traffic generated and received by wired network remains almost equal but far lesser traffics are received than what is generated by the wireless LAN. The reason for this is the same as we have in the first scenario. More packets are delayed in the wireless LAN, which results into more collision and retransmission attempt as it is indicated in Fig. 12.

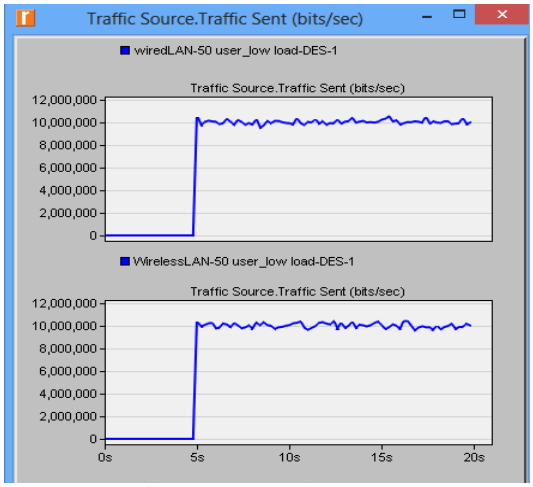


Fig. 11. Traffic sent (bits/sec) for 50 users under low load.

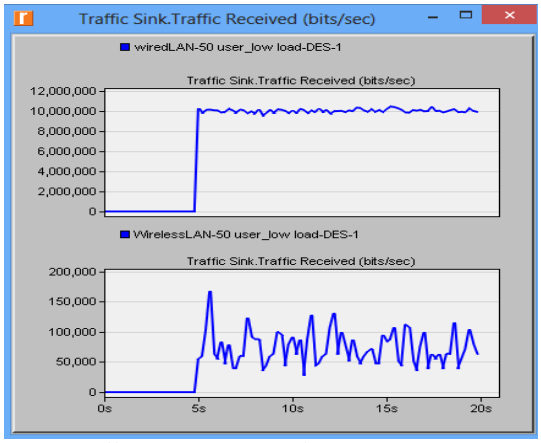


Fig. 13. Traffic received (bits/sec) for 50 users under low load.

In Fig.13, when the traffic generated by nodes is increased to 800kbps, and number of user nodes remains at fifty, number of packets received in the wired LAN is considerably low to what is sent. The network was only able to deliver averagely one-fifth of the packets sent. For the wireless network, the network only experienced little data dropped compared to what is obtainable in the wired network.

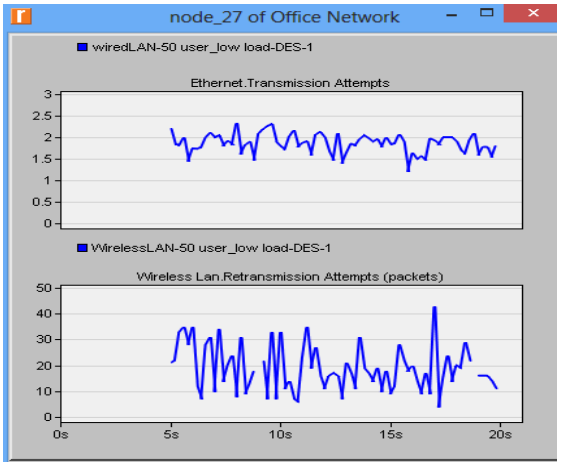


Fig. 14. Retransmission attempt for 50 users scenario under low load

As always, with more packets generated, more packets are delayed in the wired network unlike when generated packets are less, while increase in packets generated has little or no effects on the wireless network.

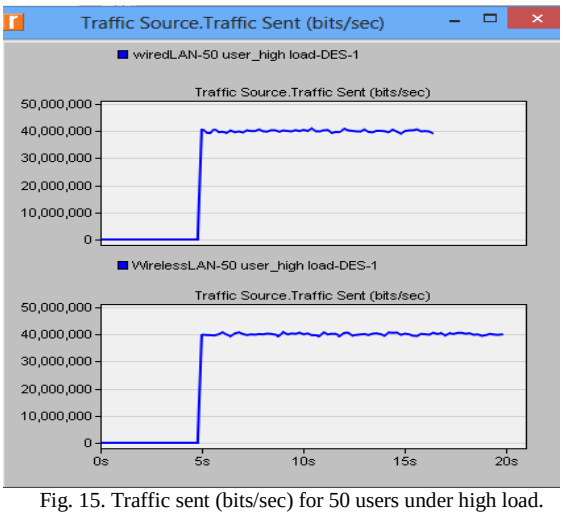


Fig. 15. Traffic sent (bits/sec) for 50 users under high load.

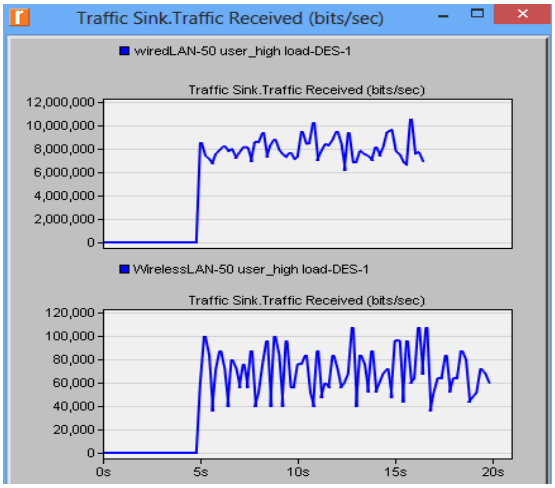


Fig. 16.Traffic received (bits/sec) for 50 users under high load.

From the simulation results obtained in Fig 15, it could be observed that as more packets are generated in a network, traffic continues to increase in the network, and this in turn reduces the throughput. In a wired local area network, it is seen that when number of network nodes is kept the same and packet generated by each network is increased, more traffic is created in the network and this ultimately affects the throughput of the network.

Fig. 16.Traffic received (bits/sec) for 50 users under high load.

For instance, in the case of wired LAN with 20 users, when each node is generating packets at the rate of 200kbps, the network was sending and delivering data at the rate of 4mbps but when packet generated by each node increases to 800kbps, the average traffic sent was at 16mbps while traffic received was at a low rate of 8mbps, this indicates about 50% drop in data delivered. This sharp drop was as result of more delay and collision that occurred within the network hubs.

Conversely, the wireless LAN, performs poorly compared to wired LAN even when subjected to the same condition. From the simulation, with the network comprising of 20 nodes and each node generating traffic at the rate of 200kbps, total traffic sent by network was 4mbps, while it was only able to deliver at an average of 250kbps, a percentage of just 6.25% compared to wired LAN which delivered at almost 100%.

Therefore, from the simulation results obtained, it could be concluded that;

i. Under the same network condition, wired local area network will always perform better than the wireless LAN.
ii. When more traffic is being generated in the network, performance of wired LAN dropped to a considerable level (only able to deliver half of the packet sent) due to more utilization of the hubs, which leads to more collision and retransmission attempts in the network.
iii. Increasing traffic generation for the same number of nodes has little effects on the wireless network compared to the effect it has on the wired network, however, the wired network still delivers more packets than the wireless network.

X. SUMMARY, CONCLUSION AND RECOMMENDATION

A. Summary

This research work identified the importance of computer networks to human living and has therefore taken time to evaluate the performances of the network. The scope of the study was limited to the local area network, being the most available and widely used network around due to their simplicity and low cost of set up. Wired and wireless local area network were further considered for study.

The evaluation technique adopted was computer simulation as it was preferred to other evaluation techniques because of its simplicity in modelling real life events compared to mathematical analysis and also the unavailability of measuring tools contributed to shifting the ground in favour of simulation techniques. From the numerous types of simulators available, OPNET IT Guru was preferred above others for its user friendliness, simplicity and most of all for the availability of a free academic version.

Parameters and metrics are vital components of performance evaluation because they enable the system analyst to carry out a successful evaluation. For this study, the parameters used for evaluation were varying number of computer nodes and packet sizes, while the metrics measured include throughput, traffic generated, collision, latency and hub utilization.

For the purpose of the simulation, four different scenarios were modelled, these scenarios are;

i. Wired and wireless LAN with 20 nodes with packet generation at 200kbps.
ii. Wired and wireless LAN with 20 nodes with packet generation at 800kbps.
iii. Wired and wireless LAN with 50 nodes with packet generation at 200kbps.
iv. Wired and wireless LAN with 50 nodes with packet generation at 800kbps.

Each simulation was run for 20 seconds and data were obtained in a graphical format, analysis of the data was also carried out to the best of my knowledge for the purpose of comparing the two networks.

C. Conclusion

Based on the simulation results obtained and from the analysis of the results, conclusion was made that wired LAN performs better than wireless LAN when subjected to the same condition. This is because wired LAN has a dedicated path (cable) for transmitting data. This path, under normal condition is relatively free from interference. In the case of the wireless networks, transmitting medium is the air and this medium is subject to interference from other waves travelling through it.

However, the low performance of wireless LAN does not totally ruled it out of usage, it is still very useful, in fact, it is integral to the future of communication network, this is due to its mobility.

C. Recommendation

From this study, the following recommendations are made;

i. For optimum performance, network administrators should make use of wired networks.
ii. Traffic on network should not exceed the maximum threshold of the transmitting link, in order to get maximum throughput possible because more traffics result into more collision and this reduces the number of packets delivered.
iii. For the sake of mobility and maximum throughput, an hybrid of both wired and wireless networks could be made.
iv. Also for the purpose of accuracy in evaluation, this research work could be expanded to include other evaluation techniques, most preferably, measurement technique as this will help to validate the results obtained from this simulation.

REFERENCES

[1] Bansal et al, "Performance analysis of wired and wireless lan using soft computing techniques – a review," Global Journal of Computer Science and Technology Vol. 10 Issue 8 Ver.1.0, pp. 67-69, September 2010.
[2] Singh, A., "Throughput analysis of ethernet and fiber distributed data interface using OPNET IT Guru Academic Edition 9.1," An International Journal of Engineering Sciences ISSN: 2229-6913 Issue, Vol. 4, pp. 455-460, September 2011.
[3] Aguilera et al, "Evaluating wireless lan access methods in presence of transmission errors."
[4] Wazid et al, "Analysis of a lan under different ethernet wiring standards with variation in time and components," Proceedings of the International Conference on Pervasive Computing and Communication (PCC), Universal Association of Computer and Electronics Engineer, 2012.
[5] Tamar, M., "Performance analysis of wireless local area networks," Unpublished Masters Thesis submitted to Faculty of Engineering, Cairo University, 2000.
[6] Jain, R., "Art of computer systems performance analysis techniques for experimental design measurements simulation and modelling," Wiley Computer Publishing, John Wiley & Sons, Inc., 1991.
[7] Song J., and Trajkovic L., "Enhancements and performance evaluation of wireless local area networks, 2010.

Data Integrity through Integrated Database Trigger Model

Itai Yaya PhD
Information Technology Control Unit
FirstBank Nigeria, Marina
Lagos- Nigeria
yaya.itai@firstbanknigeria.com

Nicolae Goga PhD
Molecular Dynamics Group
University of Groningen & Politehnica University,
Netherlands
n.goga@rug.nl

Abstract- Inability to capture real values as a result of Data Manipulation Language operations has been of major concern to research community. Numerous studies have therefore attempted to employ traditional database monitoring technique in identifying limitations responsible for the disparity. Database integrity monitoring tools have been successfully applied in various areas of research. However, they could be the springboard for improving database integrity monitoring. This paper therefore aimed at database trigger techniques that track database modifications by recording all changes and access activities at the database level as well as recommending necessary actions for improvement. A multi-dimensional approach was employed. Literature review and in-depth interviews technique was employed in identifying the present level of database integrity monitoring automation in Nigeria financial institutions. A systematic random sampling technique was used to select some participants for this purpose. Event-Condition-Action (ECA) was employed in this paper with Requirement Management Life Cycle (RMLC). Results of the trigger techniques would be measured against best standard practice. Performance from triggers techniques was measured using Mark Gurry & Associates (MGA) EagleEye model. The ECA rules of this model was illustrated at the database level, which suggested that whether existing applications were modified or new ones are created were the rules, would be executed. In conclusion, relevant database monitoring exceptions handling techniques could be incorporated into database trigger techniques that would serve as a form valuable tool for tracking database modifications. This could be assisted by recording all changes and access activities at the database level as well as recommend necessary intervention strategies.

Keywords: Integrated database trigger model, Regression, structured query language, Trigger, Oracle.

I. INTRODUCTION

Database integrity is a fundamental component of information security. In its broadest use, data integrity refers to the accuracy and consistency of data stored in a database, data warehouse, data mart or other construct. The term Data Integrity is used to describe a state, a process or a function and is often used as a proxy for data quality. Data values are standardized according to a data model and/or data type. All characteristics of the data must be correct including business rules, relations, dates, definitions and lineage for data to be complete. Data integrity is imposed within a database when it is designed and is authenticated through the ongoing use of error checking and validation routines. As a process, data integrity verifies that data has remained unaltered in transit from creation to reception (Glynn, 2012). As a function related to security, data integrity service maintains information exactly as it was inputted, and is auditable to affirm its reliability (Baker, 2009). Data undergoes any number of operations in support of decision-making, such as capture, storage, retrieval, update and transfer. Data integrity can also be a performance measure during operations based on the detected error rate.

The integrity of data is guarded by specific business rules called data integrity constraints. The rules that apply in the real world to events, actions, manipulations, and situations are translated into constraints that apply to the data handled in information systems. Data integrity means data recorded in the system does not violate these constraints. Data integrity is about the internal, system- level consistency of data collection. Without data integrity, the quality of the data would be highly doubtful (Sandhu, 2005).

Government and industry regulations such as the U.S. Sarbanes-Oxley Act of 2002, Japan's HPB 517, Basel II, and the International Organization for Standardization's 27001 standard have created challenges for database integrity monitoring experts that seem insurmountable. These regulations require, among other things, the implementation and testing of organizational database integrity controls to ensure the accuracy and integrity of financial statements and other corporate data (Ashok, 2006).

Data must be kept free from modification or unauthorized disclosure to drive any number of mission-critical business processes with accuracy (Kroenke, 2002).

The need to continuously audit database modifications cannot be overemphasized. Several techniques exist for improving data integrity (Pankaj, 2005). However one of the problems of existing approaches is their limited support for online monitoring of Data Manipulation Language (DML) activities. The solutions that are currently offered by industry or academia cannot answer effectively the question of "who-did-what-when" as the current databases audit trail logs do not contain enough information to adequately capture the details of DML activities.

II. LITERATURE REVIEW

According to Mark (2011), database triggers are used to maintain complex data validation and integrity. A trigger can also be used to rollback SQL transactions performed with invalid or incomplete data, before being applied to the database. Triggers are used to maintain security at a higher level by restricting access to privileged users. Triggers are used to maintain complex check constraints. Teradata Database does provide column level check constraints, but it cannot help validate a column against a column in another table. CRM applications can use triggers to generate logs of their best customers when they make big purchases, or frequent travelers, or buyers of high-profit items. Referential integrity on a child table requires the referenced parent key column to be unique. Triggers are used as an alternative where business applications cannot guarantee that the referenced parent key column is unique. The number of business rules that were implemented using triggers is truly limited only by your imagination or, more appropriately, your business needs. Teradata experts introduces new trigger features to Teradata database V2R6 that includes user defined functions called by triggers, stored procedures called by triggers, and large object (LOB) access in triggers.

Smith (2011), illustrated IBM's trigger technology was used to monitor changes to the database and notify the administrator or security officer through an alert. The technique hasn't gained widespread use largely because of the overhead that triggers place on system resources. With IBM's newer high-performance processors, however, such work is of little consequence nowadays, according to KST.

A trigger is a user database exit program that is activated whenever an event occurs, such as mistaken data entry, sabotage, and other vicious data manipulations. Triggers have an unlimited set of capabilities and are nicely appropriate for security applications because they are embedded into the database and can't be bypassed. The nice thing about triggers is they can alert you to breaches in real time rather than finding out about a security event a day, week, or month later.

Security breaches, unless they involve the government, generally don't make the daily news. Yet they occur every day and cost businesses and consumers millions in losses.

They could be the result of an organized crime effort or simply of a disgruntled employee. Smith believes that its Database Trigger solution leverages any existing security implementation to make learning about a breach faster and easier and thus more preventable or at least addressable before major damage is done.

According to Pedro (2012), database monitoring is an observation and the regular record of an application activity. It is a permanent data collection service executing over many different aspects of the monitored application or system. Monitoring is to check the progress of the running applications, that is, a systematic observation with well-defined goals.

The data read from sensors are used to make informed decisions about what actions to take—by a human or by reaction software. After these actions were executed, the sensors are probed again and the data collected are analyzed to check if the previously actions improved the state of the system or if other actions must be taken. A monitoring which executes in parallel with the monitored application is an intrusive monitor, sharing the computational power which would be used only by the application. This demands the monitor to be lightweight, in order to interfere the minimum as possible in the global system.

A monitoring application will keep track of the changes caused by any other running software in the system or even other systems which may interact with the monitored one. This data is then collected and should be served, as a report for administrators or to the main application use, staying aware about the system state. For example, in a medical heart rate controller software would be interesting to monitor every action that the controller may take — increasing or decreasing the heart rate. If other medical devices interact with the controller, it is necessary to record their actions also, so any change in the heart rate software behavior is registered. If suddenly the heart rate in the patience starts decreasing, a report over the monitored data can help the doctor to understand the situation and correct the problem as soon as possible. The first perturbation caused by the monitoring process is the execution of its own extra instruction.

Sonali (2012) illustrated how the race between confidentiality, integrity and availability has put terrible pressure for the need of out-of-box thinking. In his Customer relationship management system, the

increase in crucially of information to the business has inflated the challenges in protecting and managing the data Bank.

The main aspect was to increase the accessibility of information without compromising of security measures. Distributing data is one of the ways to provide business continuity. Distributed database is an aggregation of different databases that are logically distributed and stored in a variety of outsourced database system that aspects client as a centralized application where each database application may involve different database technology and different computer architecture that contribute the execution of global query. With the more promotion of internet technologies, the major problem is to preserve the integrity of data.

According to Fergal (2012), data integrity verifies that data has remained unaltered in transit from creation to reception. As a state or condition, Data Integrity is a measure of the validity and fidelity of a data object. As a function related to security, a data integrity service maintains information exactly as it was inputted, and is auditable to affirm its reliability.

Data undergoes any number of operations in support of decision-making, such as capture, storage, retrieval, update and transfer. Data integrity can also be a performance measure during these operations based on the detected error rate. Data values are standardized according to a data model and/or data type. All characteristics of the data must be correct including business rules, relations, dates, definitions and lineage for data to be complete. Data integrity is imposed within a database when it is designed and is authenticated through the ongoing use of error checking and validation routines. Data must be kept free from corruption, modification or unauthorized disclosure to drive any number of mission-critical business processes with accuracy. Inaccuracies can occur either accidentally (e.g .through programming errors), or maliciously (e.g. through breaches or hacks). Software developers must also be concerned with data integrity. They can define integrity constraints to enforce business rules on data when entered into an application. Business rules specify conditions and relationships that must always be true, or must always be false. When a data integrity constraint is applied to a database table, all data in the table must conform to the corresponding rule.

Vikas (2013) defined database trigger as a stored procedure which functions when trying to amend the data in a table like inserting, deleting or updating data. It is a database object, executed automatically and is bound to a table. Fundamentally, triggers are classified into two types mainly (Instead of Trigger and After Trigger). He illustrated how insert, delete or update operations aligned with excessively intricate views to support natively through Instead of trigger.

III. THE NEED FOR DATA INTEGRITY

Data integrity is the assurance that data is consistent, correct, and accessible, as we move into the database world, data integrity comes in many different types: null rules, unique column values, primary key values, referential integrity rules and complex integrity checking . Without data integrity, we cannot even begin to be concerned with data quality since it assumes that we might not have all the data, or that we cannot access it either physically or logically, or that we can have no certitude of its condition. It becomes important to track logical data integrity as applications are implemented on architectures where data movement, operations or migrations might altar, truncate, delete or corrupt the data. This is particularly important in the context of systems of record. These are information storage systems that are considered the authoritative data source for a given piece of data or information. As a process, Data integrity verifies that data has remained unaltered in transit from creation to reception. As a state or condition, Data Integrity is a measure of the validity and fidelity of a data object. As a function related to security, a data integrity service maintains information exactly as it was inputted, and is auditable to affirm its reliability. Data undergoes any number of operations in support of decision-making, such as capture, storage, retrieval, update and transfer. Data integrity can also be a performance measure during these operations based on the detected error rate.

Data must be kept free from corruption, modification or unauthorized disclosure to drive any number of mission's critical business processes with accuracy. Inaccuracies can occur either accidentally (e.g. through programming errors), or maliciously (e.g. through breaches or hacks) .Businesses sometimes choose to use triggers over standard Oracle audit features because triggers offer the flexibility to tailor the audit transactions in the database to the unique business requirements. Using a trigger also presents the possibility to tune the performance of auditing activities and the option to store the audit trail in a user-defined table space to alleviate maintenance burdens on the SYSTEM tablespace.

A. Application of Triggers in data integrity

- Enforcement of Business Rules
- Channel Behavioral Monitoring
- Provide sophisticated auditing

- Automatic generation of derived column values
- Prevention of Invalid Transactions.
- Fine-grained access control security policies

IV. METHODOLOGY

A multi-dimensional approach was employed in this paper. Extensive literature review and in-depth interviews technique were employed in identifying the present level of database integrity monitoring automation in Nigeria financial institutions. A systematic random sampling technique was used to select 100 professionals from three financial institutions representing 85% of system control personnel in the three financial institutions. It was observed that the financial institutions are currently migrating from basic to evolving stage of database integrity monitoring.

The Dynamic Objective Model (DOM) was used as a baseline for the Integrated Database Trigger Model (IDTM). Also employed was the Event-Condition-Action (ECA) rule with integration to TRACE model and a comprehensive Requirement Management Life Cycle (RMLC). Benchmarks were used in comparing the implemented models: DML operations real value pulling, the burden on system overhead and the model building's speed. An online IDTM model and Real-Time Performance Diagnostic Model (RTPDM) were then designed based on the outcome. The IDTM was developed using Oracle plsql on DOM technology suite with ECA rule on Cheque Management database as a prototype.

Demographic questions relating to this study were properly profiled according to the respondents' areas of specialization and the organization they represent. In this papers three institutions were considered A, B and C, which aided in cross-tabulation and comparing subgroups to see how opinions vary between these groups

Table 4.1: Institution A

Total Number of male=30, Total number of female=10 Grand total =40, representing 75% and 25% respectively from data collected from **Institution A**			
Variable		Distributed to area of specialization	%(percentage of representation)
Area of specialization	Database Administrators	3	7.5
	IT Security Personnel	20	50
	Systems Auditors	10	25
	Financial Control	7	17.5

Table 4.2 Institution B

Total Number of male=20, Total number of female=10 Grand total =30, representing 66.7% and 33.3% respectively from data collected from **Institution B**			
Variable		Distributed to area of specialization	%(percentage of representation)
Area of specialization	Database Administrators	2	6.7
	IT Security Personnel	18	60
	Systems Auditors	5	16.7
	Financial Control	5	16.7

Table 4.3 Institution C

Total Number of male=10, Total number of female=20 Grand total =30, representing 33.3% and 66.7% respectively from data collected from **Institution C**			
Variable		Distributed to area of specialization	%(percentage of representation)
Area of specialization	Database Administrators	3	6.7
	IT Security Personnel	17	60
	Systems Auditors	6	16.7
	Financial Control	4	16.7

This paper requires a technology model that has predefined trigger concepts that can be easily modified to suite users' needs. The DOM design has a structure that best suitable for handling changes [Riehle Et al 2000]. It will allow trigger and its components (Event, Condition, Action) to be changed at runtime.

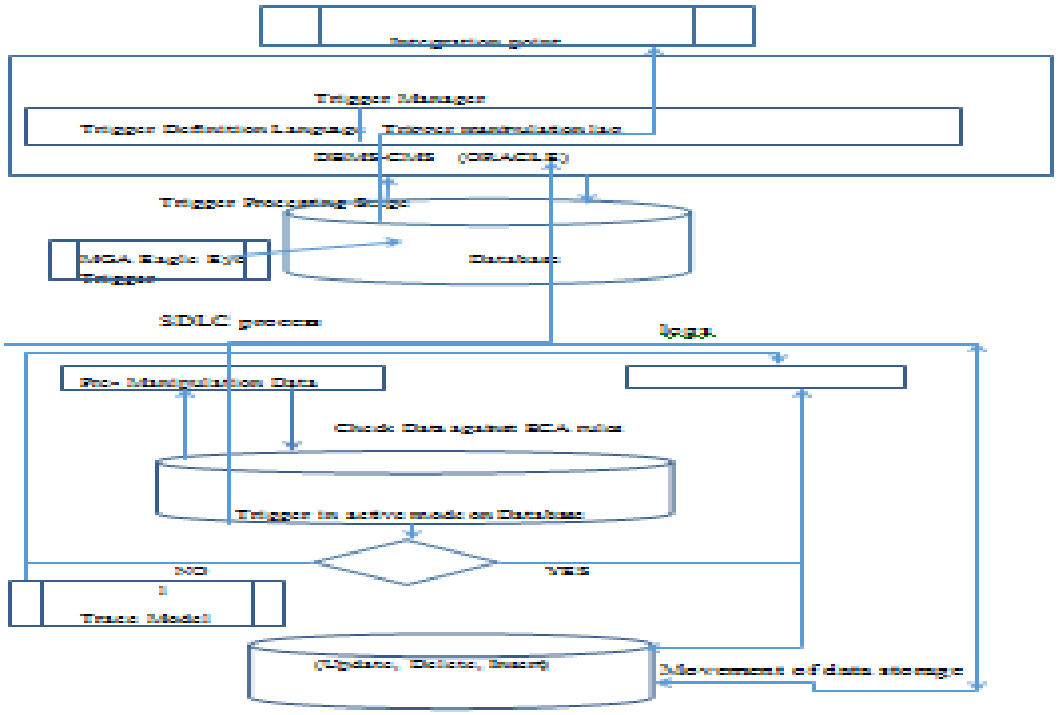


Figure 4.1: The Model System Architecture(2014)

The integrated database trigger model is based on the fact that action taken to ensure data integrity depends on business rule that the data need to satisfy. It starts with preprocessed (original) database manipulation with a variety of exceptions. Database manipulations is checked against predefined triggers (in figure 4.1) in the database with the objective of obtaining correct and consistent data, and then are loaded into data exception database store. The effectiveness of the proposed model will depends on the triggers (rules) defined in the database.

The technical process definition build around model involves:

i. Establishing business rules for the model (Event)

ii. Specification of acceptance criteria (Condition)

iii. Transformation rules (Action)

Normally a trigger is defined in the Event-Condition-Action (ECA rule) fashion. Given this structure of a trigger language, a simple trigger can be generated automatically when supplied with necessary attributes. An event part of the model (a trigger)was used to perform operations such as updating, inserting, loading, etc. Complex predicates, functions or checks are contained in the condition part for checking errors, duplicate data, inconsistency, missing values, etc. It is necessary to specify the condition criteria for accepting that values are outside the exceptions band for the specified purpose.

The condition criteria can apply to individual values as well as set of values. The condition criterion helps to identify the problem or problems. Once identification has been made then some transformation rules have to be implemented. The transformation rule specifies the action to be taken to transform data. This could be omitting a record, substituting a value, confirmation operations such as sending message alerts or other database operations, etc. The action part of a trigger will be activated when the condition is satisfied. The approach in this model consists of four stages namely Business Rule, Database predefined trigger rule, exception database store, End-User Reporting Web-view the trigger processing stages. The database defined trigger processing stage is for creation and manipulation of triggers and their components.

V. Experimental Analysis (IDTM Model)

The expected output from this study includes optimal models (IDTM and RTPDM) that addressed the inability to capture real values during DML operations and the resultant impact of database performance. The resultant output from the developed IDTM model was measured, analyzed against the existing systems Omohanmansm (2010) and Somkefe (2012) work, using Oracle toad analyzer automatic database diagnostic monitor (ADDM) on the RTPDM. Regression performance evaluation was carried out and results compared with

other existing systems using RTPDM model. End user (Financial Institutions) performance evaluations were also carried out using a structured statistical questionnaire feedback mechanism.

The Experiment shows how a wide range of security requirements for a database system for cheque management application is archived by building application specific controls on a suitable trigger model. The technique was illustrated using IDTM model build around critical tables in the cheque management system database. The requirements of this application cover most aspects of security, including confidentiality, integrity and audit trail. The research findings were also compared with similar researches Omohanmansm (2010), Somkefe (2012) including popular third-party event management systems DB Protect. The issue of real value tracking and database performance were adequately addressed in this work by IDTM and RTPDM models.

In order to illustrate this comparison, Omohanmansm (2010) work, Somkefe (2012) work and were subjected to the same data criteria (Data migration set of 2002, 2005 and 2013 from a Financial in Nigeria) and results compared with IDTM model with same data sample from CMSDB DML activities (Insert, update and delete). All the Schemas on IDTM model were considered with particular reference to some selected tables (ChequeNumber, ChequeType, SecurityNumber, and SecurityChequeNumber).

A. Result

The Experiment had shown how a wide range of security requirements for a database can be met by building application specific controls on a suitable trigger mechanism. The technique was illustrated using with a detailed trigger model on Cheque management system database. The requirements of this application cover most aspects of security, including confidentiality, integrity and audit.

Table 5.1Original SecurityChequeNumber

ChequeNumber	CusName	SecurityNumber	Operator_id	PrintStatus	PrintDate	CaptureDate	CusAcctNo
00000001	Wunti Raski	80881234	Sn01100	Y	15-03-2013	12-03-2013	00005656 48
00000002	Olufemi Ojo	97881234	Sn02400	Y	18-03-2013	11-03-2013	01238647 65
00000111	Yekinni Ola	00881234	Sn02401	N	19-03-2013	12-03-2013	9871 2354 12

When an update SQL is fired on a cheque Number filed on the Original SecurityChequeNumber table, it inserts the updated new records in the 'inserted' logical tables. The old records, i.e. records before updating are entered into the trigger table as shown below**.**

Table 5.2 Original SecurityChequeNumber

Cheque Number (old)	Cus Name	Security Number	Updated by	Updated date	ChequeNumber(new)
00000001	Wunti Raski	80881234	Aderonke Adeshina	18-03-2013	57710856

In table 5.2 above the ability to report before aspect of the DML operation in the cheque management systems database has been completely addressed. The CMC database health was examined for the most expensive sessions, and found to be normal using oracle toad utilities. An oracle third-party tool MGA's Eagleeye was used to perform regression testing to get immediate of any change in the cheque management system database, by performing load testing as well as tracking the response times that the production users are receiving to serve as the service level agreement.

6.0 Conclusion

This paper is essential for regulatory monitoring, because it provides a continuous and permanent audit trail of data access and changes, while storing this information in a centralized repository archived for long-term retention. The information gathered by the implemented integrated database trigger model help improve an organization's operational performance by identifying data-use patterns leading to increased IT efficiency and the fine-tuning of existing business processes.

The paper provide the most convenient and comprehensive way to track modifications by recording all changes and access activities at the database level.

The Integrated Database Trigger Model (IDTM) also maintains information integrity to reduce organizational risk and ensure regulatory compliance. Data integrity is an important part of an organization's daily operations and technology has a vital part to play in ensuring adequacy of data integrity. The major gaps identified in the closely related work reviewed such as: inability to capture real value, absence of exception handling and performance issue around the database has been addressed. With the rising complexity of data integrity challenges, database triggers techniques go a long way to improve the effectiveness of data integrity issues. The use of the various database

trigger techniques in IDTM model, such as prevention of invalid transactions, financial business rule enforcement, sophisticated auditing and automatic column values derivation yielded positive result. Database trigger therefore has an important role to play in the enforcement of data integrity that would be capable of standing up to data integrity challenges we face in our operating environment today. This paper is a possible step forward towards improving data integrity in Nigeria financial institutions.

REFERENCES

[1] Alex, J and Thibault, W. (2002). Scalable, secure and highly available file access in a distributed workstation environment. Distributed Computing Journal vol. 4 pp. 78.

[2] Ashok, G. (2006). Database integrity controls and accuracy of financial statements and other corporate data. Journal of Computers & Technology Volume 3 No. 1.

[3] Arzberger, P. and Schroeder, P.(2004). An International Framework to Promote Access to Data. Science, 303: 1777-8.

[4] Caruso, F., Cochinwala, M., Ganapathy, U., Lalk, G. and Ceri, S. ; Cochrane, R. J. & Widom, J. (2000). Practical Applications of Constraints Triggers: Successes and Lingering Issues. Proceedings of 26th.VLDB Conference, Cairo-Egypt VLDB Vol. 34 pp 90-99.

[5] Chandrasekaran, S. & Kehoe, B. (2003). Technical comparison of Oracle real application clusters IBM DB2 UDB ESE Vol.71 IDUESE 230-238.

[6] Chen, P. (1976). The Entity-Relationship Model: Toward a Unified View of Data. ACM Transactions on Database Systems 9-36.

[7] Cohn, D., and Hull, R. (2009) Business artifacts: A data-centric approach to modeling business operations and processes. IEEE Bulletin on Data Engineering. 32, 3-9.

[8] Craig, M. and Steve, C. (2004). Collaborative computing and Active Database Systems:Triggers and Rules for Advanced Database Processing, College & Research Libraries 67, 1, 5-14.

[9] Crossman, P. (2007). The Less-Than-Obvious Costs of Ignoring Data Architectures .Wall Street & Technology vol. 101.

[10] Cynthia, D.(2008). Differential privacy: A survey of results. In Proceedings of the 5th International Conference onTheory and Applications ofModels of Computation pages 1–19, Xi'an, China.

[11] Calvanese, G. G., and Lenzerini, M. (2004). Data integration under integrity constraints. Inf. Syst., 29(2):147–163.

[12] Deutsch, A .and Tannen, V. (2005). Xml queries and constraints, containment and reformulation. Theoretical Computer Science, 336(1):57–87.

[13] Dai J., Zhang, Y., Li X., and Luo S.(2011) A completeness and freshness guarantee scheme for outsourced database," on Second International Conference on Networking and Distributed Computing, 2011, pp.201-205.

[14] Davide M. (2005) Advanced Techniques for Efficient Data Integrity Checking. IEEE Transactions on Knowledge & Data Engineering, vol. 24(12).

[15] David, H. O., Olga, Y. and Nicole, F. (2002). Enhancing database integrated and process automation through smarts triggers , IACI 2002 490-496

[16] Kung, D. (2004). A generalized framework for monitoring changes to database relations using active capability Data and Knowledge Engineering, vol. 14, no. 10, pp. 1\

[17] Daza, V. and Herranz, J.(2002) .A Distributed and.Computationally Secure Key Distribution Scheme. ISC'02. LNCS 2433

[18] Deutsch, A. (2009). FOL Modeling of Integrity Constraints (Dependencies).Encyclopedia of Database Systems, 1155-1161.

[19] Dinn, A., Paton, N., and Williams ,M. H.(1999). Active Rule Analysis and Optimization in the Rock &Roll Deductive Object-Oriented Databaset." Information Systems (Volume 24 Number 4) page 327-353.

[20] Don, S. (2009). DML Triggers in SQL Server. Communications of the ACM,34(10):94–109.

[21] Dong, S. and Dan, L. (2010).study on method for data Confidentiality and data integrity in relational Database," in 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT),vol. (1)pp.292-295.

[22] Eddie, K. (2008). Hot crap! In Proceedings of the Workshop on Organizing Workshops, Conferences, and Symposia for Computer Systems (wowcs'08), San Francisco, ca, usa.

[23] Ezekiel, K. and Marir, F. (2006). A conceptual model for managing knowledge Represented as triggers in active databases": 4th European Conference on Knowledge Management, pp. 323-333

[24] Fan, W. and Geerts, F. (2008). A Revival of Integrity Constraints for Data Cleaning." *VLDB*, 2008: 1(2): 1522-1523.

[25] Frank, M. (2009). Privacy integrated queries: An extensible platform for privacy-preserving data analysis. In Proceedings of the 35th acm sigmod International Conference on Management of Data (sigmod'09), pages 19–30, Providence, ri, USA.

[26] Fritz, C., Hull, R. and Su, J.(2009). Automatic construction of simple artifact-based business processes. In ICDT, 225-238, 2009.

[27] Geerts, F. (2008). Foundations of Data Quality Management. Synthesis Lectures.... PVLDB 1(2):1460-1463.

[28] Glynn, F. (2012). Fundamental of database integrity integration. TACI conference vol. 10,pp

[29] Hoffman, T. (2007). Oracle performance tuning, Database management systems (2nd ed.). O'Reilly & Associates.

[30] [30]Christiansen H. and Martinenghi, D. (2006). On simplification of database integrity constraints. Fundamental Informaticae, 71:371–417.

[31] Decker, H. 2002 .Translating advanced integrity checking technology to SQL. In J. H. Doorn and L. C. Rivero, editors, Database integrity: challenges and solutions, pages 203–249. Idea Group Publishing.

[32] Indrajit, R., Donald P., Bond, K., and Emmett, W. (2009). Laminar: Practical fine-grained decentralized information flow control. In Proceedings of the 2009 ACM sigplan Conference on Programming Language Design and Implementation (pldi'09), pages 63–74.

[33] Jed, L. and Andrew, M. (2005). Defining and enforcing referential security. In submission, April 2012. Peng Li and Steve Zdancewic. Practical information-flow control in web-based information systems. In Proceedings of the 18th IEEE Computer Security Foundations Workshop, pages 2–15.

[34] Jepson, K. (2006). Data Security Getting Better All the Time, Experts Say. Credit Union Journal vol. pp78

[35] Jean, Y., Kuat, Y. and Armando, S. (2012). A language for automatically enforcing privacy policies. In Proceedings of the 39th acm sigplan-sigact Symposium on Principles of Programming Languages (popl'12), pages 85–96.

[36] Mallur, V. (2011). A Model for Managing Data Integrity. Journal of Computer science uottawa vol. 34, pp77-88. https://www.ruor.uottawa.ca/en/.../ Mallur_Vikram_2011.

[37] Martinenghi, D., Christiansen, H., and Decker, H. (2006). Integrity checking and maintenance in relational and deductive databases – and beyond. In Z. Ma, editor, Intelligent Databases: Technologies and Applications, pages 238–285.

[38] Matthew, S. M. (2009). Data integrity issues in relation to environmental and ecological.` Journal Ecological sciences vol. 16 pp. 72-80

[39] Martin, L., Matthias, H. and Erich, N. (2009). GMD – German National Research Center for Information Technology Volume 48.

[40] Martin, S. (1997) .Active Database Management Systems for Monitoring and Control, Journal of Electrical computing systems Vol. 35 pp. 45

[41] Mark, H. and Fred, K. (2011). Both of Teradata solutions group Terada Journal, Inc. vol. 3 pp. 56-67

[42] Maxwell, K. and Alexander, Y. (2007). Information flow control for standard OS abstractions. In Proceedings of the 21st ACM Symposium on Operating Systems Principles (sosp'07), pages 321–334, Stevenson, wa, USA.

[43] McCarthy, W.E. (2002). Resource-Event-Agent (REA) Model, Journal REA Vol.9 pp. 715.www.cirworld.ijssronline.

[44] Mike, C. (2012), Adoption of database normalization Basics Journal of Privacy and Confidentiality vol. 19 pp, 49,

[45] Nickolai, Z. and David, M. (2008). Securing distributed systems with information flow control. In Proceedings of the 5th usenix Symposium on Networked Systems Design and Implementation (nsdi'08), pages 293–308, San Francisco, CA, USA.

[46] Omohanmansm, T .Y (2010). Integrity constraints in a transaction database using triggers, CMM Conf. on Management of Data, pp.212

[47] Pankaj, R. (2005) i. Side-channel attacks. In Handbook of Information Security, volume 3, pages 241– 260.Wiley, 2005.

[48] Pathak, S.(2012), Integrity Checking Mechanism with Effective and Realistic Approach in Distributed Database. International Journal of Scientific and Research Publications, Volume Issue 6, June 2012 1 ISSN 2250-3153.

[49] Ross, R.G. (1999). Exploring Business Rules. Business Rule Book, Second Edition vol. 2.

[50] [50]Riehle D. (2000). Dynamic Object Model, In Proceedings of the 2000 Conference on Pattern Languages of Programming – PloP vol.9

[51] Shulman, F. (2008). Database trigger operations and content to identify subjectively interesting web usage patterns", ACM Transactions on Internet Technology (TOIT), pp.100-116

[52] Somkehfe O.Y (2012) Declarative Database triggers: Language, Automating triggers,

Proceedings of the 27th International Conference on Information Security, Pages: 71-79

[53] Steve V. and Petros, E. (2007). Labels and event processes in the Asbestos operating system. ACM Transactions on Computer Systems, 25(4).

[54] Steven, H. B. (2007). The importance maintaining data integrity and availability. The EMBO Journal Online Volume 26, Issue 2, pages 448–458

[55] Steve, C. (2004), the trigger-happy DBA series on data manipulation language. www.databasejournal.com/.../The-Trigger-Happy- DBA---Part-2.htm 29, 2010.

[56] Thibault, D. (2003), Technical comparison of Oracle real application Clusters IBM DB2 UDB ESE Journal of Computer science vol. 7 pp 78.

[57] Trillium, W. (2000). Trillium Software System for data warehousing and ERP," Trillium Software,Webpage, Date Accessed: 01/15/2000, http://www.trilliumsoft.com/products.htm, 2000.

[58] Türker, C., Gertz, M. (2001). Semantic integrity support in SQL-99 and commercial (Object-) relational database management systems pp 78.

[59] Vijayan, J. (2007). Six Ways To Stop Data Leaks. Computerworld Protecting Data becomes Top Security Priority for IT. Vol. 4

[60] Vikram, M. (2011). Database integrity constraints and design. https://www.ruor.uottawa.ca/en/.../Mallur_Vikram_2011_thesis.pdf

Strategies for Effective Information Security Management in Nigerian Universities

Babalola Y.T.
Department of Information Resources Management
Babcock University
Ilishan-Remo, Ogun state, Nigeria
yemisiolutayo@yahoo.com

Soyemi O.D.
Department of Information Resources management
Babcock University
Ilishan-Remo, Ogun state, Nigeria
yemshod1404@yahoo.com

***Abstract*--**Information systems are indispensable to virtually every personal and organizational activity, hardly can any organization function without them today; they are however subject to abuse in the hands of unscrupulous users. Information security problem has become a global phenomenon with the pervasive use of information systems across national and international networks. The situation is no better in higher education institutions especially in a developing country like Nigeria. In order to ensure the Confidentiality, Integrity and Availability (CIA) of information, information security experts have employed various technical solutions. However, the rising trend of information technology misuse suggests the need for a more comprehensive approach to effective information security. Specifically, the human factor seems to be a strong variable in determining effectiveness of information security measures. This paper highlights various forms of information system abuse, their consequences for organizations and the reasons for information system misuse as found in the literature. The paper suggests various non-technical solutions that could be combined with technical measures to combat information security problems in Nigerian Universities.

Keywords— Information security, information systems, information security policy, Nigerian Universities

I. INTRODUCTION

Information systems play pivotal roles in virtually all human endeavours as they facilitate access to information and support transactions across local and global networks. Various organizations, whether public or private, profit or not-for-profit are investing substantially in information and communication technologies having realized that their ability to compete locally and internationally depends, to a large extent, on efficient use of information and communication technologies. For business organizations, the need to increase the productivity and efficiency of their employees, and create innovative products and services that can improve their overall business performance are the impetus for their technological investments. Academic institutions are equally employing information systems not only to aid teaching, learning and research, but to extend access to higher education and increase opportunities for self development through e-learning and the Massive Open Online Courses (MOOCs) [1]. Presently, almost all academic and administrative activities in most universities, including student admission, registration, teaching and examination are internet-based.

The pervasive use of information systems in higher academic institutions implies that the volumes of personal and corporate data being trafficked over computer networks are vulnerable to unlawful access, manipulations and misuse [2]. This situation is worsened by the fact that students, who generally constitute the largest population of users could be exuberant and carefree in their management of personal data such as passwords and in the use of copyrighted materials and social network sites. Besides, students tend to patronize websites that offer free music and software downloads more than any other category of users and are more likely to fall victims of virus and malware sites. This however does not preclude the fact that faculty and staff can pose as much security risks either because of ignorance or nonchalance.

There are evidences to prove that without proper management, information systems can become deadly weapons in the hands of unscrupulous individuals who might use them as a threat to security of personal, organizational and national information. According to Sodiya, Ibrahim and Ajayi [3], eighty-two percent of organizations in Nigeria still

experience computer intrusion. Misuse of information systems is however not a national challenge but a global phenomenon with a global financial burden over one trillion dollars every year [4]. This paper highlights various forms of information system abuse, their consequences for organizations and the reasons for information system misuse as found in the literature. The paper suggests various non-technical solutions that could be combined with technical measures to combat information security problems in Nigerian Universities.

II. FORMS OF INFORMATION SYSTEM MISUSE

Various forms of information system misuse have been observed in the literature; security breaches such as virus attacks, unauthorized access to systems, theft of confidential information, system sabotage, financial fraud, denial of service attacks are on the increase in various types of organizations [5]. Apart from financial losses, academic institutions also run the risk of litigations, reputational damage, low webometric ranking, and loss of vital data such as students or staff records and research data [2,6, 7]. They are thereby compelled to spend a lot of money on securing their information systems from internal and external assaults.

Information security refers to 'the processes and methodologies which are designed and implemented to protect print, electronic, or any other form of confidential, private and sensitive information or data from unauthorized access, use, misuse, disclosure, destruction, modification, or disruption' [8]. Whitman and Mattord cited by Makhudu, Mavetera and Mavetera [9] viewed information security as the protection of information and the hardware and software that are used to store and manage it. Information security is considered in this sense in this paper.

Information security is assessed based on three major parameters- Confidentiality, Integrity and Availability. Confidentiality means that only those who should have access to particular information have access to them; Integrity means that information is without error or unauthorized modifications while Availability implies guaranteed access to information and information resources whenever they are needed. In order to ensure the CIA of information, organizations have explored various technical measures including authentication, encryption, firewall, biometric system, e-mail filtering, patch management, URL filtering, log management software, forensic software, intrusion detection software anti-virus and anti-spy software, surveillance technology etc. The rising trend of information technology misuse however suggests the need for a more comprehensive approach to effective information security management. The human factor seems to be critical in determining the effectiveness of information security measures. Insiders are particularly indicted because of their ability to circumvent established procedures with valid usernames and passwords and gain access to information and information system without being detected easily [10].

III. REASONS FOR INFORMATION SYSTEM MISUSE

Several reasons have been suggested for information system misuse; among them are:

A. Lack of Information Security Policy

An information policy is an official document that guides the use of information and information technologies in an organization. An information security policy should state unambiguously the kinds of behaviour and usage that are acceptable and those that are not as well as the penalties for non-compliant behaviours. In line with the Biblical principle which states that 'where there is no sin, there is no law', absence of a written information security policy means users are at liberty to do what they please; and the management has no moral or legal basis on which to punish or reprimand them. This has been the situation in most organizations and institutions of higher learning studied in Nigeria [8].

It is the responsibility of the management to provide information security policy as evidence of their commitment to information security [11]. Unfortunately, management support in this regard is often lacking and the information technology department is saddled with the responsibility of drafting and enforcing compliance with information technology policies. Empirical evidences have shown that availability of organizational information security policy can reduce incidences of security breaches and increase users' intention to comply with information security policies [12,13]. Therefore, there have been much emphasis on the need to adopt a comprehensive, organization-wide approach to information security management [14,6] and also explore the non-technical measures such as provision of information policy [13].

B. Lack of awareness of Information Security Policy by Users

In contrast to the many reports that have indicated a positive relationship between availability of information security policy and reduction in rate of security breaches, studies like that of Wiant [15] have reported no association between the two. This

inconsistencies has led to the conclusion that availability of information security policy in itself cannot prevent misuse of information systems, if users are not aware of its existence. An information security policy that is not publicized is non-existent. Unfortunately, several studies within and outside Nigeria report low level of awareness of information security policy in various organizations. Sodiya, Ibrahim, and Ajayi [3] reported that 82% of academic staff of government universities in south-west Nigeria were not aware of the existence of information security policy in their institutions. A more recent study conducted in South Africa showed that most of the respondents were not aware of the information security policies in their organizations and also admitted absence of security awareness programmes in their organizations [9].

C. Lack of Information Security Training

This is another prominent cause for misuse of information systems [16, 17). Information security training program is an on-going effort to educate users about the information security risks in their organization, their responsibilities as individuals and the sanctions that await information security violations [18]. Information security training involves creating awareness of a broad range of security issues including organizational security policy, and raising users' consciousness about security matters either through briefing, seminars, posters and conversations. Information security training however goes beyond creating awareness, it is an in-depth study of specific information security issues which could be delivered periodically as refresher courses either through the classroom or online platform [19, 20].

Information security training emphasizes an organization's commitment to security policy and empowers users to comply with the policies. Unfortunately, the findings of a lot of study carried out within and outside Nigeria indicate that information security training is not given due attention [21].

D. Disgruntled Employees

Lack of knowledge and basic skills as discussed above are often responsible for inadvertent breach of information security rules. However, intentional misuse of information system are not perpetrated by novices but by disgruntled expert users. According to Straub [22] employees, who have an axe to grind with their organizations either often take their anger out on the system. From two national studies, Stanton, Mastrangelo, Stam and Jolton [23] found that job satisfaction and organizational commitment are among the factors that determine information security behaviour of employees. Lack of job satisfaction naturally reduces organizational commitment which might lead to careless or carefree attitude towards information security issues. Besides, adhering to information system policies requires some efforts on the part of users and could be sometimes inconvenient. Employees who are dissatisfied with their jobs may not be willing to exert such efforts in obeying information security rules. Incidentally, insiders accounts for up to 30% of information security breaches in organizations [5].

IV. CONCLUSION AND RECOMMENDATIONS

The effectiveness of any system depends to a large extent on its users. In the same vein, the foregoing has shown that any effective information security strategy must take the human factor seriously. Incidentally, technical measures alone might not be adequate in addressing this problem. In order to enhance the effectiveness of technical information security measures in Nigerian higher education institutions, this paper recommends that:

- the Nigerian University Commission (NUC) mandate all Nigerian Universities to develop a comprehensive information security policy.
- here should be periodic awareness and training programs for both staff and students in the various institutions and finally,
- University management should identify factors that could contribute job dissatisfaction and lack of organizational commitment of their employee so they can be motivated to comply with institutional information security policy.

REFERENCES

[1] L. Paas and H. Creech, "How information and communications technologies can support education for sustainable development: Current uses and trends," International Institute for Sustainable Development," (IISD), 2008.

[2] M. Jones, "An evaluation of privacy and security issues at a small university," The Technology Interface Journal 10 (2), 2009.

[3] A. S. Sodiya, S.A. Ibrahim, and O. B. Ajayi, 'The state of information security in south-western Nigerian educational institutions," Proceedings of the international conference on Science and national development 25th-28, Oct 2004; 38-45, 2004.

[4] Q, Hu, Z. Xu, T. Dinev and H. Ling, "Methods for evaluating and effectively managing the security behavior of employees. Does deterrence work in reducing information security policy abuse by employees?" Communications of the ACM 54 (6), 54-60

[5] R.P. Hartwig and C. Wilkinson, "Cyber risks: The growing threat," Insurance information institute, 2014.

[6] A.A. Amarachi, S. O. Okolie and C. Ajaegbu, "Information security management system: Emerging issues and prospect," IOSR Journal of Computer Engineering 12 (3), 96-102, 2013.

[7] T. Oguntunde, "Challenges of managing higher academic institution information and communication technology (ICT) unit from cradle: The University of Ibadan experience," African Journal of Computing & ICT 5 (2), 25-30, 2012.

[8] A. Okunoye, L.A. Adebimpe, A. Omilabu, I.O. Olapeju and O.B. Longe, "Information security awareness among smes in the south western Nigeria – Significance of factors," African Journal of Computing & ICT 5 (5) 184-193, 2012.

[9] A. B. Makhudu, N. Mavetera and C. G. Mavetera, "Investigating information system security policy and awareness training programs in South African organizations," Innovation Vision 2020: Sustainable growth, Entrepreneurship, and Economic Development 1863-1875, 2012.

[10] M. Warkentin, M. McBride, L. Carter, and A. Johnston, "The role of individual characteristics on insider abuse intentions," Proceedings of the Eighteenth Americas Conference on Information Systems, Seattle, Washington, August 9-12, 2012.

[11] K. Holme and B. K. Slovang, "Information security policy," Computers & security Vol.21 (5), 402-409, 2002.

[12] O. Adedara, M.T. Karatu, and A. Lagunju, "Employee's conformity to information security policies In Nigerian business organisations: The case of Data Engineering Services PLC," Computing, Information Systems & Development Informatics Journal. 3 (2) 43-50, 2012.

[13] A. Imam and M.S. Hammoud, "The impact of non-technical security management factors on information security management in health informatics," International Journal of Information Technology and Business Management 26 (1), 13-28, 2014.

[14] J. Choobineh, G. Dhillon, M.R. Grimaila and J. Rees, "Management of information security: Challenges and research directions," Communications of the Association for Information Systems 20 (57), 2007.

[15] T. L. Wiant, "Policy and its impact on medical record security," Unpublished doctoral dissertation, University of Kentucky, Lexington, 2003.

[16] J. D'Arcy, A. Hovav and D. Galletta, "User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach," Information Systems Research Articles in Advance. 20 (1), 79-98, 2008.

[17] I. Kirlappos, A. Beautement and M.A. Sasse, "Comply or die is dead: Long live security-aware principal agents," Paper presented at the FC 2013 Workshops, USEC and WAHC 2013, Okinawa, Japan, April 1, pp.70-82, 2013.

[18] D.W. Straub and R.J. Welke, "Coping with systems risk: Security planning models for management decision-making," MIS Quarterly 22 (4), 441-469, 1998.

[19] F.J. Haeussinger and J.J. Kranz, "Information security awareness: Its antecedents and mediating effects on security compliant behavior," Thirty Fourth International Conference on Information Systems, Milan, 2013.

[20] IsecT Ltd., "Information security policy, awareness and compliance manager: Role of justification, job description and candidate specification," 2009.

[21] M., Kimwele, W. Mwangi and S. Kimani, "Adoption of information technology security policies: Case study of Kenyan small and medium enterprises (SMES)," Journal of Theoretical and Applied Information Technology, 18 (2), 413-416, 2010.

[22] D.W. Straub, "Effective IS security: An empirical study," Information Systems Research Vol 1 pp. 255-276, 1990.

[23] J.M. Stanton, P.R. Mastrangelo, K.R. Stam and J. Jolton, "Behavioral information security: Two end user survey studies of motivation and security practices," Proceedings of the tenth Americas conference on information systems, New York, August 1-7, 2004.

ICT and Students' Learning Effectiveness

Amanze Ruth C.
Information Technology Development Services
Babcock University
amanzer@gmail.com

Kuyoro 'Shade O.
Department of Computer Science
Babcock University
afolashadeng@gmail.com

***Abstract*— Information and Communications Technologies (ICTs) are diverse set of technological tools and resources used to communicate, create, disseminate, store, and manage information. ICT has played great role in both formal and non-formal educational settings; in programs provided by governmental agencies, public and private educational institutions, profit and non-profit corporations, and secular and religious communities. The shift towards the ICT-based education in the last two decades is said to have revolutionized and revitalized the education sector , such that giant strides are being taken towards technological progress. The focus of this paper is to address the issues relating ICT to teaching/learning effectiveness and the various ways of its application in the educational system.**

***Keywords*— Information Communication Technology; Education; Effectiveness; Tools; e-learning**

VII. INTRODUCTION

Information and Communication Technology (ICT) is defined as a diverse set of technological tools and resources used to communicate, create, disseminate, store, and manage information, which comprise internet service provision, telecommunications equipment and services, information technology equipment and services, media and broadcasting, libraries and documentation centres, commercial information provides network-based information services, and other related information and communication activities.[1][2]

In early part of history, smoke and drums and town cry, carved messages on stones, and writing were some of the various signals of communication in Africa. In the quest for means of improving communication and the need for message to travel long distance; discoveries such as telegraph line, telephone, radio, television, computer, and recently the internet-which turn the entire world into a global village- were made. Today email, teleconferencing, collaborations, unified communication have become popular methods of communication.[3]

The development of ICTs brought about a major shift in the world. Whether one discusses the emergence of global financial systems, growing citizen solidarity networks, or teleconferencing, one thing remains common and is at the core of the new society, the solicitation and exchange of the world's most valuable resource: information. Thus, defined as the new social morphology of our society, ICT is both a structure and a process that enables the exchange, the redirection, and the reception of information, on a global scale, without restraints of space or time. Distance is rendered irrelevant, allowing direct, simultaneous, decentralized, and expanding relations of collaboration, advocacy, trade, production, and innovation, generating new forms of power constellation and distribution.[4]

ICTs have become, within a very short time, one of the basic building blocks and cause for re-engineering the modern society in the recent years.[5] Educational institutions have begun to realize the importance of adoption and integration of ICT in their teaching/learning process.[6] ICT, according to [7] has a significant impact on the way teachers and students communicate. Communication in the learning process provides a transfer of information between teachers and students. Many institutions are now heavily investing in ICT infrastructure and aim to exploit it to the maximum extent possible.

The recent development in education embraces the use of ICTs which have undoubtedly affected teaching, learning, and research.[8] This adoption has been divided into two broad categories: ICTs for Education and ICTs in Education. ICTs for education refers to the development of information and communications technology specifically for teaching/learning purposes, while the ICTs in education involves the adoption of general components of information and communication technologies in the teaching learning process. The role of ICT in education is becoming more and more important and this importance will continue to grow and develop as the world move more rapidly into advanced digital media and information. ICTs are swiftly evolving technologies, and so even the most ICT fluent teachers need to continuously upgrade their teaching method and at the same time keep abreast of the latest developments and best practices in their field.[9]

This article addresses the issues relating to ICT roles in teaching/learning effectiveness and the various ways of ICT's application in the educational system. The remaining part of this article is arranged as follows section 2.0 addresses ICT in teaching/learning. Section 3.0 describes the e-learning. Section 4.0 highlights the importance of ICT in Education sector. Section 5.0 identified the advantages and disadvantages of ICT to teaching/learning while Section 6.0 gives the conclusion.

VIII. ICT IN TEACHING/LEARNING

The importance of ICT is quite evidence from the educational perspective. Though the chalkboard, textbooks, radio/television and film have been used for educational

purpose over the years, none has quite impacted on the educational process like the computer. While television and film impact only on the audio-visual faculties of users, the computer is capable of activating the senses of sight, hearing and touch of the users. ICT has the capacity to provide higher interactive potential for users to develop their individual, intellectual and creative ability. Today, computers perform a host of functions in teaching/learning as many nations are adding computer literacy, reading and writing literacy as skills students will need for succeeding in a technologically developed world.[10]

ICT is used for educational purposes, as a collaborative tool, e-mail, chatting, and as means of communication in schools.[11] It helps teachers in connecting with students in their deliveries of teaching that suits students' specific needs. Digital technology offers teachers the flexibility to personalise learning and to engage students with different learning styles.

There is no doubt that ICT provides productive teaching and learning in order to increase creative and intellectual resources especially in modern information society. Through the simultaneous use of audio, text, multicolor images, graphics, motion, ICT gives ample and exceptional opportunities to the students to develop capacities for high quality learning and to increase their ability to innovate. With the Internet and the World Wide Web, a wealth of learning materials in almost every subject and in a variety of media can now be accessed from anywhere at any time of the day and by an unlimited number of people as shown in Figure 1.[11][12]

There are various ways of incorporating ICT in teaching/learning. These include but not limited raising particular information about certain topics through visualization; forwarding learning material to students via emails; collaborating and communicating with people in other places, and in other countries in real time; exchanging information with others in different locations; and developing social skills through turn-taking, resources-sharing, and collaborative work.[13]

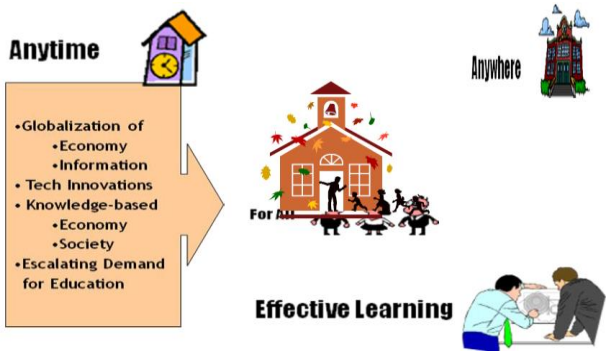


Fig. 1. ICT in teaching/learning [13]

ICT can provide speedy dissemination of education to target disadvantaged groups. ICT enhances the international dimension of educational services.[1][14] The experience of many teachers, who are early innovators, has it that the use of ICT is motivating for the students as well as for the teachers themselves.[15] It can also be used for non-formal education like health campaigns and literacy campaigns.[1]

IX. E-LEARNING

E-learning is a learning program that encompasses learning at all levels, both formal and non-formal, it makes use of an information network- such as the internet, an intranet (LAN) or extranet (WAN) whether wholly or in part, for course delivery, interaction and/or facilitation. Web-based learning is a subset of e-learning and refers to learning using an internet browser such as the Moodle, Blackboard or Internet Explorer.[16] It promotes learning in personal, constructive, ubiquitous, collaborative, and connective ways. E-learning enable students in several locations to get information about the school, teachers, documents, school fees, how and when to pay it, as well as other information according in the students' interest.

OECD [17][18] distinguishes E-learning into four different levels:

A. Web-supplemented *courses:*

These focus on classroom-based teaching but include elements such as putting a course outline and lecture notes online, use of e-mail and links to online resources.

B. Web-dependent courses:

These require students to use the Internet for key elements of the programme such as online discussions, assessment, or online project/collaborative work, but without significant reduction in classroom time.

C. Mixed-mode courses:

The e-learning element begins to replace classroom time. Online discussions, assessment, project/collaborative work replace some face-to-face teaching/learning. However, the significant campus attendance remains part of the mix.

D. Fully online courses:

Students can offer courses from a University or institution that is in another city, town, country or time zone without being physically there.

The use of E-Learning for continuous training in both informal and non-formal learning cannot be over-emphasized. The advantages of ICT usage in e-learning include but not limited to flexibility; reduced cost; improved interaction; enhanced tracking, monitoring, archiving and retrieving processes; sharing and re-use of resources; collaborative work and student-centred learning. The drawbacks include lesser control over the student adherence to the study calendars; difficulty assessing the students' practical work.[19][20]

X. ICT IN EDUCATION SECTOR

ICT in the education sector plays an important role, especially in the process of empowering technology into the educational activities. Computers, projectors and smart boards are some of the ICT tools used in educational sectors. Education sector can be the most effective sector to anticipate and eliminate the negative impact of ICT by recognizing the

importance of implementing ICT in supporting the teaching and learning process in the classroom. Educational ICT tools can be divided into three categories: Input source, Output source and others. See Figure

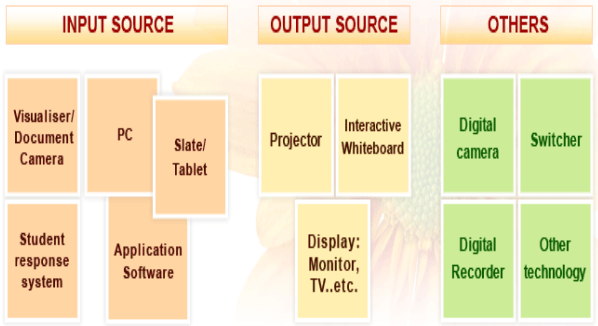


Fig. 2. Educational ICT Tools [21]

There are some important facts that pinpoint the need for ICT in the modern education. First, ICT is evolving very rapidly, hence there is need to integrate ICT into the educational activities. Second, the influence of ICT, especially internet (open source tool) cannot be ignored in the students' lives. Thus, the learning activities should incorporate the availability of the open source tool on the internet. Third, the presence of multimedia games and online games by internet is another measure that can be applied to teaching/learning in education sector to encourage the students in learning. This can be achieved by edutainment or educational games. Fourth, there should be an active participation, initiative and good will of the schools and the government institutions to enhance ICT implementation at school. Fifth, the teachers as the main motivator and initiator should be aware of the social change in their teaching activities; be the agent of change from the classical method into the modern one; and be the part of the global change in learning and teaching modification.[22]

The objectives of ICT implementation in education include implementing the principle of life-long learning/teaching education; increasing a variety of educational services and medium/method; promoting equal opportunities to obtain education and information; developing a system of collecting and disseminating educational information; promoting technology literacy of all citizens, especially for students; developing distance education with national contents; promoting the culture of learning at school (development of learning skills, expansion of optional education, open source of education, etc.); and supporting schools in sharing experience and information with others.[13][22]

XI. BENEFITS AND DRAWBACKS OF ICT IN TEACHING/LEARNING

According to Unwin [23], ICT technologies can enable people learn effectively at whatever level they choose, it encourages new way of studying and assessment that can be enabled to be part of the learning process. The benefits of ICT include but not limited to access to updated information on various subjects; an exciting learning experience; improvement in giving students assignments; access to remote learning resources anytime and anywhere; improvement in teaching through prior and a-prior knowledge; increased communication across geographical barriers; better team work for teachers and students; better class interactivity, availability of resource materials for research and other educational purposes; easier way of handling difficult and complex problems; aid for improving legible writing and clear drawing; and so on.[24]

Teachers who are adept users of ICT will gain the following benefits: improved resource sharing; greater flexibility in when and where tasks are carried out; better ICT literacy skills, confidence and enthusiasm; easier planning and preparation of lessons and designing materials; and enhanced professional image. The ICT's benefits to students include more focused and higher quality lessons; greater collaboration; improved pastoral care and behaviour management through better tracking of students; improved reading and comprehension; better development of writing skills; better independent and active learning, higher level learning styles, increased self-confidence, self-esteem and motivation to learn; and easier communication with teachers.[25][26]

Though, the advantages of ICT to teaching/learning are numerous, there some drawbacks that one cannot but emphasized. In distance education, the teacher-to-students physical interaction is almost nil. In an environment where ICT is entirely new to both teachers and students, learning may be relatively difficult. Other disadvantages include difficulty in setting up device due to cost implication; people's negative attitude to change in technology; encountering technical problems in the course of teaching; poor maintenance and lack of helpdesk support; and the problem of plagiarizing other people's work.[24][27]

VI. CONCLUSION

ICT contribution to teaching/learning cut across every subject, race, age, national origin, gender, time and space. The opportunities to collaborate with other people within and without the school environment motivate high quality teaching as well as better learning. . ICT when put into effective use facilitates the learning needs and styles of individual student such that good success can be achieved in teaching/learning to better enhance the educational system. Some major setbacks include the human attitude to change and lack of resources.

REFERENCES

[10] UNESCO (2002) Information and Communication Technology in Education–*A Curriculum for Schools and Programme for Teacher Development*. Paris: UNESCO.

[11] UNESCO 2008b. ICT Competency Standards for Teachers: Policy Framework [Online]. Available from UNESCO at: http://cst.unesco-ci.org/sites/projects/cst/The%20Standards/ICT-CST-Policy% 20 Framework.pdf

[12] http://www.teachthought.com/wp-content/uploads/2013/07/21stcenturylearning-synthesis.png
[13] Castells M. (2000). The Rise of the Network Society, Second Edition. Oxford: Blackwell Publishing. Original publication 1996.
[14] Daniels J.S. (2002) "Foreword" in Information and Communication Technology in Education–A Curriculum for Schools and Programme for Teacher Development. Paris:UNESCO.
[15] Romaniello, A., Rey, U., Carlos, J., & Medlin, D. (2010). "Higher education success and ICT".Bulletin of Applied Computing and Information Technology, 7(1), 109. Retrieved from http://books.google.com/books?id=6S0napW1w2kC&pgis=1
[16] Zelenekova, M., Pavolová, H., & Bakalár, T. 2012. "Internet Communication in the Process of Education at Universities". Procedia - Social and Behavioral Sciences , 46, pp. 2711–2715.doi:10.1016/j.sbspro.2012.05.552
[17] Yusuf, M.O. (2005). Information and communication education: Analyzing the Nigerian national policy for information technology. International Education Journal Vol. 6 No. (3), Pp; 316-321.
[18] Syed N. (2009) An Effective use of ICT for Education and Learning by Drawing on Worldwide Knowledge, Research, and Experience: ICT as a Change Agent for Education (A LITERATURE REVIEW) An unpublished Ph.D Research work at the Department Of Education, University Of Kashmir. Available at http://www.nyu.edu/classes/keefer/waoe/amins.pdf
[19] Aduwa-Ogiegbaen, S. E., & Iyamu, E. O. S. (2005). Using information and communication technology in secondary schools in Nigeria: Problems and prospects. *Educational Technology & Society*, *8*(1), 104-112.
[20] ELNORD (2006) "eLearning Nordic 2006", Edited by Pedersen, S.G., Malmberg, P., Christensen, A.J., Pedersen, M., Nipper, S., Græm, CD., Norrgård, J. & Ramboll Management. Copenhagen: Ramboll Management. www.ramboll-management.com
[21] Adeyemo, S.A (2010) The Impact of Information and Communication Technology (ICT) On Teaching and Learning of Physics *International Journal for Educational Research and Technology volune1 [2]* pp 48-59
[22] Joshi, L (2014). The Role of ICT for Improving Education in Schools and Higher education. IOSR *Journal of Engineering (IOSRJEN)* ISSN (e): 2250-3021, ISSN (p): 2278-8719 Vol. 04, Issue 01 (January. 2014), ||V3|| PP 30-33. www.iosrjen.org[Accessed Sep15, 2014]
[23] Mikre, F (2011). The Roles of Information CommunicationTechnologies in Education. Review Article with Emphasis to the Computer and Internet. *Ethiopia Journal for Education and Science.* Volume 6 no.2 [Accessed Sep. 29, 2014].
[24] Plomp, T.; Pelgrum, W. J. & Law, N. (2007), 'SITES2006—International comparative survey of pedagogical practices and ICT in education', *Education and Information Technologies* Vol.12, No. (2), Pp; 83- 92.
[25] Tinio, V.L. (2002). ICT in Education: UN Development Programme. Available: http://www.eprimers.org.
[26] OECD (2000) Learning to bridge the digital divide (Paris, OECD, Centre for Educational Research and Innovation, National Centre on Adult Literacy) [Accessed July. 21, 2014]
[27] OECD (2005) "E-learning in Tertiary Education. Where do we stand?" OECD: Paris.
[28] http://ec.europa.eu/education/programmes/elearning/doc/studies/ict_in_vocational_en.pdf
[29] Perraton,H.(2000b) Open and distance learning in the developing world , (London and New York, Routledge).
[30] http://www.elmoglobal.com/en/html/ict/01.aspx
[31] Saverinus, K. (2008). The Role of ICT in Education Sector. Avaliable at http://verykaka. wordpress.com
[32] Unwin, T. (2004). Towards a framework for the use of ICT in teacher training in Africa. *Open Learning: The Journal of Open and Distance Education* [Online]. 20, pp 113-129. Available from: http://www.gg.rhul.ac.uk/ict4d/ict%20tt%20africa.pdf[Accessed Sep. 30, 2014].
[33] http://ict-adv-disadv.blogspot.com/
[34] Hooker, M. (2009). The Use of ICT in Teacher Professional Development. Global e-School and Community Initiative. Available from :http://www.gesci.org[Accessed Sep. 13, 2014].
[35] Olakulehin, F. K. (2007). Information and communication technologies in teacher training and professional development in Nigeria. *Turkish Online Journal of Distance Education, 8*(1) Retrieved from http://tojde.anadolu.edu.tr/tojde25/pdf/article_11.pdf
[36] http://unpan1.un.org/intradoc/groups/public/documents/APCITY/UNPAN005832.pdf. Accessed August 03, 2014

Digital Divide and the Imperative of Bridging the Gap in the Developing Countries (The Nigerian Perspective)

Osisanwo Temitope A
Department Of Library and Information Science
Tai Solarin University of Education,
Ijagun Ijebu Ode, Ogun state.
osisdetopsy@yahoo.com

Abstract **-** The Digital Divide is the gap that exist between those with regular, effective access to digital technologies in particular the internet and those without. It is a phenomenon in the developing countries caused by various factors like economic, social, political, level of education and literacy factors. Various ways of bridging the digital divide were suggested among which include easy access to technological tools, policy implementation among others and benefits of bridging the digital gap was also suggested.

Keywords: Digital divide, Bridging, Gap

I. INTRODUCTION

The origin of digital divide has been associated with various individuals in the United State of America during the first part of the 1990s. The former US president, Bill Clinton is thought to have been the first to use the term in discussion of the National Informative Infrastructure (NII) Advisory Council in 1993. The Concept of digital divide is associated with the gap that exists between developed and the developing Countries with regards to access to information and Communication technology. It is seen according to Spectar 2010 as inequitable access to information and Communication technologies such as PCs, internet, telephone cable and other internet- related technologies by individual or group of people in a country. Warschauer 2002 described it as much more than providing internet and computer connections, because access to ICT is embedded in a complex array of actors encompassing physical, digital, human and social relationships.

Norris (2001) describes it as a multidimensional phenomenon which encompasses three distinct aspects, namely: the global divide which refers to the divergence of internet access between industrialized and developing societies; the social divide which is concerned with the gap between information rich and information poor in each nation; and finally the democratic divide which signifies the difference between who do and do not use the panoply of digital resources to engage, mobilize and participate in public life.

James (2003) defines the global digital divide as the strikingly differential extent to which rich and poor countries are enjoying the benefits of Information technology and as the unequal distribution of computers, internet connection. Also, Dubey and Devanand (2011) defined the digital divide as the gap between individuals, households, businesses and geographic areas at different socio-economic levels with regards both to their opportunities to access ICTs and to their use of the Internet for a wide variety of activities. As the internet is rapidly growing to become the basis for virtually all aspect of the global economy. The term digital divide which is often referred to as the internet access is a divide that affect and strengthen fundamental economic and social divide between and within countries and is threatening to further exacerbate these inequalities

Digital divide is also seen as the gap that exists between people at the community, institutional or national level with regards to ability to access, adapt, understand and use emerging technological tools.

According to UN secretary General Kofi Annan (1999) Communication and access to communication technologies are just like social, security fundamental human rights and that the digital divide is a pressing humanitarian issue. Soon the World's population will pass the six billion mark. Five out of the six billion live in developing countries. For many of them, the greatest scientific and technical achievements of our era might as well be taking place on another planet. The capacity to receive, download and share information through electronic networks, the freedom to communicate freely across national boundaries must become realities for all people.

Manuel Castells (2002) defines the digital divide as inequality of access to the internet. Access to the internet is moreover a requisite for overcoming inequality in a society which dominant functions and social group are increasingly organized around the internet while Van Dijk one of the most important theorist of the network society defines the digital divide as the gap between those who do and do not have access to computers and the internet.

II. FORMS OF DIGITAL DIVIDE

According to Okpaku (2011) Digital divide can further be sub divided into different aspects and configuration. They include:

a) The Infrastructural Digital Divide

The *Infrastructural Digital Divide,* which is what most people mean when they use the expression "digital divide", refers to the vast differential in the extent, quality and versatility of information and communications infrastructure and networking capacity between the industrial world and the developing world. This differential exists not only between, but also within, regions and countries, and even within cities and local communities. The infrastructural digital divide mirrors, approximately, the patterns of wealth in society, and the correlative distinctions of demand, affordability, and political, social and economic clout required to drive the new process from demand to access. There is general consensus as to the nature and topography of the infrastructural digital divide. This is not likely to be the case when we come to other aspects of the digital divide.

b) The Information Digital Divide

The *Information Digital Divide* is the vast differential not only in the access to information and data, but in the ability to put a spin on both and post them on the global tableau as the predominant information and interpretation of facts and events. This definition is deliberately loaded with subtle implications because it seeks to challenge commonplace assumptions. While digital transmission capacity greatly facilitates the process of accessing and transmitting information and data, it does not, *per se*, bestow accuracy or authenticity to such information (or data, for that matter), and definitely has no intrinsic correlation with truth, value or meaning. In fact, the enormous ability of digital network capacity to manipulate information is such that, more and more, it appears to be fast becoming inversely proportional to truth, authenticity, value and meaning.

The disproportionate nature of this ability by one sector of the world to post myopic information duly vested with self-interest-motivated interpretation, and its own spin on facts and events, especially where the interests of others are at the core of the issues at stake, makes the asymmetrical nature of digital capacity a serious threat to the value of information in the global arena of complex and often conflicting perspectives, perceptions and purposes. It also most often blanks out the information from a major proportion of human society.

c) The Intellectual Digital Divide

Intellectual capacity is essentially the ability and wherewithal to acquire and manage information, analyze experience, and draw on existing knowledge, and combining all these processes to create new insights and knowledge, thereby enhancing our ability to cope with the challenges of life and to enrich human culture. The differential in this capacity, especially as it relates to the role of digital capacity in the process, constitutes the *Intellectual Digital Divide*. Here, again, the topography is not along traditional contours of industrial versus developing societies, popular wisdom notwithstanding.

d) The Human Resource Capacity Divide

Directly related to this is the *IT Human Resource Capacity* Divide, or the differential in the expert capacity, especially at the cutting edge, of societies to acquire, master and manipulate the enabling tools of Information and Communications Technology to facilitate the management of life, the conduct of human interaction across vast distances, and the creation of innovations to solve the material and intellectual challenges to quality existence. While the ability of societies to derive maximum benefit from digital innovations depends on how widespread such capacity is, the critical Digital Human Resource Capacity of any society at the level of creating dedicated innovations and solutions need only consist of a core of cutting edge expertise to jumpstart the process of creating solutions to bridge the digital divide.

III.CAUSES OF DIGITAL DIVIDE IN DEVELOPING COUNTRIES

There are many factors that causes digital divide in the developing countries. Although poor ICT infrastructure and effective access to the internet among others are seen as the major cause of digital divide. Others will be discussed as follows:

a) *Economic factor:* This is one of the major causes of digital divide in Africa. The level of economic development of a nation determines to a large extent the level of ICT usage in homes and communities. The disparity in the economic development between communities within and between countries has been responsible for the inability of communities or individuals to purchase computer systems, buy modern or pay for internet bandwidth or get an internet enabled phones and pay for electricity or generator needed to power the digital technologies.
Also, according to (Murtala 2003) ICT is widely perceived as a new engine of economic development but the people of Africa in general, especially those living in rural areas do not easily understand how ICTs can provide them with their basic needs such as food, basic healthcare, shelter, clean water e.t.c, when they cannot even afford a single meal a day.

b) *Political factors:* This is another major factor that is responsible for digital divide in African Countries. Inadequate budgeting for Education at the primary, secondary and tertiary levels. Most African countries merely gives 1% of its global expenditure on education. (Johnson 2001). And when much money is allocated such fund is always diverted either for selfish reasons by the Politicians. Also, the political environment of the developing countries will determine the rate at which foreign investors will establish different infrastructures that will bridge the digital gap.

Migration factors: Most of the African Countries

have been hit by the migration of its key professionals looking for greener pastures, quality education as well as scholarship programmes in the developed Countries but they never return home after achieving their goals or completing their education. These professionals and students would have made a more meaningful contribution towards digital divide if they had returned to their countries.

c) Level of Education and Literacy factor: The level of Education and literacy of Adults especially in the developing countries is another factor that contributes to digital divide. Developing countries especially in Africa are in the bottom rank especially with regards to adult literacy (Mutula 2008). In addition most of these countries lack post literacy programmes to guard against relapse into illiteracy and as a result functional literacy achievement cannot be sustained.

d) ICT Policies Implementation: Most Institution and Organizations are having enough ICT facilities in place but they are highly under-utilized. The under-utilization of available ICTs is attributed partly to inimical policies such as unnecessary restrictions on access to the technologies in various organizations e g restricting students from entering the computer laboratory, limiting or reducing the number of hours a student can spend on the internet per day, restricting students or selected students from gaining entrance to the electronic library section etc.

e) The fear of technology: This is another factor that contributes to the digital divide. Many people do not use technology because they are not confident in their own ability to use computer skills. They are afraid that others will laugh at their attempts therefore it is easier not to try it and also think that their poor attempt might cause a damage or problem to the computer and therefore prefer not to attempt it.

f) Lack of motivation: This is another factor that tends to contribute to digital divide in Africa or developing countries many people are not prepared to put in the time and effort to learn something new as some lisle ICT use with irrelevant activities such as. Computer, game, playing eyes damage e.t.c

g) Cultural Factor: Africa and the developing countries are yet to effectively integrate local languages into ICT to enable ready access to ICT by the people. In schools Colleges and Universities, the language of delivery is largely English and local language are hardly used. Also the curriculum used in schools has more of foreign content which is in English language. Research has shown that people with limited English speaking skills lag behind in their English speaking counterparts in access to computers and the internets

h) Telecommunication Constraint: The problem of inadequate bandwidth to facilitate effective access and use of the internet resources is another problem causing digital divide in developing countries. The bandwidth problem emanate from inefficiencies in the regional telecommunication network which affect access to and use of the internet. The quality of the internet connectivity on account of limited bandwidth has hampered effective access to the network.

IV. BRIDGING THE DIGITAL GAP

Digital divide is an issue that should concern the government, development agencies and non- governmental organizations not only in developing countries but also in developed world. Practical steps are needed to ensure equal access to information technologies for all. Some of the measures needed to put in place to ameliorate digital divide are discussed below:

a) Provision of Technological tools for the developing countries

James 2003 agrees that one at the possibly way of solving the global divide is to transport. Old computers from rich to poor countries due to rapid technological progress and the non-upgrade ability of most hardware which causes. People to buy new computer every 2 or 3 years as well as heavy profits of the hardware and software industry.

b) Information and Communication Technology to be seen as a fundamental human right

Also, information and communication is just like social security a fundamental human right. This right is explicitly mentioned. In article 19 of the universal Declaration of human rights. "Everyone has the right to freedom of opinion and expression, this right includes freedom to hold opinion without interference and to seek recare and impart information and ideas. Through and media and regardless of frontiers" In and information societies opinions are increasingly expressed an d articulated with the help of the internet and other news media. Hence, materials usage and skills access to new technologies is a contemporary expression of a fundamental human right. It is unjust that western citizens enjoys more human right and economic, social cultural and technological resources than the developing countries citizen. So government in the developing countries needs to ensure of the provision of unequitable access to information and communication tools for the citizens just as security health and other social amenities are at a fundamental human right to the citizens.

c) *Competition and Promotion*

Wilson (2006) agrees that one of the measure for solving the digital divide should be the enhancement of completion and the promotion of investment through the public or the communal ownership of telecommunication infrastructures.

d Attracting foreign capital investors by creating an enabling environment

Liberalizing telecommunications market in developing countries will attract western corporation to invest in the ICT sector in these region and that this will result in economic growth that benefits all and lowers internet and phone prices due competition. Also ICT application in the areas of e'commerce e travelling, e government, e transport, e health, e-education, e-learing etc are mainly developed in western countries and benefit under current conditions mainly western corporation if they are exported to developing countries.

e) Easy access to technological tools

Access to technologies should be made universal guaranteed to all people by the public, free of cost and based on open source. This can be best, achieved of provided not by private organizations but by public ones because the letter are not based on profit .Interests that might undermine universality but on the common interest in common goods.

Also, managing all spaces such as schools, libraries information centres, public and community centre ,a primary location for accessing and teaching computer skills for adult and children.

The issue of high cost of internet protocol bandwidth should be carefully addressed by renegotiating the existing global telecommunication agreement and restructuring of the worldwide web connects in the developing countries.

f) Policy implementation

There is the need for government to implement proactive policies in institutions and organization that will enable them to implement the latest low cost technologies and be empowered to grow the ICT skills base within the institution and also enable the creation of awareness about ICTS to enhance understanding of the use of digital technologies.

V. Benefits that will be derived from Bridging the digital divide

There are numerous benefits that will be derived from bridging the digital divide. Information and Communication Technology is an enormous engine of development and it has the potential for commercial use by the local enterpreneurs which may generate employment and economic growth.

According to Mutula (2008) A vibrant ICT service sector can provide better paid skilled employment for example by increasing both the demand and ability to pay for better education health and other social services.

Also, it promote academic excellence by enabling every student to gain access to instructional resources which include technology and support services with needed skills to harness digital opportunities.

Another benefit that could be derived from bridging the digital divide is that it can open doors of economic opportunity for individuals with different disabilities because working with computer requires low mobility or physical strength.

It could also be helpful in addressing the HIV/AIDS, cancer, malaria etc crises in Africa and other developing countries through improved Communication associated with national ICTs by creating awareness and providing easy assistance and needed information.

VI. Conclusion

Digital divide is phenomenon in Africa and some other developing countries that has to do with the gap that exist between people with regards to access to and use of various technologies through various factors of economic, political, cultural, telecommunication constraint etc

Digital divide in a developing countries is a question of differences in possibility of access to the internet and ICT in the rural and urban areas. It is basically a technological divide but the problem when it comes to bridging this divide seem to be a combination of political, cultural, literacy level and financial obstacle. Our recommendation is therefore to provide more technological tools for the developing countries, make ICT to be seen as a fundamental right, provide an enabling environment to attract foreign investors so that people will be able to utilize internet for their own human development and thus bridge the divide.

References

[1] Annan K 1999 Speech at the ITU Telecom Opening Ceremony October 9 1999. Available from www.itu/telecom-wt99/press service /information for the press/presskit/speeches/annan

[2] Dubey. P Jeevanjyoti and Devanand (2011). A study to examine the Digital Divide factors:Jamma and Kashmir Perspective. New Dehli. Journal of Information Technology

[3] Euchs & Horak 2008 Africa and the digital divide. Science Direct Australia

[4] James. J . 2003. Bridging the global Digital Divide .Edward. Egar ,Chetterham

[5] Murelli . E 2002. Breaking the Digital Divide implicetions for developing countries. Commonwealth Secretariat, London

[6] Mutula S (2008) Digital Divide in Africa: its Causes and Amelioration Strategies. Third World Information Service Limited. Ibadan

[7] Norris .P . 2001 .Digital Divide Civic Equipment, Information Poverty and the internet Worldwide. Cambridge University Press, New York

[8] Okpaku. J (2011) Ownership of problem, Intellectual Property and Digital Divide. The Enabling challenge of solution. WIPO Second International Conference on Electronic Commerce and Intellectual property. Geneva. September 2001. Available from http:// www.wipo_ec_conf_01_spk_12c.

[9] Peters T. (2003) Bridging the Digital Divide Global issues. An electronic journals of the U.S. Depatrtment of state.

[10] Spatar .J .2000 Bridging the Global Divide: Framework for Access and the world. Wireless Web. North Carolina Journals of international law and commercial Regulation.

[11] Van Dijk 2005 The Deepening Divide. Inequality in the Information Society.SAGE London

[12] Van Dijk (2006). The Network Society. Social Aspects

[13] World Summit on the Information Society (WSIS) 2003 Declaration of principles. Building the Information Society: A Global Challenge in the New Millennium. Available from: www.itu.int/wsis/docs/geneva/official/doc

A Comparative Analysis of Software Engineering Models: Classic and Modern Rapid Models

Akintola K.G.
Computer Science Department,
University of Houston-Victoria,
Texas 77901
United States of America.
akintola2087@yahoo.com

Abstract- Prominently there are two schools of thought in software engineering methodologies in today IT industry. They are classified as the classic and modern schools of thought. While the classic believe that software must pass strictly through developmental processes with each process completed before the next process begins as outlined by the waterfall model, the modern approach on the other hand believe that software should be evolutionary as a whole and not in parts or stages. Agile development belongs to the second school of thought. These development strategies include extreme programming, scrum, and the variants. This paper studies the two models. The waterfall model and the EXTREME programming software models that have been adopted in software development over the years.

I. BACKGROUND

Software systems pass thorough some stages during their life cycle just as many systems undergo changes before extinction. The stages of software systems are in phases which can vary somewhat but generally include the following: conceptualization; requirements and cost/benefits analysis; detailed specification of the software requirements; software design; programming; testing; user and technical training; and maintenance.
Software Engineering is an approach to developing software that attempts to treat it as a formal process more like traditional engineering than the craft that many programmers believe it is. We talk of crafting an application, refining and polishing it, as if it were a wooden sculpture, not a series of logic instructions. The problem here is that you cannot engineer art. Programming falls somewhere between an art and a science. Therefore, Software Engineering is intended to mean the best-practice processes used to create and/or maintain software, whether for groups or individuals, in attempt to rid ourselves of the usual haphazard methods that have plagued the software industry. This would include subjects like configuration management, project planning, project tracking, software quality assurance, risk management, formal inspections, etc

Since inception of information systems, many models of systems life cycle have been adopted in recent years.

II. WATERFALL MODEL

Royce (1970) proposed the software life cycle using a waterfall model. This classic approach to software life cycle usually includes some of the following activities:

a. Analysis
b. Design and construction
c. Implementation (and continuing Operations)

a. Analysis

i. Initiation (*e.g.*, an RFP)
ii. Feasibility study
- Technical – *can* we build it?
- Economic – *should* we build it?
- Operational – if we build it, will it be *used*?
- Schedule – will it be ready *in time*?

iii. Requirements definition
iv. Specifications
v. Project plan

b. Design

vi. Logical design (*i.e.*, the external view)
vii. Physical design (*i.e.*, the internal view)
viii. Coding (or code acquisition)
ix. Testing

c. Implementation

x. Documentation – ouch! This should have been done all along!
xi. Conversion
- Direct
- Parallel
- Pilot
- Phased

xii. Training – both initial and continuing
- Users
- I/S staff
- Management

d. Installation

e. Operations
f. Production
g. Post-implementation audit
h. Maintenance

The waterfall model consists of the following phases:

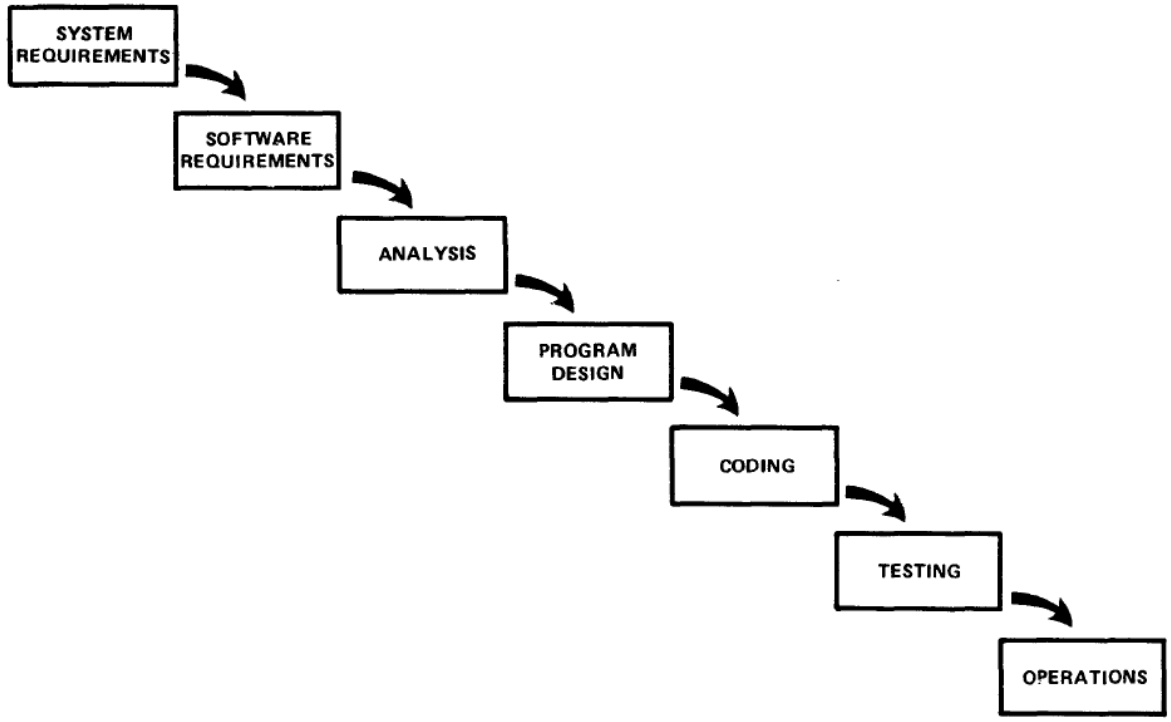


Figure 1.0 The Waterfall Model proposed (Royce 1970)

The essence of waterfall software development is that complex software systems can be built in a sequential, phase-wise manner where all of the requirements are gathered at the beginning, all of the design is completed next, and finally the master design is implemented into production quality software. This approach holds that complex systems can be built in a single pass, without going back and revisiting requirements or design ideas in light of changing business or technology conditions. Szalvay (2004).

A. Advantages of the waterfall model

If in the beginning of the project failures are detected, it takes less effort (and therefore time and money) for this error to be fixed. In the waterfall model lays much emphasis on documentation. In the newer software development methodologies less emphasis is laid on documentation. This implies that in the new setting, when new people are involved in the project, it will be difficult to transfer knowledge. The way of waterfall ensures that there are specific phases. This tells you what stage a project is, and therefore projects can be adequately planned for. Milestones can be used to monitor the progress of the project. The waterfall model is common in the developers' cycle.

B. Disadvantages of the waterfall model

It is non-iterative in nature. The waterfall methodology assumes that up-front planning is enough to take into account all variables that could impact the development process. In fact, waterfall projects allocate copious effort detailing every possible risk, mitigation plan, and contingency. Waterfall therefore equates software development to an assembly line. Almost no software system is so simple that the development can be entirely scripted from beginning to end. The inherent uncertainty and complexity in all software projects requires an adaptive development plan to cope with uncertainty and a high number of unknown variables Szalvay [2004]. It is a disadvantage that the waterfall model assumes that the requirements will not change during the project. When a requirement changes in the construction phase, a substantial number of phases will be affected. It is very difficult to time and cost estimate. The phases are

very large, it is therefore very difficult to estimate how much each step cost.

III. THE AGILE SOFTWARE DEVELOPMENT

The simple ability to revisit the "phases" of development dramatically improves project efficiency. The idea of revisiting phases over and over is called "incremental and iterative development" (IID). The development lifecycle is cut up into increments or "iterations" and each iteration touches on each of the traditional "phases" of development. For example, with IID requirements is an ongoing process that is periodically revisited. As new requirements surface and as the scope changes, IID processes continually capture the requirements iteration after
iteration Interestingly, Winston Royce (of waterfall process fame) later noted that his ideas were incorrectly interpreted and that a "single pass" framework would never work (his article actually advocates at least a second pass). IID allows for multiple "passes", or iterations, over a project lifecycle to properly address complexities and risk factors Szalvay [2004]. Agile methodologies embrace iterations. Small teams work together with stakeholders to define quick prototypes, proof of concepts, or other visual means to describe the problem to be solved (Serena, 2007).

One of the most important differences between the agile and waterfall approaches is that waterfall features distinct phases with checkpoints and deliverables at each phase, while agile methods have iterations rather than phases. The output of each iteration is working code that can be used to evaluate and respond to changing and evolving user requirements. Waterfall assumes that it is possible to have perfect understanding of the requirements from the start. But in software development, stakeholders often don't know what they want and can't articulate their requirements. With waterfall, development rarely delivers what the customer wants even if it is what the customer asked for. Agile methodologies embrace iterations. Small teams work together with stakeholders to define quick prototypes, proof of concepts, or other visual means to describe the problem to be solved. The team defines the requirements for the iteration, develops the code, and defines and runs integrated test scripts, and the users verify the results. Verification occurs much earlier in the development process than it would with waterfall, allowing stakeholders to fine-tune requirements while they're still relatively easy to change (Serena, 2007).

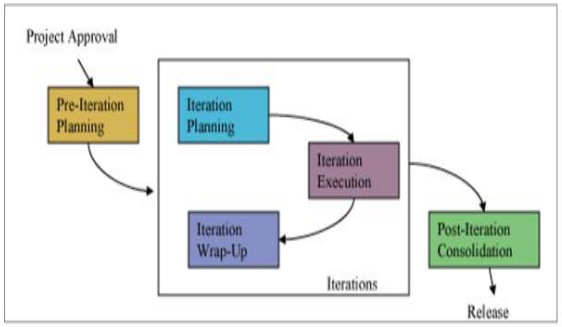


Figure 2.0 **Agile development process (Serena, 2007)**

IV. EXTREME PROGRAMMING (XP)

Traditional software engineering means have been characterized by a rather predictable process in the past. Users tell once and for all exactly what they want. Programmers design the system that will deliver those features. They code it, test it, and all is well. But all was not always well. The users didn't tell once and for all exactly what they wanted. They changed their minds. And the users weren't the only problem. Programmers could misjudge their progress. The academic software engineering community took the high cost of changing software as a challenge, creating technologies like relational databases, modular programming, and information hiding. This is where Extreme Programming comes in. Rather than planning, analyzing, and designing for the far-flung future, XP exploits the reduction in the cost of changing software to do all of these activities a little at a time, throughout software development. According to the author, XP is by no means a finished, polished idea. The limits of its application are not clear. To try it today would require courage, flexibility, and a willingness to abandon the project if necessary.
(Beck 1999).
Extreme Programming (XP) is one of many new lightweight software development life cycle methodologies . These methodologies have evolved from previous failures to produce complex software of acceptable quality in a reasonable amount of time. XP can provide many of the same benefits in a SAS programming environment that it does for small teams using object-oriented languages for commercial software development. It includes many of the same features as other lightweight processes but its emphasis on testing is unique .
Extreme programming has been has been advocated recently as an appropriate programming method for the high–speed, volatile world of Internet and web software development. Extreme can be characterized as a lightweight or agile methodology that is usually attributed to Kent Beck, Ron Jeffries, and Ward Cunningham. XP is targeted towards small to medium size teams building software in the face of vague and/or rapidly changing requirements (Paulk,2001). XP teams are expected to be co-located, typically with less than ten members.

The critical assumption underlying XP is that the high cost of change has been (or can be) addressed by technologies such as objects/patterns, relational databases and information hiding. As a consequence of this assumption, the resulting XP process is intended to be highly dynamic. (Paulk 2001).

The four basic activities in the XP life cycle are coding, testing, listening, and desining. The dynamism is demnstarted via four values. Continual communication, with the customer, and within the team, simplicity by always focusing on the minimalistic solution, rapid feedback, via unit and functional testing, (among other mechanisms), and the courage to deal with problems proactively (Paulk,2001).

Most of the principles espoused in XP, such minimalism, simplicity, and evolutionary life cycle, short cycles time, user involvement, good coding standards etc are commonsense and appropriate practices in any disciplined process. The 'extreme' in XP comes from taking common sense practices to extreme levels. Extreme Programming provides the mechanism to legitimize and sustain creativity throughout the development process.

Extreme Programming recognizes that requirements evolve as customers and users gain a better understanding of what they want. The approach invites change as it happens, reducing negative impact on the development cycle.

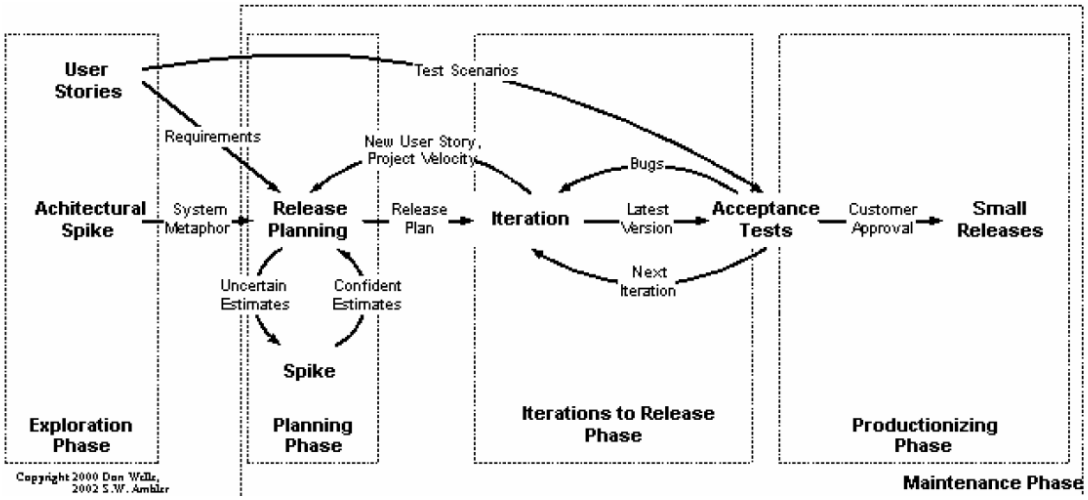


Figure 3.0 Xtreme programming.

A. The principles of Extreme programming

1) Programming:

Testing

Write unit tests before and while you develop to form the specification, and the proof of completion

2) Refactoring

Change the organisation of the code whenever necessary to simplify or to make it easy to add the next feature.

It is called refactoring, which means rewriting or rearranging or reorganizing or transforming code

- The trick is to wait, and do it only when it pays off, and do it in small stages, keeping the program working
- So you live with what you have until the next stage of development, then decide whether the change is worth it to make the next stage easier
- This depends vitally on unit testing, so you know that the changes haven't broken anything

3) Simple design

What is the simplest thing that could possibly work? to avoid risking effort on features that never get used.

4) Continuous integration

build and test the whole project about twice a day to ensure that the team is all pulling in the same direction

B. Teamwork Practices

1) Pair programming

All programming is done with two people using one computer communication is better and team members can be replaced.

2) Collective ownership

Any pair can change any code anywhere, anytime works because behaviour is defined by tests

3) Coding standards

The code (& tests) form a communication medium conventions are shared simply to increase communication

• Fixed working week

Never work overtime two weeks in a row if you need a

heroic effort, something has already gone wrong.

C. Business Practices

1) The planning game

Managers say how much a feature is *worth* developers say what a feature will *cost*

2) On-site customer

Available all the time to answer analysis questions helps write functional tests to prove features are done

3) Small releases

Deploy a simple version as early as possible release new versions on a very short cycle

4) Metaphor

Another word for "system architecture" simple shared story of how the whole system works

V. CONCLUSION

All the Agile developments methods have prototyping, iteration, continue release in common. Extreme programming in contrast with waterfall model supports:

- **Individuals and interactions** over processes and tools
- **Working software** over comprehensive documentation
- **Customer collaboration** over contract negotiation
- **Responding to change** over following a plan.

While Extreme is in support of rapid application development, the waterfall is in support of adequate planning and execution. We can see that there would be a better synergy between the two models in efficient project management in future.

REFERENCES

[1] Paulk M.C. (2001) Extreme Programming from a CMM Perspective XP universe Raleigh, NC.

[2] Royce W. (1970). Proceedibgs of IEE WESCON August 1970.

[3] Serena (2007) An Introduction to Agile software development http://www.serena.com/docs/repository/solutions/intro-to-agile-devel.pdf

[4] Szalvay V. (2004) An Introduction to Agile Software Development http://www.danube.com/docs/Intro_to_Agile.pdf

[5] Williams L., Crebs W., Layman L., Anton A.I., Abrahamsson P. (2004). Towards a framework for evaluating Extreme Programming Empirical Assessment in Software Engineering (EASE) pp. 11-20

An Assessment of ICT Knowledge and Skill Levels of Nigerian Colleges of Education Lecturers- The Example of Emmanuel Alayande College of Education, Oyo, Nigeria

AYOADE, Olusola Bamidele
Computer Science Department
Emmanuel Alayande College of Education, Oyo
Nigeria
ayoadebamidele2014@yahoo.com

***Abstract*-** This study was carried out to investigate the Information and Communication Technology (ICT) knowledge and skill levels of the lecturers in a Nigerian College of Education. The sample of the study consisted of (150) lecturers randomly selected from the five schools in Emmanuel Alayande College of Education, Oyo, Nigeria. This comprises 80 males and 70 females' lecturers. A modified instrument tagged Lecturers' ICT skills and usage survey adapted from ICT survey indicator for teachers and staff were used for the collection of data. The data collected are analyses by simple percentage and ANOVA. The results showed that lecturers have positive attitude toward the use of ICT. Gender and years of teaching experience do not significantly affect the ICT knowledge and skill levels of the lecturers. Also lecturers with less teaching experience have more competence on the use of basic ICT applications than the older ones. The study also revealed that both age and teaching subjects are significantly affecting the ICT knowledge and skill levels of the lecturers while lecturers teaching science and science related courses are more competent than older lecturers and lecturers teaching non-science courses. Finally, it was also found from the study that there are no adequate ICT facilities in the college; Lecturers can not adequately used ICT tools. Based on these findings, the college management should provide adequate internet and intranet facilities for both lecturers and students and organize training for lecturers on how to integrate ICT within their learning area.

Keywords: ICT, Competency, Gender and Lecturer

I. INTRODUCTION

The ubiquitous nature of Information and Communication Technologies (ICT) has transformed almost every facet of our society and radically changed how people live, work and play [9]. The use of ICTs in Nigeria and African countries generally is increasing and dramatically growing. However, while there is a great deal of knowledge about how ICTs are being used in developed countries, there is not much information on how ICTs are being introduced into schools in developing countries [1]. Looking at the developing countries according to these authors, there is generally limited access time per month using ICTs by both the teachers and students, and even less time spent with reliable Internet access. It should be noted that availability of ICTs vis-à-vis access in term of ratio of teachers and students differs significantly [15]. Despite this, the new and emerging technologies challenges the traditional process of teaching and learning, and the way education is managed. While information communication technology is an important area of study in its own right, it is having a major impact across all curriculum areas. Easy worldwide communication provides instant access to vast array of data, challenging assimilation and assessment skills [5]. Rapid communication plus increased access to ICTs in the home, at work, and in educational establishment, could mean that learning becomes a truly lifelong activity- an activity in which the pace of technological change forces constant evaluation of teaching process itself.

Meanwhile, it is observed that some studies have been conducted on the uses of ICTs by teachers particularly on the issue of their professional development. Most of these studies were carried out in developed countries where the use of ICTs has come of age, and where there are resources and material to maintain them. However, the use of ICTs by teachers in Nigeria is just beginning to gain popularity and researches in the area have just started emerging. Emphatically, the use of ICTs by teachers to teach the students is highly advantageous. This is because its enable them to demonstrate understanding of the opportunities and implications of the uses for learning and teaching in the curriculum context; plan, implement, and manage learning and teaching in open and flexible learning environment.

The idea that teaching and learning can successfully take place through the application of electronic communication facilities between teachers and students is one which had generated, sometimes,

hope and dismay and at other times, excitement and fear. Hope that many more learners can be reached at a more convenient pace that had erstwhile been the case, dismay that the infrastructures necessary for deploying an effective ICT platform is lacking in low-income countries like Nigeria [12].

It is a common misconception that access to technology on its own motivates teachers to apply it in their teaching. The biggest barriers to the use of computers identified by teachers participating in the 1998-1999 survey assessing the World Links schools programme were the lack of time available in classes, and in their own schedules for planning; and the lack of a national policy on the use of computers in schools [10]. Relatively few teachers identified infrastructure problems, such as the lack of computers in working condition, unreliable electricity or lack of access to the internet, although these varied by country. As less technologically advanced countries joined the programme in 1999-2000, the major barriers to ICT classroom use became the lack of computer hardware (60%), software (56%) and reliable internet connections (52%), particularly in African countries such as Mauritania, Ghana and Zimbabwe. Lack of access to technology is inevitably a major barrier in its application, but availability does not necessarily translate into use. However, in the study of Nigerian secondary school teachers by [15], a lack of technical support in the schools, and teacher's lack of expertise in using ICT, was the prominent factors hindering teacher's readiness and confidence in using ICT.

The lack of incentives and support for teachers are other factors hindering their use of ICT. The SRI-World Links evaluation [10] shows that teachers enthusiastically engage in collaborative projects and often portray constructivist pedagogy. However, school administrators offer very little structural support and few incentives to use the technology effectively in the classroom. Too often the curriculum in developing countries is rigid and overloaded, leaving little time for innovative classroom practices. National policies need to make more commitment to helping teachers effectively integrate computers and internet technologies into the classroom by aligning curricula, exams, and incentives with the educational outcomes that they hope to gain. In the end, computers by themselves bring very little to the learning process – they are only tools for teaching and learning.

Teachers need to be supported to get the most from using ICT in classrooms. Particularly where resources are limited, as is common in Africa. ICT initiatives need to be driven by the provision of appropriate technological solutions for the challenges faced by communities, rather than by an interest purely in these physical technologies themselves. Those implementing technological solutions need to ensure they are context-specific, and adapted to local needs and conditions. It is also imperative that ICT initiatives are sustainable or effective by ensuring that the technologies embedded within them meet the demands of users in appropriate ways. It is essential that potential users have a sound understanding of how to use new ICTs beneficially, and a cultural view of the relationship between learning and technology.

The main purpose of this study was to investigate the information communication and technology knowledge and skill levels of lecturers in Emmanuel Alayande College of Education, Oyo.

II. RESEARCH QUESTIONS

The following research questions were formulated:

i. What is the status of the ICT infrastructural facilities in the college?
ii. To what extent do lecturers use ICT to achieve personal/professional objectives?
iii. What is the general attitude of lecturers towards the use of ICT?
iv. To what extent do lecturers use ICT in the classroom to achieve learning outcomes?
v. Do gender, age, teaching experience and teaching subjects have any effect on the ICT competence of the lecturers on the use of basic ICT application?

III. METHODOLOGY

A. Sample

The participants were 150 lecturers randomly selected from the five schools at Emmanuel Alayande College of Education, Oyo. Thirty (30) lecturers were selected from the following schools: Science, Vocational & Technical Education, Education, Languages as well as Arts & Social Sciences. The gender of the lecturers was stratified into male and female, their ages were classified into 25-34,35-44,45-54 and 55 and above. Their teaching experiences were also classified into 0-9, 10-15, 16-20 and 21 and above while the teaching subjects were classified into science courses, science related courses, and non-science courses.

B. Survey Instrument

The survey instrument used for this research work is a questionnaire and it was developed by the researcher based on established procedures in literature. The survey instrument contained six sections. Section A included four items and it focused on demographic information of lecturers: gender, age, teaching experience, and teaching subject. Section B focused on school ICT capacity available for students and lecturers. The section contained 14 items and the response modes were "Available and easily accessible" (AE coded 4);

"Available but not easily accessible" (AB coded 3); and "Not Available" (NA coded 1). Section C contained 9 items on how often the lecturer uses ICT in the classroom to achieve learning outcomes and the response were "Regularly" (RG coded 4); "Sometimes" (SO coded 3); "Rarely" (RA coded 2); and "Never" (NE coded 1). Section D contained 7 items on how often lecturer use ICT to achieve the personal/professional objectives and the response were "Regularly" (RG coded 4); "Sometimes" (SO coded 3); "Rarely" (RA coded 2); and "Never" (NE coded 1). Section E of the questionnaire which contained 10 items addressed the attitude of lecturers to information and communication technology (ICT) and the Likert response mode of Strongly Agree (SA), Agree (A), Disagree (D) and Strongly Disagree (SD) were used. Lastly, Section F was designed to know the level of competence of lecturers in the use of ICT, specifically, basic ICT competence and not the educational ICT competence. The section contained 35 items and the response modes were: "I am fully competent with this application/operation" (FC coded 4); "I am a regular and confident user of this application" (RCU coded 3); "I have used this occasionally but need further training" (OU coded 2); "I do not use" (DU coded 1); and "I am not aware of this application/operation" (NA coded 1). The instrument was given to experts in the field of both computer and test and measurement to comment on their suitability as well as the construction. The reliability coefficient was calculated to be 0.82 using the Kunder-Richardson formula 21.

C. Procedure and Data Analysis

The instrument was administered on the sample during the first semester of the 2011/2012 academic session. One hundred and forty copies were returned. The responses of the respondents were tabulated and compared, and descriptive analysis (percentages and means) were done to present the details about the ICT knowledge and skill levels of the lecturers. Furthermore, simple percentage and ANOVA statistic were used to test the hypothesis generated in the study.

IV. RESULTS

Table 1. Demographic Information of Participants

Variables		N	%
Gender	Male	80	57.1
	Female	60	42.9
Age	25-34	05	3.6
	35-44	85	60.7
	45-54	35	25.0
	55-above	15	10.7
Teaching Experience	0-9	42	30.0
	10-15	53	37.9
	16-20	28	20.0
	21-above	17	12.1
Teaching Subject	Science	25	17.9
	Science related courses	65	46.4
	Non-science	50	35.7

The demographic information of the participants is given in Table 1. The table indicates that 57.1% were male lecturers while female lecturers were 42.9%; this shows that both male and female lecturers were fairly represented. With regards to age of the lecturers, the table shows that young lecturer was more represented than the older ones (i.e. 3.6% and 60.7% is more than 25.0% and 10.7%). Likewise, the lecturer with 0-9 and 10-15 years of teaching experience were more represented than those with 16-20 and 21-above (i.e. 30.0% and 37.9% is more than 20.0% and 12.1%). Finally, lecturers teaching science and science related courses were more represented than those teaching non-science courses (i.e. 17.9% and 46.4% is more than 35.7%).

Research Question One: What is the status of the ICT infrastructural facilities in the college?

The results in table2 shows that item 12 was the only one that is available and easily accessible by the lecturers and six items (2,3,5,9,11 and 13) out fourteen items of the school ICT capacity were not available for use based on the responses of the lecturers(100%,100.0%,100%,100.0%,100% and 100.0%) while seven items (1,4,6,7,8,10 and 14) were available but not easily accessible by the lecturers based on their responses (78.6%, 80.0%,90.7%, 80.0%,90.7%,86.4% and 57.9%). This result indicates that the ICT capacity were not adequately available in the college which in turn incapacitated the ICT skills and knowledge of the lecturers.

Research Question Two: To what extent do lecturers' used ICT to achieve personal/professional objectives?

The results in table 3 shows that item 4 (86.4%) was the only one that lecturers regularly used ICT to achieve while items 3,6 and 7(54.3%, 70.7% and 51.4%) were never used by the lecturers. Also, items 1 (35.0%) and 2 (40.0%) were rarely used by the lecturers. This result indicates that lecturers were not adequately using ICT to achieve personal/ professional objectives.

Research Question Three: What is the general attitude of lecturers towards the use of ICT?

Based on the results in Table 4, there is every indication that response to the positive statements (items 1,2,3,5,6,7,9 and 10) shows that more than 50% of respondents have a positive attitude towards ICT. It is seen that more respondents believe that ICT could generally provide better learning experience. However, for negative statement item 4, the result shows that only about 70.0% of the respondents disagree that ICT has limited capacity to provide benefits in the classroom while 30.0% agreed that ICT has limited capacity to provide benefits in the classroom. Item 8 shows that about 56.4% of respondents agreed that ICT can't address

the needs of school system while about 43.6% of the respondents disagree on the fact that ICT can't address the needs of school system. As seen from the analysis in Table 4, student-teachers generally have positive attitude towards ICT.

Research Question 4: To what extent do lecturer uses ICT in the classroom to achieve learning outcomes?

The result in Table 5 shows the extent to which lecturers incorporated ICT into learning so as to achieve positive learning outcomes. The results in table 5 shows that items 4(47.1%), 5(54.3%) and 8(48.6%) were regularly used by the lecturers to achieve learning outcomes while items 1(42.1%), 2(53.6%) and 7(44.3%) were rarely used by the lecturers. Also, items 3(65.0%), 6(38.6%) and 9(65.7%) were never used by the lecturers. This result indicates that lecturers were not adequately using ICT to achieve learning outcomes.

Research Question 5: Do gender, age, teaching experience and teaching subjects have any effect on the ICT competence of the lecturers on the use of basic ICT application?

From the Table 6, at $P<0.05$ and degree of freedom 8, 131, the F_{cal} obtained for gender, age, teaching experience and teaching. Subject are 2.93, 6.01, 1.68 and 8.04 respectively while the F_{tab} obtained for gender, age, teaching experience and teaching subject are 3.37, 3.37, 3.02 and 3.02 respectively. This result indicates that the F_{cal} obtained for both age and teaching subject was greater than F_{tab} for the two demography variables while the F_{cal} obtained for gender and teaching experience is less than F_{tab} for the two remaining demography variables. Therefore, it can be conclude that both age and teaching subject of the lecturers' affect the ICT competence of the lecturers on the use of basic ICT application while both the gender and their teaching experience does not have any effect on the ICT competence of the lecturers on the use of basic ICT application.

V. DISCUSSION

The potentials of Information and Communication Technology (ICT) as an educational tool in teacher education had been well established by several studies [11].

This study examined the Information and Communication Technology (ICT) knowledge and skill levels of lecturers at Emmanuel Alayande College of Education, Oyo, Nigeria. Results from this study (table 2) revealed that six out of fourteen items (i.e. notebooks, interactive whiteboards, school intranet, specialist software applications, desktop computers for students' use in classroom and laptop computers for students) on the ICT infrastructural were not available in the college but seven items (i.e. desktop/laptop computer for personal use, personal e-mail account, internet, printer, digital cameras, digital projectors and photocopier machine) were available but not easily accessible while desktop computers for students' use in the computer laboratory was the only one that is available and easily accessible. The implication of this result is that the ICT infrastructure facilities are not adequately available for use by the lecturers and this will affect their competence and at the same time promote high computer anxiety.

Also, the findings as shown in table 3 revealed that lecturers do not adequately using ICT to achieve personal/professional objectives. For instance, three out of seven items were never used by the lecturer (i.e. use of ICT for curriculum administration e.g. planning, monitoring, evaluation, and reporting, use of ICT to post information to a website to assist students in their work and use of ICT for online professional learning) and lecturers rarely used ICT to create materials for students' use and to access research and best practices for teaching while lecturers only used ICT to communicate with colleagues/other professional. Also, the findings (table 5) revealed that lecturers do not adequately using ICT to achieve learning outcomes. For instance, three out of nine items were never used by the lecturers (i.e. use of ICT to express themselves clearly, use of ICT to analyze information and use of ICT in learning to work collaboratively). Lecturers rarely used ICT in mastering skills just taught, remediation of skills not learned and presentation of information to an audience. Lecturers regularly using ICT to communicate with other people, find out about new ideas and information and to improve their computer skills. This unforeseen result is on the fact that lecturers do not adequately have access to enough ICT infrastructures in the college and most of them had never received any training on the use of ICT for curriculum development and this will definitely affect their competence on the use of ICT to achieve professional objectives.

Results from table 4 also revealed that lecturers at Emmanuel Alayande College of Education, Oyo seem to have positive attitude comparable to that of lecturers in other countries [6], [2]. This study (table 6) also revealed that gender does not have any significant effect on the ICT skills and knowledge of the lecturers on the use of basic ICT application. The result (table 6) further indicated that age have significant effect on the ICT skills and knowledge of the lecturers on the use of basic ICT application. This finding corroborates that of [8] which shows that younger teachers appear to be less skeptical about the benefits of ICT in learning.

Table 2. Analysis of Results of the availability and accessibility of the school ICT capacity

S/N	Items	Available and easily accessible	Available but not easily accessible	Not available
1	Desktop/laptop computer for personal use	24 (17.1%)	110 (78.6%)	06 (4.3%)
2	Notebooks for Teachers program	0 (0.0%)	0 (0.0%)	140 (100.0%)
3	Interactive whiteboards	0 (0.0%)	0 (0.0%)	140 (100.0%)
4	Personal e-mail account	13 (9.3%)	112 (80.0%)	15 (10.7%)
5	School intranet	0 (0.0%)	0 (0.0%)	140 (100.0%)
6	Internet	04 (2.9%)	127 (90.7%)	09 (6.4%)
7	Printer	09 (6.4%)	112 (80.0%)	19 (13.6%)
8	Digital cameras	02 (1.4%)	127(90.7%)	11 (7.9%)
9	Specialist software applications (e.g. CAD, HTML editors, etc)	0 (0.0%)	0 (0.0%)	140 (100.0%)
10	Digital projectors	05 (3.6%)	121 (86.4%)	14 (10.0%)
11	Desktop computers for student use in the classroom	0 (0.0%)	0 (0.0%)	140 (100.0%)
12	Desktop computers for student use elsewhere at school (e.g. computer lab)	113 (80.7%)	07 (5.0%)	20 (14.3%)
13	Laptop computers for student	0 (0.0%)	0 (0.0%)	140 (100.0%)
14	Photocopier Machines	36 (25.7%)	81 (57.9%)	23 (16.4%)

Table 3. Analysis of Results on how often the lecturers use ICT to achieve personal/professional objectives

S/N	Items	Regularly	Sometimes	Rarely	Never
1	I am using ICT to create materials for students' use (e.g. handouts, tests)	26 (18.6%)	37 (26.4%)	49 (35.0%)	28 (20.0%)
2	I am using ICT to access research and best practices for teaching	38 (27.1%)	27 (19.3%)	56 (40.0%)	19 (13.6%)
3	I am using ICT for curriculum administration (e.g. planning, monitoring, evaluation and reporting)	0 (0.0%)	05 (3.6%)	59 (42.1%)	76 (54.3%)
4	I am using ICT to communicate with colleagues/other professional	121 (86.4%)	09 (6.4%)	06 (4.3%)	04 (2.9%)
5	I am using ICT to communicate with student(s) and / or students' parent(s)	12 (8.5%)	11 (7.9%)	68 (48.6%)	49 (35.0%)
6	I am using ICT to post information to a website to assist my students in their work	0 (0.0%)	08 (5.7%)	33 (23.6%)	99 (70.7%)
7	I am using ICT for online professional learning	20 (14.3%)	19 (13.6%)	29 (20.7%)	72 (51.4%)

Table 4. Analysis of Results of the attitudes of lecturers towards the use of ICT

S/N	Items	SA	A	D	SD
1	Student's use of ICT has the capacity to strongly support student-centred, inquiry based learning	87 (62.1%)	39 (27.9%)	09 (6.4%)	05 (3.6%)
2	ICT provides valuable resources and tools to support student learning	109 (77.9%)	11 (7.9%)	12 (8.5%)	08 (5.7%)
3	ICT provides students with efficient presentation and communication tools	91 (65.0%)	23 (16.4%)	22 (15.7%)	04 (2.9%)
4	ICT has limited capacity to provide benefits in the classroom	20 (14.3%)	22 (15.7%)	32 (22.9%)	66 (47.1%)
5	I like the challenge of exploring technology and new software and its possibilities	58 (41.4%)	49 (35.0)	19 (13.6%)	14 (10.0%)
6	Teacher education should include ICT	80 (57.1)	40 (28.6%)	11(7.9%)	09 (6.4%)
7	Student will learn more from ICT than from book	67 (47.9%)	45 (32.1%)	17 (12.1%)	11 (7.9%)
8	ICT can't address the needs of school system	41 (29.3%)	38 (27.1%)	21 (15.0%)	40 (28.6%)
9	The state of facilities discourages me from using ICT	69 (49.3%)	52 (37.1%)	13 (9.3%)	06 (4.3%)
10	ICT skill is worthwhile	80 (57.1%)	34 (24.3%)	15 (10.7%)	11 (7.9%)

Table 5. Analysis of Results on how often the lecturers uses ICT to achieve learning outcomes

S/N	Items	Regularly	Sometimes	Rarely	Never
1	Use ICT in mastering skills just taught	28 (20.0%)	32 (22.9%)	59 (42.1%)	21 (15.0%)
2	Use ICT for remediation of skills not learned well	13 (9.3%)	29 (20.7%)	75 (53.6%)	23 (16.4%)
3	Use ICT to express themselves clearly	08 (5.7%)	21 (15.0%)	20 (14.3%)	91 (65.0%)
4	Use ICT to communicate with other people	66 (47.1%)	35 (25.0%)	20 (14.3%)	19 (13.6%)
5	Use ICT to find out about new ideas and information	76 (54.3%)	39 (27.9%)	13 (9.3%)	12 (8.6%)
6	Use ICT in analyzing information	17 (12.1%)	25 (17.9%)	44 (31.4%)	54 (38.6%)
7	Use ICT to present information to an audience	05 (3.6%)	38 (27.1%)	62 (44.3%)	35 (25.0%)
8	Use ICT to improve their computer skill	68 (48.6%)	44 (31.4%)	19 (13.6%)	09 (6.4%)
9	Use ICT in learning to work collaboratively	14 (10.0%)	23 (16.4%)	11 (7.9%)	92 (65.7%)

Table 6: One way ANOVA of the Independent variables on Students access and utilization of both computer and Internet

Variables	Source of Variance	Degree of Freedom	Sum of the Squares	Mean of Squares	F_{cal}	F_{tab} α=0.05	Remark
Gender	Between Group	8	82.325	5.924	2.93	3.37	**NS**
	Within Group	131	89.419	3.572			
	Total	**139**	**171.744**				
Age	Between Group	8	16.236	4.237	6.01	3.37	**S**
	Within Group	131	74.264	2.293			
	Total	**139**	**90.500**				
Teaching Experience	Between Group	8	60.939	7.912	1.68	3.02	**NS**
	Within Group	131	287.379	4.748			
	Total	**139**	**348.318**				
Teaching Subject	Between Group	8	35.515	9.146	8.04	3.02	**S**
	Within Group	131	194.250	6.302			
	Total	**139**	**229.765**				

Also the result (table 6) revealed that teaching experience does not have any significant effect on the ICT skills and knowledge of lecturers on the use of basic ICT application. This findings contradicted that of [16] who asserted that younger less experienced teachers use computers more, because they are more likely to be computer fluent, had more technologically richer teacher training and are less likely to be limited by previous habits, perceptions or attitudes, than older teachers.

Finally, the study (table 6) revealed that teaching subject have significant effect on the ICT skills and knowledge of lecturers on the use of basic ICT application. This result is not unexpected because lecturers teaching science and science related courses acquainted the use of technology and also have high competence and confidence in the use of computer technology.

VI. CONCLUSION

In this study, it was discovered that lecturers have positive attitude towards the use of ICT. The results revealed that lecturers indicated high competency in the use of the internet resources than other basic ICT applications. However, the same lecturers lacked required competence in the use of spreadsheet, presentation, database and also the use of peripheral ICT equipment [11].

Gender and years of teaching experience had no statistically significant influence on the ICT skills and knowledge of the lecturers but both age and teaching subjects had significant influence statistically on the ICT skills and knowledge of the lecturers.

Finally, it was also established from the study that there is no adequate ICT infrastructural facilities in the college. Lecturers were not adequately using ICT to achieve personal/professional objectives and similarly they were not adequately acquainted with the use of ICT to achieve learning outcomes in the classroom.

VII. RECOMMENDATION

Based on the findings revealed from this study, the following suggestions are recommended:

i. The college management can provide adequate Internet and Intranet facilities for both lecturers and students in the college so as to enhance smooth communication between lecturers and their colleagues outside the campus.
ii. Adequate provision of printers, digital cameras, consumables and digital projectors to allow lecturers to create materials for class lecture.
iii. Training could be organized for the lecturers on how to integrate ICT within their learning area, development of lesson plan, using ICT to monitor, evaluate and report within the school and having a clear sense on direction for ICT.
iv. The management should ensure that they motivate lecturers' interest in ICT especially online learning.
v. The management should try to provide adequate ICT resources (such as interactive whiteboards, digital projectors, desktop computers, etc. for lecturers to use within the classrooms.
vi. Government could include in the universities, polytechnics and colleges of education curriculums the introduction of more courses on ICT with needed hand-on experiences so as to promote effective integration of ICT [11].

REFERENCES

[1] Beukes-Amiss, C.M. &.Chiware, E.R.T. (2006).The impact of diffusion of ICTs into educational practices, how good or how bad? A review of the Namibia situation. Retrieved on 12 November 2011 from: http://www.dspace.unam.na:8443/dspace/bitstream/1995/244/impact+diffusionICTedupdf

[2] Chai, C. S., Hong, H., & Teo, T. (2008). Singaporean and Taiwanese pre-service teachers' beliefs and their attitude towards ICT: A comparative study. *Paper presented at the 16th International Conference on Computers in Education*. Retrieved on 12 November 2011 from:http://www.apsce.net/icce2008/papers/ICCE2008paper19.pdf

[3] Derbyshire, H. (2003). *Gender issues in the use of computers in education in Africa*. Retrieved on 12 November 2011 from: http://imfundo.digitalbrain.com/imfundo/web/learn/documents/ Gender%20Report.pdf

[4] Fabry, D. and Higgs, J. (1997) 'Barriers to the effective use of technology in education', *Journal of Educational Computing Research*, 17(4), 385-395.

[5] Fowowe, S.S. (2006).Information technology: A veritable tools for sustaining the Universal Basic Education Programme in Africa. Paper presented at the African conference on Achieving Universal Basic/Primary Education in Africa in 2015. University of Botswana. 16- 19 October.

[6] Kennewell, S. & Morgan, A. (2003). Student teachers' experiences and attitudes towards using interactive whiteboards in the teaching and learning of young children. *Presented at the IFIP Working Groups 3.5 Conference: Young Children and Learning Technologies.* UWS Parramatta. Retrieved on 12 November 2011 from http://crpit.com/confpapers/CRPITV34Kennewell1.pdf

[7] Kirkpatrick, H. & Cuban, L (1998). Should we be worried? What the research says about gender differences in access, use, attitudes, and achievements with computer. *Educational Technology, 38* (4), 56-60.

[8] Korte, W. B. and Husing, T. (2007) 'Benchmarking Access and Use of ICT in European Schools 2006: Results from Head Teacher and A Classroom Teacher Surveys in 27 European Countries', *eLearning Papers*, 2(1), 1-6.

[9] Kozma , R. (2005) *"National Policies that Connect ICT-Based Education reform to Economic and Social Development"* Human Technology [Electronic], 5(4): 358-367. Available at: < www.humantechnology.jyu.fi/current/abstracts/kozma05.html >[Accessed on 12 November 2011]

[10] Kozma, R., McGhee, R., Quellmalz, E., & Zalles, (2004). Closing the digital divide: evaluation of the World Links Program. *International Journal of Educational Development, 25*(4), 361-381.

[11] Mudasiru, O. Y and Modupe, R. B (2011). Student-Teachers' Competence and Attitude towards Information and Communication Technology: A Case Study in a Nigerian University. *Contemporary Educational Technology*, 2(1), 18-36 .

[12] Olakulehin, F.K. (2007).Information communication technologies in teachers training and professional development in Nigeria. *Turkish Journal of Distance Education TODJE 8, (1), 133- 142.*

[13] Sara,. H. ,David, H. , & Leaonard, W. (2010). Teacher Factors Influencing Classroom Use of ICT in Sub- Sahara Africa. *Itupale Online Journal of Africa Studies*, 2, 39-54.

[14] Schoepp, K. (2005) 'Barriers to Technology Integration in a Technology-Rich Environment, *Learning and Teaching in Higher Education: Gulf Perspectives*, 2(1), 1-24.

[15] Tella, A., Tella, A., Toyobo, O. M., Adika, L. O., & Adeyinka, A. A. (2007). An Assessment of Secondary School Teachers Uses of ICTs: Implications for Further Development of ICT's Use in Nigerian Secondary Schools. *Online Submission, 6*(3).

[16] UNESCO, (2004). ICT pedagogy. UNESCO office.

[17] Young, B. J. (2000) 'Gender differences in student attitudes toward computers', *Journal of Research on Computing in Education*, 33(2), 204-217.

Mobile Learning with Real-Time Video Service over Ad-Hoc Network

Onwodi Gregory
School of Science and Technology
National Open University of Nigeria
Lagos, Nigeria
gregonwodi@hotmail.com

Adeyinka Adekunle
School of Computing and Engineering
Babcock University
Ogun, Nigeria
adekunleya@gmail.com

***Abstract*-** The demand for education has continued to increase, unabated. Consequently there is a corresponding need to identify and design more affordable ways or models for increasing access to learning opportunities. Coincidentally, similarly being witnessed is an unprecedented explosion in the types of mobile devices being produced, particularly mobile phones. The technology which is arguably the commonest means of communication, could play a pivotal role in extending the possibilities for teaching, learning and research in formal and informal settings. Published research works have suggested prospects and technological challenges of mobile learning. This paper reviews mLearning models and evaluates implementing mlearning with real-time video service over ad-hoc network. Our result shows that packet loss was decreased when the Dynamic Source Routing (DSR) protocol with extension was used. In real-time video a lost packet will cause error propagation into following frames and the number of frames having a low quality was therefore significantly much higher than the number of lost frames.

Keywords: mLearning, DSR, Mobile Devices, ad-hoc network, real-time video service

I. INTRODUCTION

Mobile learning has been simply defined as learning that takes place with the help of mobile devices Quinn (2003). Similarly, Traxler (2005) commented that mobile learning can be defined as educational provision where the sole or dominant technologies are handheld or palmtop devices.

Historically, Sharples (2002) singled out Alan Kay's Dyna book conceived in the early 1970s, as the first serious attempt to design a computer-mediated mobile learning platform. Basically, mobile learning deals with the use of mobile or wireless devices for the purpose of learning while in motion. It is also said that electronic organizers capable of three line text-only display were available in the 1990s. Palm Pilot PDAs, introduced in 1996, were the first multi-purpose, customizable handheld devices suitable for a range of creative learning activities. The grant of more than 100 palm education pioneers (PEP) awarded in 2001 by SRI International (USA) to US teachers who had a vision of how Palm handhelds could be used to improve teaching and learning was another milestone. Many of the findings of the PEP grants have been confirmed by later "handheld learning" studies. Examples of pertinent findings include the strengths and weaknesses of various models for allocating handheld computers to students, to the degree of success with which various learning activities can be accomplished using handheld devices (Saipunidzam M. et al (2010)).

Ellen Wagner (2005) proclaimed that the mobile revolution had finally arrived:

> "Wherever one looks, evidence of mobile penetration is irrefutable: cell phones, PDAs, MP3 players, portable game devices, handhelds, tablets, and laptops abound. No demographic is immune from this phenomenon. From toddlers to seniors, people are increasingly connected and are digitally communicating with each other in ways that would have been impossible only a few years ago".

Recently, many researchers have focused on m-Learning and its environment, such as users' acceptance of m-Learning (Phuangthong & Malisawan, 2005; Liu, 2009), setting the environment for m-Learning (Chao & Chen, 2009; Brown et al., 2006; Liu, 2008), and the application of m-Learning in developed countries (Paul, 2008). Simil arly, several researches have been carried out in developing countries as well. In Saudi Arabia, a recent paper reports on the results of the survey conducted at King Saud University as reported by chanchary (2010), the author attempted to determine how the mobile technology can be utilized to improve student's retention at bachelor degree. The result indicated that the introduction of mobile teaching and learning can enhance the teaching and learning situation in that country. Although most of the developing countries are still in the first phase or perhaps in the research and development phase in implementing this type of learning environment, Kyun Baek and Uk Cheong as well as Barker, Krull and Mallinson as reported by Saipunidzam Mahamad et al. (2010) had proved that developing countries will soon catch up with this new learning paradigm.

This shows that this new learning paradigm will evolve mobile devices with the rapid usage and ownership among the users.

In Nigeria, mobile learning have been experimented by tutors and experts through partnerships between the University of Ibadan and Educational Advancement Centre to guarantee outstanding results in the Join Admission and Matriculation Board (JAMB UTME) available for secondary school students (SS1, SS2, SS3 and retake students) (www.mobi**learn**.mobi/index.php). Research on mobile learning has emphasized and repeated the advantage of using mobile tech-nology because of the characteristics and opportunities that mobile learning offers. Common characteristics that are brought up in many research papers around m-learning are those of ubiquity, nomadicy, personalization, interactivity, and collaboration. The technology itself allows for these characteristics because of its size, weight and portability. The small screen size has been noted as a drawback. Being a part of most people's daily lives, mobile technology doesn't need to be taught and so learning can be integrated quickly and easily into a persons' everyday life (Al - Hmouz, Shen, Yan, & Al-Hmouz, 2010; Cavus & Uzunboylu, 2009; Ozdamli & Cavus, 2011; Rogers, Connelly, Hazlewood, & Tedesco, 2009).

II. SIGNIFICANCE OF STUDY

A seemingly effective way of rapidly building a critical mass of the necessary human capacity is through a distance learning scheme. It offers the best opportunity for African countries to take their rightful place in the global economy. It is now recognized that distance education is an additional method to provide quality education to a large number of students spread over wide geographical areas in a short period of time (Akinyede, 2010)

M-learning is gradually taking a central stage, and this work will facilitate the implementation of an M-learning System via satellite, this will increase access to education for several learners across the nation, including the nomads. Consequently a large mass of human capacity/capital would be developed. This project if implemented will facilitate access to education and educational resources anytime anywhere and consequently enhance the expansion and spread of knowledge globally.

III. STATEMENT OF THE PROBLEM

There is an endemic crisis in Sub-Saharan Africa's teaching and learning development systems. This can be seen in the shortage in the number of qualified teachers and in the depth of motivated teachers who are capable of delivering quality teaching and learning within a twenty-first-century educational context. The infrastructural deficit associated with the region as well as the level of illiteracy poses great danger to the educational and general development of the continent. It is widely believed that addressing these challenges requires a range of interventions. One of such intervention includes the integration and use of Information and Communications Technology (ICT) and Open and Distance Learning (ODL) opportunities. The use of e-learning strategies as a means of addressing these issues is a step in the right direction especially in ODL environment. But these strategies have not sufficiently broadened access to education especially to the rural dwellers in Nigeria. Access to internet technology, high cost of computer systems, non portability of computer systems etc remain obstacles to e-learning. In Nigeria, the problem of accessing education poses a major challenge to development. Lack of adequate ICT (Information and Communications Technology) infrastructure hampers the efforts of governments to give citizens some essential services, such as education. Some states in Nigeria have terrains that make it difficult for Governments and Non-Governmental Organizations (NGOs) to make delivery of conventional education easy. Some places are not easy to reach in the highlands therefore resulting in lack of essential basic services as electricity which could be used to power some teaching aids e.g. computers. The limitation of conventional learning locations reduces access to knowledge and formal education. Nigerian's population is about 170 million, greater percentage of this population are young adults. Very few of these Nigerian youths have access to higher education due to difficulty in gaining admission into the limited tertiary Institutions available and cost of conventional education.

The global geometrical growth of mobile phone subscriptions in recent years has sparked interest in how mobile phones in particular might enhance ODL opportunities for the professional development of teachers, their support in delivering teaching in their pedagogical practices and administrative duties as well as development of teachers, their support in delivering teaching in their pedagogical practices and administrative duties as well as the delivery of learning to the students.

IV. REVIEW OF LITERATURE

Mobile learning is a new educational paradigm which is more flexible than learning using the desktop computers. It is moveable from one environment to another such as classroom to outdoors and vice versa. Hence, a seamless learning potential is realized, where the technology mediates between the learners and the learning content; the learning is flexible and adaptable so that teachers and students are not bound to a particular learning space. Learning while being mobile and through the use of mobile devices (such as Mobile phones, PDAs, iPhone, iPod, Palmtop, Laptop or even digital cameras) are considered to be

independent of time and location, as it could occur at any time and in any place. It also provides access on demand learning content to learners.

Mobile learning technology is portable as the name suggests and it is wireless. Learners are able to move around with these learning tools. Mobile learning can be viewed as the focal point, where mobile technologies and web-based learning intersect to offer anywhere anytime instant on-demand educational information. It can be generally defined as the acquisition of any knowledge and skills through the use of mobile devices anywhere anytime that results in an alteration in behavior. According to (Brown,2005) "mobile technologies have the power to make learning even more widely available and accessible than we are used to in existing web-based learning environments".

According to the Nigeria Communication Commission report of 2011, more than 90% of Nigerians own at least one mobile phone and mobile phone usage peaks in the range from 16 to 45 years, where 70% of this age group regularly uses a mobile phone. The majority of Nigerian University Students fall in this age group thus making mobile phones the most commonly used device among University Students. Therefore, National Open University of Nigeria should take up the challenge to use these mobile devices, specifically mobile phones to provide and offer learning services. New technologies are constantly being developed and produced; examples of different mobile devices that could be used for learning today include; Mobile phones, iPods, Mp3, Mp4, PDAs, GPS, and Organizers to mention a few. With the rapid advancement of new technologies, you generally will only need one mobile device which will be capable of integrating many features to support learning. Therefore, learners need not have multiple devices but rather one mobile device which will be able to meet their needs, as a standard phone in today's market is equipped with much functionality which if used effectively, can make learning an easy process. Most hi-tech mobile phones, currently available like the 3G iPhone, the Android, the Google G-1 phone or any smart phone, contain the functionalities to help the users perform many tasks and allow them have access to information as required anytime and from anywhere. This access to information is normally achieved in an interactive way, where the mobile phone user feels in control of handling the accessibility, management and sharing of resources. Today mobile devices have some unique features which make them an efficient device capable of providing, sharing and exchanging of learning content. Some of these feature as are as follows:

- Portability and Mobility: Mobile devices are small size lightweight devices, students find them convenient to use and carry with them.
- Flexibility: Students and teachers can access the server from almost anywhere.
- Convenience: Whenever students need to connect teachers or urgently need information from the internet, remote access is available to them. In particular accessing the server through WAP (Wireless Application Protocol).
- Remote Accessibility: Students can add or update information remotely.
- Ease of Use: Using an iPhone provides a larger screen than an ordinary mobile phone and it's convenient for learners input.
- Utility: Most of the smart phones provide almost all of the capabilities of a standard computer such as processing and storing data.

V. MLEARNING MODELS

Table 1: mLearning models: strengths and weaknesses

Model Name	Strength	Weakness	Remark
Model for m-learning adoption NEduSat).	It contains a wireless access point which enables communication between the mobile devices. Learning can also occur through desktop PCs. Mobile devices can be used as academic support for learners via online assessment, providing course content and access to the Internet.	mLearning operates in traditional learning environment	This limits the operational mode or range of the mLearning model
Information System (IS) model	The success of an IS can be represented by the quality characteristics of the IS itself (system quality) the quality of the output of the IS (information quality); consumption of the output of the IS (use); the IS users' response to the IS (users' satisfaction); the effect of the IS on the behavior of the user (individual impact); and the effect of the IS on organizational performance (organizational impact)	The IS effectiveness performed in the traditional data processing environment cannot be used directly in the m-learning environment in the Delone and McLean's model which is an updated IS model.	There is no interoperability between the IS model and the Delone and MacLean's model. (which is an updated version of the IS model)

Model Name	Strength	Weakness
Shih's mobile learning model	The model was created to support instructional design for mobile learning. The learning cycle of the ARCS include: Attention, Relevance, Confidence, and Satisfaction (ARCS). 1. Sendings multimedia message to mobile phones to trigger and motivate learners. 2. Searching the Web for related information by using embedded hyperlinks (URLs) in the message received in the phone. 3. Discussing with learning peers by text, voice, picture, or video messaging 4 Producing a digital story telling of what they learn by audio or video diary. 5. Applying what they learn in the simulated environment, such as online educational gaming.	This learning model mainly relies on the mobile computing communication infrastructure, and would be most suitable for applications in blended learning environments
A model for mLearning in Africa	Downloading of content access to articles/study materials/other resources e-mail/bulletin board/chat room (communication and interaction) Learners use mobile phones on a regular basis Academic support for learners via SMS communication and interaction. Administrative Information (reminders, notifications, urgent information, etc) access to examination and test marks via mobile service number, etc.	Learners only have periodic access to the Internet via PCs at learning or community centres
Pedagogical model developed for mobile tutoring	The model is focused on the characteristics of mobile learning. AEFIRP stands for Activation, Externalization, Focusing, Interpretations, Reflection and Information Processing. Here mobile technology is a trigger and platform that includes guidance and support for learning methods and the learning process.	
Älykkö("egg head").	It consist of intelligent tutoring agents that provides semi-automatic and automatic tutoring as well as indicators based on student's learning process enabling individualized tutoring process for students, even without teacher's virtual presence.	

VI. THE PLATFORM

Here a platform for providing a practical and reasonable priced communication network is presented. This platform is based on wireless ad-hoc networks and can be used to provide high quality m-learning. The platform consists of a set of nodes that are connected by wireless links. In this network the nodes are free to move and the topology of the network may therefore change. The nodes can forward data for each other in multiple hops and the path between two nodes is set up dynamically depending on the connectivity between the nodes and without the support of any central coordinator. The coverage of this network is increased when more users are added to the network. This type of network is very robust since nodes can be added or removed from the network dynamically and the routing will adjust rapidly to the new conditions. Even though the capacity for end to end communication is decreased when data is forwarded in multiple hops (Jinyang, 2001) the capacity provided by an ad-hoc network is high enough to enable real-time video communication.

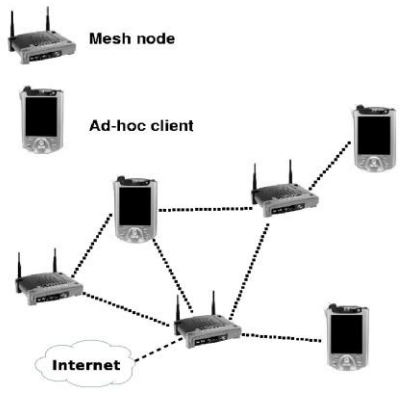


Fig. 1: An overview of a network consisting of mesh nodes and mobile ad-hoc clients.

Our network is built up using two different types of nodes, see Fig.1. One type of nodes is the ad-hoc clients in the network. These nodes can be for example a laptop equipped with an IEEE 802.11 wireless interface, a mobile phone or Tablet equipped with WiFi. This type of nodes can relay traffic for other nodes, but also be sink or source for traffic. They are both routers and hosts. These nodes are mobile and can also be switched on and off randomly thus creating fast changes in the network topology. The other type of nodes is mesh nodes that will only relay traffic for other nodes. They are needed to guarantee connectivity even at times when the number of mobile clients is few. These nodes can be located at rooftops or other positions providing a high level of coverage.

A. Mesh nodes

The requirements for the mesh nodes are a CPU capable of running the routing protocol, a wireless communication interface and possibilities to harvest energy from the environment, and a battery to compensate for the variation in energy collected from the environment. The node should be able to operate even if there is no available connection to a power line. The mesh nodes used in this test-bed are built using a Linksys WRT54GL router, a solar panel and a battery, see Fig.2 for a picture of a deployed mesh node. All components are high volume consumer products to keep the cost for a node as low as possible. The Linksys WRT54GL router has a 200MHz Broadcom 5352 CPU, 4MB Flash, 16MB RAM and an IEEE 802.11 b/g wireless chipset. It is capable of running a Linux distribution enabling the possibility to add customized ad-hoc routing protocols. It uses a switched power regulator and at 12v the measured current was 240mA.

Fig. 2: A picture of a deployed mesh node.

B. Ad-hoc clients

For the mobile ad-hoc clients node a laptop with a 7.5“ TFT display, a VGA camera, and a wireless network card capable of creating an ad-hoc network even when the CPU is off is suitable. The Classmate PC from Intel is also developed for this same purpose. Mobile Phones or Tablets with Wi-Fi can equally be used.

C. Routing

Sending real-time video over wireless ad-hoc networks is a challenging problem. Video is very sensitive for packet loss and wireless ad-hoc networks are error prone due to node mobility and weak links. The most popular routing protocols today for ad-hoc networks, for example Dynamic Source Routing (DSR) (David B. J, 1996) and Ad-hoc On Demand Distance Vector (AODV) (Perkins, 2003) are reactive routing protocols that are focused on scalability for very large networks and to reduce the routing load (Perkins, 2001). The major problem when sending real-time video over reactive routing protocols is however the long interruption in packet delivery when a route breaks and a new route has to be established. This is a drawback of the reactive design. To achieve high robustness for real-time video over ad-hoc networks issues like the delay and packet delivery ratio are very important. We have developed a preemptive routing protocol by adding some extensions to standard DSR. We call this version of DSR for real-time DSR (RT-DSR). There are basically two problems with the standard DSR that this new
routing protocol solves. The first problem is that DSR will not search for new routes until a route breaks. During the time it takes to discover new routes no data can be delivered. This is an unacceptable delay for the real-time video. The other problem is the gray-zone problem (Henrik, 2002). The standard DSR routing protocol often discovers routes that contains very weak links. The route packets are broadcasted and they are relatively small compared to the data packets. We start with the most basic form of DSR. The following optimizations were removed from DSR:

- Automatic route shortening
- Caching overheard routing information
- Replying to route requests using cached routes
- Overhearing of packets
- Adding routes to intermediate nodes
- Gratuitous Route Replies
- All nodes process all of the Route Error messages they receive

If the event of routes that are about to break can be discovered it is possible to establish and start to use new routes before the old routes break. By simply monitoring the signal strength for each data packet we can detect if a route is about to break. To avoid using very weak links when establishing new routes the Route Request (RREQ) packets are filtered based on received signal strength.

The following two extensions were added to DSR: • Send a special ”weak RERR” packet when a data packet having low signal strength is routed • Signal strength threshold for RREQ packets. The modifications to the data structures, packet formats and algorithms are described in the following sections.

D. Data structures

When the source node receives information about a weak link on a route this route should still be used until new routes have been discovered. The old routes must however be removed after the new routes have been discovered. If this is not done, the old routes will be continued to be used until a packet is dropped on the route. We therefore add a 1-bit flag, weak Signal, to the route cache. This flag is used to mark the routes to be deleted after new routes have been discovered.

E. Packet formats

A control packet is used to inform the source of the route when a weak link on the route is detected. If we use a regular RERR packet all intermediate nodes and the source node receiving the RERR packet will remove all routes in its route cache containing the weak link. This should not happen until new routes have been discovered. Instead of creating a complete new control packet a weak Signal flag is added to the RERR packet. The RERR packet has an 8-bit Error Type field and a 4-bit reserved field that can be used for this.

F. Route maintenance

A RERR packet with the weak Signal flag set is initiated every time a node receives a data packet having signal strength below −80 dBm. When the source or intermediate node receives this RERR packet they will not delete any routes in the route cache. When the source receives this RERR packet it will first set the weak-Signal flag for all routes in the route cache for this destination. The source node will then initiate a route discovery by sending a Route Request (RREQ) packet. When the node receives a Route Reply (RREP) it will remove all routes in its route cache having the weak-Signal flag set for the given destination.

G. Route discovery

When a route discovery is performed it is desired to avoid discovering routes containing weak links. Because the control packets are small and broadcasted they are less error prone than the data packets. If all RREQ packets are forwarded it is likely that many of the discovered routes will contain

links having a very high packet loss. This is referred to the gray-zone problem in wireless ad-hoc networks. We therefore discard RREQ packets received with signal strength below −80 dBm.

H. Simulations

The GloMoSim 2.03 (Karlsson 2002) network simulator is used in our simulations. This is an event driven network simulator developed at the UCLA parallel computing laboratory. We have one version for the original DSR routing protocol and one version for the modified RT-DSR routing protocol. We also compared our results with the AODV (Perkins, 2003) routing protocol.

For all the simulations we used the following settings:

PROPAGATION-LIMIT -111.0
PROPAGATION-PATHLOSS
TWO-RAY NOISE-FIGURE 10.0
TEMPARATURE 290.0
RADIO-TYPE RADIO-ACCNOISE
RADIO-FREQUENCY 2.4e9
RADIO-BANDWIDTH 11000000
RADIO-RX-TYPE SNR-BOUNDED
RADIO-RX-SNR-THRESHOLD 10.0
RADIO-TX-POWER 15.0
RADIO-ANTENNA-GAIN 0.0
RADIO-RX-SENSITIVITY -91.0
RADIO-RX-THRESHOLD -81.0
MAC-PROTOCOL 802.11

The two-ray path loss model was used (3). This is the maximum of the plane earth (1) and the free space path loss (2),

$$P = \frac{d^2}{t \times r} \quad (1)$$

$$S = \frac{4 \times \pi \times d}{\lambda} \quad (2)$$

$$T = \max(P, S) \quad (3)$$

Where d is the distance between the sender and receiver, t is the tx antenna height, r is the rx antenna height and λ is the signal wavelength. In GloMoSim both the rx and tx antenna height is 1.5 meters and the wavelength in this simulation was set to 0.125 meters. When the distance is 356 meters the received signal strength will be −80.01 dBm and the link will be detected as a weak link. When the distance is 377 meters the received signal strength will be −81.01 dBm, this is below the radio rx threshold and the packet will be dropped.

In our simulations we used the theora video codec and the ogg container format. We used two different videos in our simulations. One was the commonly used foreman video sequence in qcif resolution (176x144 pixels). The original video clip was 400 frames at 30 frames per second and we looped this sequence 45 times to create a 600 second and 18 000 frames long video clip. The other video we used was the first 600 seconds from the opening session of the biomedia course given within the ICT-LEAP consortium. This course is given for five universities over video conferencing networks. We used the recorded video from a video conference session and encoded the video from one of the parties at qcif resolution. For both of the video clips the video was encoded using a target bitrate 100 kbit/s, keyframe interval 120 and 30 frames per second. The actually encode bitrate, including the ogg headers, was 157.1 kbit/s for the foreman sequence and 101.4 kbit/s for the biomedia course sequence. First a trace file was created. This file contained the size of each packet and the time for transmission. This file was used in the GloMoSim simulations and the lost packets were recorded. This information was then used when the video was decoded. If a packet was lost or had a delay of more than 50 ms the packet was dropped at decoding.

The error concealment method used at the decoder was to copy the previous successfully decoded frame if a frame was lost. To compare the new routing protocol to AODV and standard DSR we run three different simulation scenarios. In each simulation we recorded the number of lost packets, the average PSNR, the number of frames having a PSNR below 30 dB and the total number of routing packets transmitted. An error free transmission had an average PSNR of 34.87 dB for the foreman video sequence and 40.79 dB for the biomedia course video sequence. The foreman video sequence had much more motion, including camera motion, compared to the video sequence from the biomedia course and since the video was coded using constant bitrate the encoded quality of the forman video was lower.

I. Roaming node

In this scenario we forced a number of route breaks at the last hop. We used a total of seven nodes. The nodes were placed on a row having a distance of 200 meters between them, see Fig. 3. The source and the destination node had the same initial position. The destination node was then moving along the row of nodes at a speed of 2 m/s until it was located 200 meters to the right of node six. First there was a one-hop link between the source node and the destination node. When the destination node was located between node two and node three the one-hop route was lost and a two-hop

route had to be established. Similarly the two-hop, three-hop, four-hop and five-hop route was lost. Finally before the destination node reached its final destination 200 meters to the right of node six the five-hop route was lost and a six-hop route had to be established. There were a total of five route breaks in this simulation. The simulation time was 600 seconds and we run the simulation ten times for each video sequence with different initial seed.

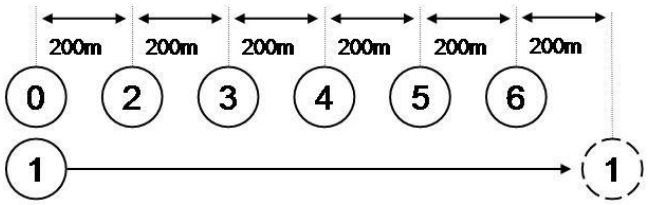


Fig. 3: The roaming node scenario.

Table 2: Route Performance in the Roaming Node Scenario ar
Foreman Forman Video Clip

	RT-DSR	DSR	AODV
Lost packets	0.00%	11.05%	0.30%
PSNR (dB)	34.87	25.68	34.10
PSNR < 30dB	0.00%	14.07%	1.20%
Control packets	149	5004	61

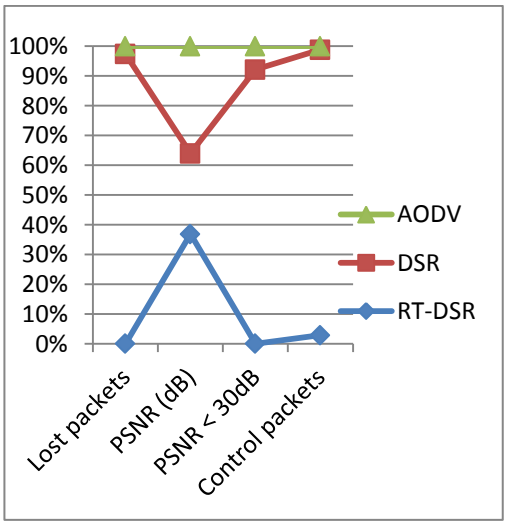


Fig.4: chart of Route Performance in the Roaming Node Scenario and Forman Video Clip.

Table 3: Route Performance in the Roaming Node Scenario and Biomedia Course Video clip.

	RT-DSR	DSR	AODV
Lost packets	0.00%	11.52%	0.29%
PSNR (dB)	40.79	30.89	40.26
PSNR < 30dB	0.81%	12.97%	0.92%
Control packets	148	5224	61

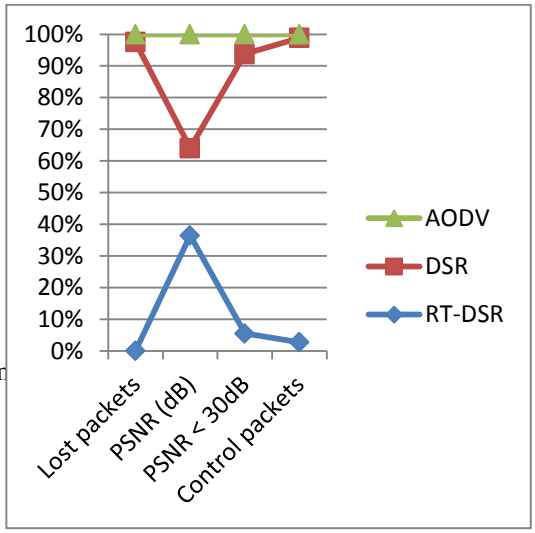


Fig.5: chart of Route Performance in the Roaming Node Scenario and Biomedia Course Video clip.

In table 2, Fig.4 and table 3, Fig.5; we can see that RT-DSR had much lower packet loss compared to standard DSR for both video sequences. Most of the packet losses when DSR was used occurred at the last route breaks. The RT-DSR routing also outperformed the AODV routing protocol for both video sequences, the difference was however much smaller compared to DSR.

Random waypoint

In this simulation we let a number of nodes move randomly within an area. We used a very high node density to make the network connected with high probability. This is because we did not want to have dropped packets because there were no routes available between the source and the destination. We used 60 nodes distributed within an area of size 1000x1000 meters. The nodes were initially distributed using uniform node placement. The mobility model used was random waypoint. Both the max and min speed were set to 4 m/s. The waypoint pause time was set to 1 second. The simulation was run for 600 seconds and each simulation was repeated ten times for each video sequence with different initial seed.

In table 4, Fig. 6 and table 5, Fig. 7; we can see that the RT-DSR had less dropped packets compared to both DSR and AODV. Although we used high node

density it is likely that some of the packets were lost because no route existed between the source and the destination. The RT-DSR routing protocol had much higher routing overhead compared to both DSR and AODV in this simulation.

Table 4: Route performance in the Random Waypoint Scenario and the Foreman Video Clip

	RT-DSR	DSR	AODV
Lost packets	0.04%	3.25%	0.15%
PSNR (dB)	34.77	31.05	34.37
PSNR < 30dB	0.26%	4.72%	1.03%
Control packets	1483	751	382

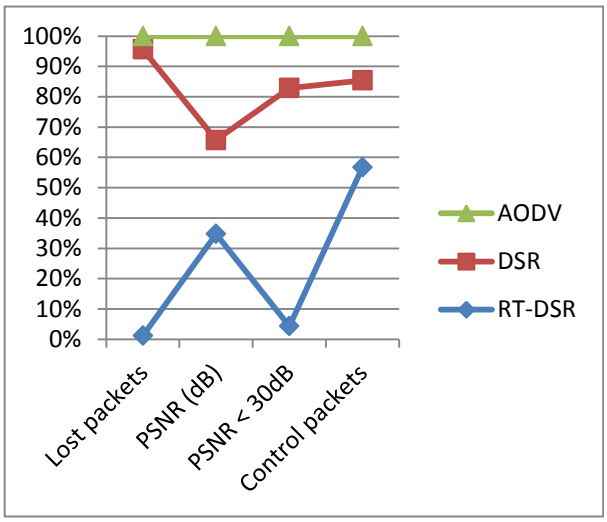


Fig.6: chart of Route Performance in the Random Waypoint Scenario and the Foreman Video Clip

Table 5: Route performance in the Random Waypoint Scenario and the Biomedia Course Video Clip

	RT-DSR	DSR	AODV
Lost packets	0.03%	2.71%	0.17%
PSNR (dB)	40.71	38.10	40.15
PSNR < 30dB	0.82%	2.87%	0.97%
Control packets	1778	838	422

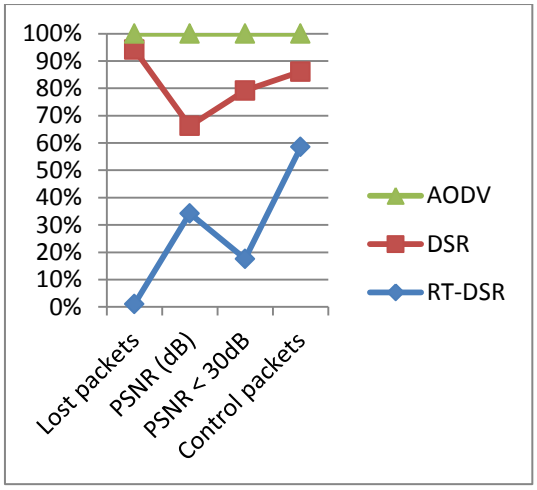


Fig.7: chart of Route Performance in the Random Waypoint Scenario and the Foreman Video Clip

VII. CONCLUSION

mLearning can be seen as a combination of mobile technology and its affordances that create a unique learning environment and opportunities for learning that can span across time and place. Mobile technology is growing at a rapid rate around the world, and becoming ever more accessible. Technology is emerging and closing the gap towards being a viable tool for mobile learning. It's a social platform that creates environments for communication, understanding, and transfer of information. Its ability to cater for varying learning styles through various features, its link to communication and its social context make it a very attractive tool for learning. It is important that mLearning exists within pedagogy and that designers have an understanding of the contexts that learners exist in. mLearning can cater for various learning styles through varying delivery methods. Learning is knowledge that is transferred through communication and mLearning has the ability to support varying learning styles. In this paper we have presented a possible solution for providing high quality mLearning to people living in areas with no infrastructure for Internet and mobile communication or where the cost to use the existing infrastructure is too high. This solution is based on wireless ad-hoc networks. The hardware platform presented is built using low cost high volume consumer products.

A preemptive ad-hoc routing protocol is also presented in this paper. This routing protocol was developed to increase the real-time video performance in ad-hoc networks. Our result shows that the packet loss was decreased when our routing protocol was used. In real-time video a lost packet will cause error propagation into following frames and the number of frames having a low quality was therefore significantly much higher than the number of lost frames. Applying this system in practice

would probably make it possible to reach everybody interested in education independent of the place they live or surrounding environment.

VIII. RECOMMENDATION AND FUTURE WORK

In order to reduce packet loss and increase the real-time video performance in any ad-hoc network that is set-up to implement an mLearning system or otherwise, we recommend using a preemptive ad-hoc routing protocol such as real-time DSR (RT-DSR) which is developed by adding some extensions to standard DSR (Dynamic Source Routing).

From the review of mLearning models, it is obvious that mLearning is still at its infancy, and researches done in this field is still at individual or group basis. There is not yet, a clear standard model. In future work, an enhanced model for mLearning for low bandwidth networks that uses intelligent mobile agents would be appropriate.

REFERENCES

[1] Adesope, O., Olubunmi, S. and McCracken, J. (2007). Implementing MLearning in Developing Countries: Prospects and Challenges. In C. Montgomerie & J. Seale (Eds.), Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2007 (pp. 1249-1254). Chesapeake, VA: AACE.

[2] Boyinbode O.K, Akinyede R.O.(2008). Mobile learning: An application of mobile and wireless technologies in Nigerian learning system. International Journal of computer science and network security. Vol. 8 No 11, 15-18.

[3] Brown, T.H. (2005). Towards a model for MLearning in Africa. International Journal on E-Learning ISSN 1537-2456 Volume 4, Issue 3.

[4] Chen, G. D., Chang, C. K., & Wang, C. Y. (2008). Ubiquitous learning website: Scaffold learners by mobile devices with information-aware techniques. *Computers & Education*, *50*(1), 77–90. doi:10.1016/j.compedu.2006.03.004.

[5] David B. J and David A. M, (1996), "Dynamic source routing in ad hoc wireless networks," in Mobile Computing, vol. 353. Kluwer Academic Publishers, 1996.

[6] Ghassan F. I, Shakir M. H, and Hussein, A (2011), "Unified MLearning Model through Interactive Education Satellite: A proposal for an Arab Homeland Education Satellite". doi:10.3991/ijim.v5i2.1600.

[7] Jinyang L., Charles B., Douglas S. J., De-Couto, Lee, H., and Robert M., (2001) "Capacity of ad hoc wireless networks," in Proceedings of the 7th ACM International Conference on Mobile Computing and Networking, Rome, Italy, July 2001, pp. 61–69.

[8] Johannes K., Adi A., and Haibo L. (2009) Enabling Real-Time Video Services Over Ad-Hoc Networks Opens the Gates for E-learning in Areas Lacking Infrastructure, doi:10.3991/ijim.v3i4.911,9 Oct 2009.

[9] Karlsson, A., Anani, Li H, Henrik L, Nordstrm E, and Tschudin C (2002) "Coping with communication gray zones in ieee 802.11b based ad hoc networks," in 5th ACM international workshop on Wireless mobile multimedia (WoWMoM 2002. 2002, pp. 49–55, ACM Press.

[10] Kearney, M., Schuck, S., Burden, K., & Aubusson, P. (2012). Viewing mLearning from a pedagogical perspective. *Research in Learning Technology*, *20*(1063519), 1–17. doi:10.3402/rlt.v20i0/14406.

[11] Keller, J. (1987), Development and use of the ARCS model of motivational design . Journal of Instructional Development, 10(3), 2-10.

[12] Olufemi, M. (2008). Improving MLearning with enhanced Shih's model of Mobile Learning. US-China Education Review, 5(11). pp 22-28.

[13] Perkins C, Belding-Royer E, and Das S, 2003, "Ad hoc on-demand distance vector (aodv) routing," RFC 3561, July 2003.

[14] Saipunidzam M., et al. (2010). M-learning: a new paradigm of learning Mathematics in Malaysia International Journal of Computer Science & Information Technology (IJCSIT) Vol.2, No.4, DOI : 10.5121/ijcsit.2010.2407, 76-78.

[15] Sharples, M., Taylor, T. & Vavoula, G. (2005). Towards a Theory of Mobile Learning. International Journal of Continuing Engineering Education and Life Long Learning, 12(5/6), 504-520.

[16] Udanor, C. & Nwodoh, T. (2011). A Review of MLearning Models. Indian Journal of Computer Science and Engineering, Vol 1 No 4 426-435.

Enhancing Security on Mobile Banking through Biometrics

Akanni Adeniyi
Computer Science Department,
Babcock University, Ilisan-Remo,
Ogun State, Nigeria
adeniyiakanni@gmail.com

Idowu Sunday A.
Computer Science Department,
Babcock University, Ilisan-Remo,
Ogun State, Nigeria
saidowu06@hotmail.com

Awodele Oludele
Computer Science Department,
Babcock University, Ilisan-Remo,
Ogun State, Nigeria
delealways@yahoo.com

***Abstract*-**Cashless Policy is gaining grounds in recent times. The list of nations embarking on the need to reduce cash grows from time to time. As more countries join the league of cashless policy crusaders, alternative channels, especially mobile banking is bound to witness an upsurge in patronage all over the world in years to come. Some security measures are already in place with some level of success to mitigate against the risk of identity theft. Personal Identification Number (PIN) is being used to initiate transactions. However, it has proven insufficient. This only suggests that the security on mobile transactions has to be taken a step further. This paper offers an additional security through biometrics to protect customers, banks and the banking industry from losses emanating from identity theft. This hybrid solution uses the existing PIN and fingerprint to authenticate mobile banking transactions.

Keywords—PIN, ATM, biometrics, I & A, cashless, mobile banking, SPOF, IdM

I. INTRODUCTION

Banking has taken a new dimension in recent times. This is however, not unconnected with the technological advancement witnessed in all spheres of human endeavors. Banking is not different. It has also undergone rapid changes from traditional banking to alternative channels such as Automated Teller Machine (ATM), Internet Banking and Mobile Banking. These electronic banking would not involve visiting the banks within the banking hours. They can be carried out at one's convenient time once initial process is concluded in the banking hall. Each of these variants of electronic banking has undergone various stages of improvements arising from different challenges encountered along the way. Mobile banking is the main focus in this paper. Mobile banking, as an aspect of banking essentially, involves carrying out allowable banking transactions such as funds transfer, bill payment, airtime top up among others via mobile devices. As more features are added to the devices, banks also roll out products that would not ordinarily sit on mere phones due to basic requirements of the mobile applications to support such products. However, many bank customers are scared when they hear of electronic banking and this has prevented them from taking the full advantages offered by this development. Unless these security concerns are addressed, the mobile banking may not be able to achieve the desired results. While a single security method may not always be sufficient for protection, this paper provides a hybrid authentication for mobile banking transactions via biometrics along with the Personal Identification Number (PIN).

II. TRADITIONAL AND ELECTRONIC BANKING (e-BANKING)

The main difference between Traditional banking and e-banking is that the former involves going to the banking hall to transact business whereas the latter does not. Traditionally, banking business is about deposits and withdrawals, mainly cash-related. Over these two hinged other banking products designed to cater for various needs of customers. Other banking activities like funds transfer and loans stemmed out of deposit and withdrawal. The advent of electronics has re-focused and modernized the way and manner banking activities are run today. Prior to electronic banking, transactions were done intra-banks and with much hassles. Errors and irregularities were the order of the day. Technological advancement has also impacted on banking in no small measure - helping to improve efficiency and reducing leakage. Thus, banks could leverage on this to do inter-bank transactions irrespective of region or distance. For instance, after consummation of a business involving a large sum of money many bags may be required to move such money from the buyer to the seller or beneficiary. Such movements involve so many difficulties like time wastage security and other logistics. Various improvements had been introduced to solve these problems at one developmental stage or the other. Initially, variants of cheques were in use – bank cheques/drafts, manager cheques. Further improvements led to transfers – locally and internationally through electronics howbeit, through the banks and within the banking hall. ATM is also playing a vital role in solving some of the identified problems by extending transactions beyond banking hours and halls. Funds can now be transferred, bills can be paid, cash can be deposited and many other useful products are available and running on the ATM. Electronic transfers between countries was

also developed through Telex or SWIFT. With the introduction of internet, banks also cash in to make available internet banking. Mobile phones and devices are now found everywhere. Banks have identified these and are reeling out products that can be enjoyed by users of mobile devices. With government's initiatives such as financial inclusion of the unbanked, mobile banking has a major role to play. Thus, mobile banking is fast gaining ground. However, more efforts would be required to strengthen the security concerns relating to mobile banking. Once this is achieved the financial inclusion goal will gain more mileage.

III. IDENTIFICATION AND AUTHENTICATION (I & A)

I & A is a form of security to prevent unauthorized access to a facility. Identity is a process whereby a system recognizes the person accessing it. Each system keeps a list of allowable users. Ant time a user attempts login on to the system, it validates who the individual is relating to the stored data of allowable users otherwise there will be a denial. Authentication is a process whereby the system makes a further confirmation that the user is actually who he said he is. Authentication can be done through any or combination of these: "something you have", "something you know" and "something you are". Thus, authentications are categorized along these. Where authentication involves any of "something you have" (like a token) or "something you know" (such as PIN), it is termed a single-factor authentication (SFA). It is called a multi-factor authentication (MFA) where it involves any two or more of something you have, something you know or something you are (such as biometrics).

IV. SECURITY ISSUES ON MOBILE BANKING AND PRESENT CONTROLS

Identity theft has been a major issue to be resolved. Mobile devices store so many details especially the ones that are needed without having to memorize them. Sometimes, because we are so used to them, passwords and PINs are also stored in them. When these devices are lost or carelessly handled mobile banking details are compromised leading to frauds.

Several attempts have been made to reduce cases of fraud by instituting relevant controls. These include usage of Password, PIN and Token. Presently, most of the authentication methods are built around something you have and something you know. Although these give some level of comfort, single point of failure (SPOF) makes a mockery of the control. There have been instances of social engineering, too. Customers are tricked to divulge their login details (token generated digits, password and user ids) and the result is fraud. Biometrics have given some leeway on applications and guaranteed that it is rightful owner that accesses. Examples of these include biometric-controlled doors, biometric controlled access to laptop among others.

V. BIOMETRICS

Biometrics can be seen as the use of physical characteristics, behavior or skills to identify a person [8]. These include palm, finger, iris, voice, DNA, face and a host of others. The basic idea about biometrics is that its features such as palm, iris or face become permanent shortly after birth and cannot be shared like password. A comparative analysis of biometric features carried out revealed that fingerprint ranked high in terms of acceptability, accuracy and cost of deployment [11]. Hence, this paper chooses fingerprint as the biometric feature for the research.

VI. RELATED WORKS

Sitekey app was implemented in 2006 and the bank made bold to tell customers that they are 100% guaranteed in the event of any fraud loss [4]. Human face was implemented at the apps level. There are issues with this. Firstly, there is no back up for human face and a slight change may lead to denial. Secondly, the biometric authentication was done at apps level hence, any compromise at this point of access will be sufficient to perpetrate fraud. Mobile banking otherwise known as m-banking has a rich combination of social usage of mobile phones with financial services [1]. So many activities such as calls, text messages and financial transactions go on simultaneously. Therefore, there must be a robust database to accommodate these without performance degradation and address the security concerns. In 2010, token was introduced on the mobile device as a second level means of authentication [12]. The issue of no back up was taken care of. However, because it was stored on the device, an outright theft can lead to fraud. In 2011, registration of each process client as a registered user was introduced. This was to keep a unique account ID in the configuration database with a view to creating separate configuration data for the process involved. This was able to prevent packet replay. The registration procedure should be able to prevent the open accessibility of the business process. Hence, the biometric enrolment should be safeguarded while necessary encryption strategy should be in place to preserve confidentiality and integrity of data. It was maintained that today's mobile devices make use of SIM cards which are forms smartcards. They are capable of carrying network identity information as well as storing different applications [6]. They can be seen as computer on their own. They can process and store data. We can then leverage on this to store mobile banking apps which interfaces with the bank's host. In 2012, the capabilities of carrying out so many activities with mobile devices were observed but customers of banks are scared of making banking out

of these devices because of security worries [2]. However, solving the security issues in terms of Identity Management (IdM) will lay more credence to the usage of mobile banking platforms. In 2012, mobile banking was seen as an emerging channel with a lot of benefits [14]. However, a major identified challenge why customers are not embracing it is the identified lack of confidence in the security of the services. Confidence can only be restored when the platform is able to take care of these security concerns. In 2012, fingerprint authentication was explored for payment cards rather than the traditional PIN that is susceptible to shoulder surfing or compromise from where it was scribbled [3]. Howbeit, it was proposed for cards. In 2013, voice was explored as a means of biometric authentication for preserving identity in mobile [9]. This is a major breakthrough and gives so much comfort. However, false rejection rate (FRR) is higher with this approach. May be due to change in whether or cough or voice cracks, the voice tends to change thereby causing a mismatch between the pre-recorded and that of the life user. Most mobile apps used by most banks are laden with varying weaknesses that are capable of compromising customers' data [7]. It was further stressed that a recent test conducted on 275 Android-based mobile banking apps, selected from 50 major financial institutions, revealed that 8 out of 10 apps had security issues. Thus, security concerns should be taken beyond apps level. A recent research showed that there was a sharp increase (of about 54%) in the number of customers who consider mobile banking "unsafe" or "very unsafe" between 2009 and 2010 [13]. This corroborates the fact that there are security worries as far as mobile banking is concerned. In 2014, there was a review of recent development in Apple's iPhone5 and Samsung S5. Both allow fingerprint authentication on the device [10]. The devices have brought some improvement by employing Touch ID technology rather than swipe. The technology permits enrolment of up to five fingers (five for iPhone, three for S5) to be captured on the device where authentication is done. The goal is to securely lock the device. This approach also affords users the ease of usage and the choice of biometrics is made optional. This has provided a leeway.

VII. HYBRID AUTHENTICATION METHOD

The existing method of carrying out mobile banking transactions is done by logging on to a mobile device whose Subscriber Identity Module (SIM) card had been previously registered with both the telecommunication company (Telco) and the bank. SIM provides the identification while authentication is done by a 5-digit PIN. Essentially, mobile banking transactions are done using SIM and PIN. This makes mobile transactions vulnerable to attacks by stealing the device and guessing or stealing the PIN. However, hybrid solution allows mobile application to be accessed but both PIN and biometric input – fingerprint to complete a transaction. To achieve this, a database of customers' biometric details is kept. PIN and biometric inputs are supplied via the mobile device (where SIM had been previously linked to customer's account number) and routed through the internet to the bank for authentication. Where there is a match, the transaction is finalized. Otherwise, an error message is displayed.

REFERENCES

[1] Anyasi, F. and Otubu, P. (2009). Mobile Phone Technology in Banking System: Its Economic Effect. Research Journal of Information Technology 1(1): 1-5, 2009 ISSN: 2041-3114

[2] AuWerter, S. (2012). Is mobile banking really safe? www.money.cnn.com. Accessed January 20, 2014

[3] Awodele, O. and Akanni, A. (2012). Combating automated teller machine frauds through biometrics. International Journal of Emerging Technology and Advanced engineering. Vol. 2, issue 11, 2012.

[4] Bank of America (2006). Mobile banking security from Bank of America. www.bankofamerica.com. Accessed January 20, 2014

[5] Central Bank of Nigeria (2014). Circular on the acceleration of bank verification number (BVN) project. Banking and payment department, Ref: BPS/DIR/GEN/CIR/01/013, September 18, 2014.

[6] Edsbacker, P. (2011). SIM Cards for Cellular Networks**:** *An introduction to SIM Card Application Development.*

[7] Higgins, K. (2013). Weak Security In Most Mobile Banking Apps. www.darkreading.com. Accessed January 4, 2014

[8] Jacobs, B. and Poll, E. (2010). Biometrics and Smart Cards in Identity Management.

[9] Krinsbruner, E. (2013). Key Considerations for Testing Voice Recognition in Mobile Banking Applications.

[10] Kypreos, E (2014). Galaxy S5 Fingerprint Scanner vs iPhone 5S Touch ID

[11] Nandakumar, K., Ross, A. and Jain, A. (2009). Biometric fusion: does modeling correlation really matter? IEEE 3rd Intl. Conf. on Biometrics: Theory, Applications and Systems (BTAS 09), Washington DC

[12] Scotiabank (2010). Mobile Banking Security and Privacy. www.scotiabank.com. Accessed January 20, 2014

[13] Siciliano, R. (2013). What is Mobile Banking? Is it Safe? www.huffingtonpost.com. Accessed January 20, 2014

[14] Ullrich, J. (2012). Security of Mobile Banking and Payments.

Effecting Secure Mutual Agreement in a Mutually-Suspicious Vicious Party

Odule Tola John
Department of Mathematical Sciences
Olabisi Onabanjo University, Ago-Iwoye
tola.odule@oouagoiwoye.edu.ng

Goga Nicholae
Department of Bioinformatics
University of Groningen, The Netherlands
n.goga@rug.nl

Awodele Oludele
Department of Computer Science
Babcock University, Ilishan-Remo
delealways@yahoo.com

Abstract- In some scenarios, the power to generate valid signatures is shared among a set of users or computers. In a guild, for instance, it is not convenient that a single person has the knowledge of the secret key and signs all the *works* on behalf of the guild. This fact decreases *security*: single point of attack and *reliability:* remains inoperative if the player has some technical problem. A solution to solve these problems is to distribute the power of signing among a set $P = \{p_1, p_2, \cdots, p_n\}$ of ℓ players, where a monotone increasing family of qualified subsets $\Gamma \subset 2^P$ must be defined. In our scheme, we have proposed a signature scheme that is secure even in the presence of an adversary who corrupts and controls the behaviour of some subset of dishonest players. Our scheme assumes a *passive* adversary structure $\Lambda \subset 2^P$ that is monotone deceasing in which the adversary structure Λ is determined by its basis $\Lambda_0 = \{B \in \Lambda \mid B \cup \{P_i\} \notin \Lambda, \text{for all } P_i \notin B\}$. Our proposed scheme is secure both in the *computation-* and *information-theoretic* models.

Keywords- **Secret sharing, distributed signature, passive adversary, perfect security**

1.0 INTRODUCTION

Scenario *Suppose a team of 13 members comprising 7 groups is working on a prized film project. The 7 groups are made up of 3 members from the guild of actors, 3 from the guild of marketers, 2 from the guild of directors, 2 from consumer rights protection agency and 1 each from copyright agency, censures board and national arts & gallery. These groups of mutually-suspicious individuals want to ensure that it takes at least 7 members from the team to cooperate to have access to the master tape. How can this problem be resolved such that no individual is marginalised?*[3]

Clearly there are two approaches around this problem: the *combinatorics* approach which is physically impractical and the *secret sharing* paradigm as introduced independently by A. Shamir [Sha79] and G. Blakely [Bla79].

Secret Sharing Scheme The main idea of secret sharing is to have a dealer distribute a secret s among more than one player. Each player will only have a share of the secret, not the secret itself. The secret can be reconstructed and used for a specified purpose by recombining a certain number of the total shares, depending on the scheme used. The security is based on the fact that each share is useless when used alone, but can be used for its purpose when combined.
Secret splitting scheme is a simpler variant of secret sharing schemes as defined in [15]

Definition 1. *Secret splitting is done by giving each player a share of the secret in such a way that it takes all the players to reconstruct the secret.*

Similarly, the privacy of a secret sharing protocol is defined by [5]:

Definition 2. *A protocol is t-private if any set of at most t players cannot compute after the protocol more than they could jointly compute solely from their set of private inputs and outputs.*

Secret splitting is therefore $(n-1)-$private and can be conducted in many ways, all of which have in common that all the players need to input their shares to be able to reconstruct the secret.

1.1 ADDITIVE SECRET SHARING SCHEME ADDITIVE

schemes involve that all players input their shares to reconstruct the secret. More formally, given n players with a secret s in a finite field Z_p and a dealer D D picks $n-1$ random numbers $\{r_1, r_2, \cdots, r_{n-1}\}$ from Z_p. D then computes

[3] This is an adaptation of the problem in [23]

$$s_n = s - \sum_{i=1}^{n-1} r_i \pmod p$$

Then player 1, player 2,..., player $n-1$ receives the shares $s_i = r_i$ from D (through secure channels). Player n receives the share s_n as calculated above) also from D (through a secure channel). The reconstruction of the secret is done simply by adding the shares from all the players in Z_p:

$$s = \sum_{i=1}^{n-1} s_i \pmod p$$

As shown above, no single shareholder knows anything about the actual secret, only a random integer. Suppose an adversary should get hold of $n-1$ shares, this would yield nothing about the actual secret, because the last random integer does only make sure that the secret s is in the range [0, p), which is already given by the finite field used.

Alternatively, the XOR scheme may be used as follows: Let M be the message to be secret shared among several players, where M is of length ℓ-bit. D gives the first $n-1$ players a random ℓ-bit sequence each, and gives the last player the ℓ-bit sequence such that the XOR of all bit sequences equal the bit sequence of the message M. More formally it can be written:

$$m_i = \{0,1\}^{\ell} \; for \; i \in [1, n-1]$$
$$m_n = M \oplus m_i \oplus \ldots \oplus m_{n-1}$$

As can be seen from the equations above, each player has only a random bit sequence, but when they are all XOR-ed together, they will yield the message M. The same privacy applies here as for the additive scheme, $(n-1)$-private, knowing all except one share of the secret yields nothing for an adversary. The adversary can XOR all obtained shares, but this only yields a random bit sequence, and knowing no bits of the actual secret makes every ℓ-bit sequence a possible last share, that is, M can be any ℓ-bit sequence, which is already given. This leads to another definition for perfect security in a cryptographic system [29]:

Definition 3. *Perfect security is a cryptographic system in which the ciphertext yields no possible information about the plaintext (except possibly its length).*

1.2 THRESHOLD SHARING SCHEMES.

Definition 4. *Given a finite field Z_p of possible secret values, a (t,n)-threshold secret sharing scheme is a secret sharing scheme that can divide a secret $s \in Z_p$ into shares $\{s_1, s_2, \cdots, s_n\} \in Z_p$ so that $t \le n$ and:*

1. Given any set of t or more shares s_i, s can be reconstructed.
2. Any set of fewer than t shares gives no information about s.

Using Definition 2 it can be verified that threshold schemes are $(t-1)$-private, where t refers to the threshold used. More formally, There is a secret s and a set of players $P = \{p_1, p_2, \cdots, p_n\}$. s is broken into n shares $\{s_1, s_2, \cdots, s_n\}$ and each s_i is given to p_i $(1 \le i \le n)$ such that

(1) If $P' = \{p_{i1}, \cdots, p_{it}\} \subset P$ is a qualified subset, then s can be reconstructed from their shares $\{s_{i1}, \cdots\cdots, s_{it}\}$, and

(2) If $P' = \{p_{i1}, \cdots, p_{it}\} \subset P$ is not a qualified subset, then s cannot be reconstructed from their shares $\{s_{i1}, \cdots\cdots, s_{it}\}$.

$\Gamma \subset 2^{\mathrm{P}}$ must be defined and must be *monotone increasing*; that is, if $A_1 \in \Gamma$ and $A_1 \subset A_2 \subset \mathrm{P}$ then $A_2 \in \Gamma$. As a result of this property, an access structure is fully determined by its basis

$$\Gamma_0 = \{A \in \Gamma \mid A - \{\mathrm{P}_i\} \notin \Gamma, \text{for all } \mathrm{P}_i \in A\}$$

The family of all the qualified subset is called the access structure of the scheme.

1.3 MOTIVATION

Adaptive chosen-message attack: Suppose an adversary chooses tow messages m_1 and m_2 of his choice and queries the signing oracle for their respective signatures s_1 *and* s_2,he can later create a new message $m = m_1 \times m_2$ and claim that it is a valid signature on m because of the following multiplicative property of a modular-arithmetic-based signature scheme such as RSA.

$$s = (s_1 x\, s_2) \bmod n \;=\; (m_1^d x\, m_2^d) \bmod n \;=\; (m_1 x\, m_2)^d \bmod n \;=\; m^d \bmod n$$

This is a case of selective forgery since the adversary can manipulate m_1 and m_2 to get a useful m.

Security and Reliability In a coalition such as the one mentioned in our scenario, it is not convenient that a single player has the knowledge of the secret key and signs all the messages on behalf of the company. This fact decreases security, because an adversary must attack only a single point to obtain the full secret information. It also decreases reliability, because the signature system remains inoperative if the player has some technical problem.

A solution to solve these problems is to distribute the power of signing among a set $P = \{p_1, p_2, \cdots, p_n\}$ of ℓ `players, where a monotone increasing family of authorized or qualified subsets $\Gamma \subset 2^{\mathrm{P}}$ must be defined. This family will be called the access structure of the system. Each player will have a share of the secret key of the set. To compute a signature, each player of an authorized subset will use his secret share to compute a partial signature. Finally, a combining algorithm will convert the partial signatures from the subset into a valid standard signature of the message, that can be verified by using the single public key which matches with the shared secret key.

Our goal in this paper is to practically demonstrate, in a formal manner, how to solve the problem of key distribution for survivable systems such that a passive adversary even with a large window of vulnerability will not learn any information from the partial keys to compromise the legacy (copyrighted and patented) data stored on such systems.

2.0 RELATED WORKS

Blakley and Shamir invented threshold sharing schemes independently [4][30]. In Blakley's scheme, the intersection of m of n vector spaces yields a one-dimensional vector that corresponds to the secret. In Shamir's scheme, the interpolation of an $m-1$ degree polynomial through m of n points yields a constant term in the polynomial that corresponds to the secret

Chor et al. present a verifiable secret sharing (VSS) scheme in which the dealer and shareholders perform an interactive secure distributed computation [7]. Rabin and Ben-Or [27] propose schemes in which the dealer and share-holders participate in an interactive zero-knowledge proof of validity; the scheme of Rabin and Ben-Or, is information-theoretically secure.

Frankel et al. [13] and Rabin [26] propose threshold PSS schemes in which each shareholder periodically distributes a subshare of its share to all the other members. Each shareholder then combines the subshares to generate a new share. A drawback of these protocols is that the shareholders rely on commitments received during the initial distribution of the secret to verify the validity of the new shares, and thus one cannot redistribute between disjoint sets of shareholders. Also, the commitments depend on (m, n) , and thus one cannot redistribute between different access structures.

Desmedt and Jajodia present a secret redistribution protocol that does not require the intermediate re-construction of the original secret [10]. Their protocol allows redistribution between different (possibly disjoint) sets of shareholders with different access structures. Unfortunately, a faulty old shareholder can undetectably distribute "subshares" of some random value instead of subshares of a valid old share, and thus cause new shareholders to generate invalid shares.

Frankel et al. propose a proactive threshold sharing scheme for RSA private keys [12]. The protocol uses a poly-to-sum redistribution from a (m, n) to (m, m) sharing scheme, and a sum-to-poly re-distribution back to a (m, n) scheme. During redistribution, each old shareholder broadcasts a commitment to its share, which new shareholders use to verify the validity of their generated share. Unfortunately, during redistribution to a disjoint set of shareholders, it is not enough for the old shareholders to broadcast the commitment to their respective shares, since a faulty shareholder can broadcast a random "commitment."

Ostrovsky and Yung introduce the concept of mobile adversaries [12] that corrupt participants in a distributed protocol at a constant rate. Zhou, Schneider, and van Renesse propose a PSS scheme for asynchronous, wide-area networks, and employ it in an on-line certification authority [35].

In the area of distributed signature schemes, different proposals have appeared throughout the last fifteen years. With respect to schemes whose security is based on the difficulty of solving the Discrete Logarithm problem, we can cite the proposals in Gennaro et al [18] and D.R.Stinson and R. Strobl [32]. With respect to schemes based on the RSA primitive (related to the difficulty of factoring large integers), the most significant proposals can be found in A. De Santis et al [9], R. Gennaro,et al [16], V. Shoup [31], I. Damg°ard and M. Koprowski [11] and P.A. Fouque and J. Stern, [14].

J. Herranz and G. Saez. [19] used distributed Schnorr's signatures, where the delicate point is to extend verifiable secret sharing schemes, already proposed for the threshold case, to a general framework. In that work, a fully distributed proxy signature scheme was also proposed, where a distributed entity delegates its signing capabilities to a distributed proxy entity; the proxy entity can sign messages on behalf of the original entity, and the recipient verifies at the same time the delegation of the original entity and the signature of the proxy entity. The initial proposal in [19] was revised and extended with a formal security analysis in J. Herranz and G. Saez. [20].

A threshold version of Schnorr's signature scheme can be found in D.R.Stinson and R. Strobl [32], and a threshold version of Digital Standard Signature (DSS) scheme can be found in R. Gennaro, et al [16]. For this kind of schemes, the generation of the keys can be jointly executed by the players, by following the protocol explained in R. Gennaro et al [17] for the threshold case. Another example of threshold signature scheme can be found in A. Boldyreva [6]. Its security is based on the Computational Diffie-Hellman Assumption.

A variant of implementation of distributed signature protocol which is fast emerging as an alternate model of computation is *multi party computation* (MPC). Even though initially considered to be of theoretically value only, in 2008, P. Bogetoft et al [1] reported some 1200 farmers in Denmark employed an MPC protocol to determine the market price of sugar beets contracts without having to reveal their (sensitive) selling and buying prices and without resorting to an external trusted party.

This alternate model of computation may be implemented as protocols that either allow to evaluate Boolean circuits or arithmetic circuits with addition and multiplication gates. It is also possible to use combined approach, where parts of the computation are performed using different representations as shown in V. Kolesnikov, A. Sadeghi and T. Schneider [22].

Virtually every MPC protocol that allows to evaluate a boolean circuit is based on Yao's garbled circuits by A. C. Yao [34]. However, Yao's protocol is only secure against passive adversaries. The most promising attempt of making Yao's protocol secure against active adversary in an efficient way can be found in Y. Lindell and B. Pinkas [24]. R. Bendlin et al [3] used an additively homomorphic encryption (additive secret sharing) to propose an efficient implementation in their BeDOZa protocol.

3.0 DESCRIPTION OF THE PROTOCOL

Usually, distributed schemes are designed from a standard (individual) scheme. The secret key of an individual user in the standard scheme is distributed in shares by means of a secret sharing scheme. Each participant of the set receives a share of the secret key. Later, each one uses his partial secret information to perform his part of the task, such as signing a message. Our implementation uses the Shamir's scheme to realize the proposed solution. The essential idea of Shamir's threshold scheme is that it takes t points to define a polynomial of degree $t-1$.

Suppose we want to use a (t, n) threshold scheme to share our secret s; without loss of generality assumed to be an element in a finite field F of size P where $0 < t \le n < P; s < P$ and P is a prime number.

Choose at random $t-1$ positive integers $a_1, \cdots\cdots a_{t-1}$ with $a_i < P$, and let $a_0 = s$ Build the polynomial

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_{t-1} x^{t-1}$$

Let us construct any n points out of it, for instance set $i = 1, \cdots, n$ to retrieve $(i,\ f(i))$. Every participant is given a point (an integer input to the polynomial, and the corresponding integer output). Given any subset t of these pairs, we can find the coefficients of the polynomial using interpolation. The secret is the constant term a_0.

Share Creation Dealer D first picks a random polynomial of degree $t-1$, where t is the threshold. Each player has a different id, $P_{id} = x_i$, which is constant for that player in the current scheme and is known by everyone. Normally, this id is given in increasing order for simplicity, such that $x_i = 1, 2, 3, \cdots$ for player 1, 2, 3, ... D now gives each player a share $f(x_i)$ such that player 1 gets the share, $s_1 = f(1)$ player 2 gets the share $s_2 = f(2)$ and so on.

Share Reconstruction Reconstructing the secret can be performed by any t number of players using their values x_i and $f(x_i)$. The reconstruction is done by using Lagrange's interpolation on t (or more) shares. Lagrange interpolation is defined as follows ([Kre99]):

$$f(x) = \sum_{i=1}^{n} L_i(x) f_i = \sum_{i=1}^{n} \frac{l_i(x)}{l_i(x_i)} f_i$$

where $l_i(x)$ and $l_i(x_i)$ are defined as follows:

$$l_i(x) = \prod_{j=1}^{n, i \neq j} (x - x_j)$$

$$l_i(x_i) = \prod_{j=1}^{n, i \neq j} (x_i - x_j)$$

$$f(0) = \sum_{i-1}^{t} s_i \cdot \prod_{j=1}^{t, i \neq j} \frac{0 - x_j}{x_i - x_j} \bmod p$$

By multiplying both the numerator and the denominator by -1, which interchanges the indexes for i and j in the denominator, and putting the denominator as a power of -1 instead, the final expression for calculation is:

$$s = f(0) = \sum_{i-1}^{t} s_i \cdot \prod_{j=1}^{t, i \neq j} x_j \cdot (x_j - x_i)^{-1} \bmod p \quad (1)$$

Notice that the indices i and j refers to the P_{id}'s of the players participating in the reconstruction, but are written as 1 to t here for simplicity.

3.1 A SMALL EXAMPLE

To share a secret $s = 8971$ among five players with a threshold $t = 3$ using the above scheme, D chooses, say, a prime $P = 9929$ and two random variables, say, $a_1 = 5$ and $a_2 = 7$ which yields the following polynomial

$$f(x) = 7x^2 + 5x + 8971 \bmod 9929$$

$$s = \sum_{i-1}^{t} s_i \cdot \prod_{j=1}^{t, i \neq j} x_j \cdot (x_j - x_i)^{-1} \bmod 9929$$

By substituting the expressions for $l_i(x)$, $l_i(x_i)$ and setting $f_i = s_i$, $f(x)$ can be rewritten as:

$$f(x) = \sum_{i-1}^{n} s_i \cdot \prod_{j=1}^{n, j \neq j} \frac{x - x_j}{x_i - x_j}$$

Since the threshold is t, n can be substituted by t. The secret is found for $x = 0$, therefore the desired expression is found by finding $f(0)$, and all values are calculated $\bmod p$ in secret sharing schemes, which yields the following:

D then calculates one share for each player and gives it to that player in a secure manner as follows:

$$s_1 = f(1) = 7 \cdot 1^2 + 5 \cdot 1 + 8971 \bmod 9929 \quad = 8983$$

$$s_2 = f(2) = 7 \cdot 2^2 + 5 \cdot 2 + 8971 \bmod 9929 \quad = 9009$$

$$s_3 = f(3) = 9049$$

$$s_4 = f(4) = 9103$$

$$s_5 = f(5) = 9171$$

The secret s is now secret shared among the 5 players, each of them having their own distinct share of s. It would require at least $t = 3$ players in order to reconstruct the secret, as shown below.

Suppose players 1, 3 and 5 want to reconstruct the secret s Each of these players has two values x_i, and $f(x_i)$ that is used in the reconstruction:

Player 1 :(1, 8983)
Player 3 (3, 9049)
Player 5: (5, 9171)

These values can be used as input to Equation (1) to reconstruct the secret as shown below:

$$8983 \cdot \prod_{j=1}^{t, i \neq j} x_j \cdot (x_j - x_i)^{-1} + 9049 \cdot \prod_{j=1}^{t, i \neq j} x_j \cdot (x_j - x_i)^{-1} + 9171 \cdot \prod_{j=1}^{t, i \neq j} x_j \cdot (x_j - x_i)^{-1} \bmod 9929$$

$$= 8983 \cdot (3 \cdot (3-1)^{-1}) \cdot (5 \cdot (5-1)^{-1}) + 9049 \cdot (1 \cdot (1-3)^{-1}) \cdot (5 \cdot (5-3)^{-1}) +$$
$$9171 \cdot (1 \cdot (1-5)^{-1} \cdot (3 \cdot (3-5)^{-1)} \bmod 9929$$
$$= 8983 \cdot (34965) \cdot (5 \cdot 7447) + 9049 \cdot (1 \cdot 4964) \cdot (5 \cdot 4965) +$$
$$9049 \cdot (1 \cdot 4964) \cdot (5 \cdot 4965) \bmod 9929$$
$$= 4982109464475 + 1115120033700 + 3389779884224 \bmod 9929$$
$$= 6436207486599 \bmod 9929$$
$$= 8971$$

This small example shows that three out of the five shares are enough to recover the secret s using Lagrange's interpolation. It should be noted, however, that the calculations require finding inverses $\bmod p$ using the Extended Euclidean Algorithm [8], [28], [33].

4.0 SECURITY ANALYSIS OF THE PROTOCOL

We hereby claim that the above scheme is both *information-* and *computation-theoretic* secure in the *random oracle* model.

Definition 5 *A* (P, Γ)*-distributed signature scheme consists of three polynomial-time and probabilistic protocols:*

- *Dist-Key-Gen: this protocol can be executed jointly by the players themselves, or by a trusted and external authority. The input is a security parameter. The public outputs are pk (the public key of the scheme) and some verification key vk, whereas the private output of each player P_i is a share sk_i of the secret key sk related to pk.*

We hereby omit, without loss of generality, discussion of this phase for the sake of space.

- *Dist-Sig: if m is the message to be signed, each player P_i uses his private information to compute and broadcast his partial signature $\sigma_i(m_i)$. The correctness of the partial signatures can be verified using the verification key vk. Finally, a combiner algorithm takes valid partial signatures corresponding to an authorized subset $A \in \Gamma$ and produces from $\{\sigma_i(m_i)\}_{p_i \in A}$ a valid standard signature $\sigma(m)$.*
- *Ver: this protocol is executed by the recipient of the signature. The inputs are the public key pk, the message m and the signature $\sigma(m)$. The output will be \true" or \false".*

The adversary structure is hereby denoted by $\Lambda \subset 2^{\mathrm{P}}$, which is *monotone decreasing*, i.e., if the scheme remains secure when an adversary corrupts a subset $B_1 \in \Lambda$, then it must also remain secure if an adversary corrupts players B_2 for $B_2 \subset B_1$. Thus informally, we say that a (P, Γ)*-distributed signature scheme is* Λ*-secure* if it is *robust* and *existentially unforgeable under adaptive chosen message attacks*, considering an adversary allowed to corrupt players of any subset in the structure Λ.

By robustness we refer to the fact that the scheme provides mechanisms to detect corrupted players who do not follow the protocol correctly. Furthermore, the protocol must always produce a valid signature from the partial signatures of the honest players. Again, we defer discussion of robustness to our work-in-progress on distributed RSA signature and concentrate here on existential forgery.

Existential unforgeability under adaptive chosen message attacks is defined in relation to the following game $\mathcal{G}1$:

1. The adversary is given a set of players P a monotone increasing access structure $\Gamma \subset 2^P$ and an adversary structure $\Lambda \subset 2^{\mathrm{P}}$.
2. The adversary chooses a subset of players $B \in \Lambda$ to corrupt.
3. The *Dist-Key-Gen* protocol is executed. The adversary obtains all the information that is made public in the execution of this protocol, as well as private information corresponding to the corrupted players $P_j \in B$ (in particular, their shares sk_j of the secret key sk).
4. The adversary can adaptively choose Q_s messages to be signed. For these messages, the *Dist-Sig* protocol is executed. The adversary sees all the information that is made public in the required executions of this protocol, as well as private information corresponding to the corrupted players.

Claim 1: *An adversary is a* $(P,\Gamma,\Lambda,T,\varepsilon,Q,Q_s)$*-- forger against a distributed signature scheme if its total running time is at most* T *and it obtains, with probability* ε *a valid (message, signature) pair, different from the ones that it has received during the game* $\mathcal{G}1$.

Claim 2: *(Exact unforgeability of distributed signature schemes) A distributed signature scheme is* $(P,\Gamma,\Lambda,T,\varepsilon,Q,Q_s)$*-unforgeable if there does not exist any* $(P,\Gamma,\Lambda,T,\varepsilon,Q,Q_s)$*-forger against it.*

From Definition 2 we note that a (t, n)*-threshold scheme* is *t-private*, meaning that it is *computationally-secure*. Also, since in both the additive and Shamir's scheme presented in our protocol the shares are as large as the secret itself, it offers *perfect security*, in the context of Definition3. This coupled with the fact that the *information rate* of our scheme is unity in the sense that the quotient between the length of the secret s (in bits) and the maximum length of the distributed shares is at most 1 offers an adversary no knowledge of what is being

transmitted thus making our scheme *information-theoretic secure.*

5.0 CONCLUSION

We hereby note the following useful properties of our scheme:
Secure: It is information-theoretic secure as shown in section 4

Minimal: The size of each piece does not exceed the size of the original data.

Extensible: When t is kept fixed, s_i pieces can be dynamically added or deleted without affecting the other pieces.

Dynamic: Security can be easily enhanced without changing the secret, but by changing the polynomial occasionally, keeping the same free term and constructing new shares to the participants.

Flexible: In organizations where hierarchy is important, we can supply each participant different number of pieces according to their importance inside the organization. For instance, the president can sign a message alone, whereas 3 secretaries are required together to do same.

This paper proposed a method to securely distribute and reconstruct a signature in an insecure and hostile asynchronous environment. We, however, note the following assumptions made in our scheme: There is a (somewhat impractical but reasonably modified) secure channel for communicating with the players during *share distribution* and *reconstruction* phases. The existence of a *trusted third party* (dealer D) who is charged with the responsibility of keeping and sharing the secret among the non-trusting players and who is deemed trusted by all the players to be *available* and *reliable*. Finally, the adversary model considered in this work has been rather simple in that only *passive adversaries*, who for most of the time are *static* and faithfully follow the protocol is assumed.

We hope to address these observations and remove the restrictions in our future work that is on-going.

REFERENCES

[1] P. Bogetoft, D. Christensen, I. Damg°ard, M. Geisler, T. Jakobsen, M. Krøigaard, J. Nielsen, J. Nielsen, K. Nielsen, J. Pagter, M. Schwartzbach, and T. Toft. Secure multiparty computation goes live. In Financial Cryptography, 2009.

[2] W. J. Bolosky, J. R. Douceur, D. Ely, and M. Theimer. Feasibility of a serverless distributed file sys-tem deployed on an existing set of desktop PCs. In *Proc. of SIGMETRICS 2000, the Intl. Conf. on Measurement and Modeling of Computing Systems,* pp. 34–43. June 2000.

[3] R. Bendlin, I. Damgrd, C. Orlandi, and S. Zakarias. Semi-homomorphic encryption and multiparty computation. In EUROCRYPT, 2011.

[4] G. R. Blakley. Safeguarding cryptographic keys. In *Proc. of the Natl. Computer Conf.,* vol. 48 of *American Federation of Information Processing Societies Proceedings,* 1979.

[5] Michael Ben-Or, Shafi Goldwasser, and Avi Wigderson. Completeness theorems for non-cryptographic fault-tolerant distributed computation. In STOC '88: Proceedings of the twentieth annual ACM symposium on Theory of computing, pages 1–10, New York, NY, USA, 1988. ACM Press.

[6] A. Boldyreva. Threshold signatures, multisignatures and blind signatures based on the Gap-Di±e-Hellman-group signature scheme. Proceedings of the PKC'03, LNCS 2567, Springer-Verlag, pp. 31{46 (2002).

[7] B. Chor, S. Goldwasser, S. Micali, and B. Awerbuch. Verifiable secret sharing and achieving simultaneity in the presence of faults (Extended abstract). In *Proc. of the 26th IEEE Ann. Symp. on Foundations of Computer Science,* pp. 383–395. Oct. 1985.

[8] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. Introduction to Algorithms, Second Edition. The MIT Press, 2001.

[9] A. De Santis, Y. Desmedt, Y. Frankel, and M. Yung. How to share a function securely. Proceedings of STOC'94, pp. 522{533 (1994).

[10] Y. Desmedt and S. Jajodia. Redistributing secret shares to new access structures and its applications.Technical Report ISSE TR-97-01, George Mason University, Fairfax, VA, July 1997.

[11] I. Damgºard and M. Koprowski. Practical threshold RSA signatures without a trusted dealer. Proceedings of Eurocrypt'01, LNCS 2045, Springer-Verlag, pp. 152{165 (2001).

[12] Y. Frankel, P. Gemmell, P. D. MacKenzie, and M. Yung. Optimal resilience proactive public-key cryptosystems. In *Proc. of the 38th IEEE Ann. Symp. on Foundations of Computer Science,* pp. 384–393. Oct. 1997.

[13] Y. Frankel, P. D. MacKenzie, and M. Yung. Adaptive security for the additive-sharing based proactive RSA. In *Proc. of PKC 2001, the 4th Intl. Workshop on Practice and Theory in Public Key Cryptography, vol. 1992 of Lecture Notes in Computer Science,* pp. 240–263. Febraury 2001.

[14] P.A. Fouque and J. Stern. Fully distributed threshold RSA under standard assumptions. Proceedings of Asiacrypt'01, LNCS 2248, Springer-Verlag, pp. 310{330 (2001).

[15] Laurence Grant and Brian Fleming. Secret Sharing and Splitting. University of Notre Dame, Indiana, USA, 2002.

[16] R. Gennaro, S. Jarecki, H. Krawczyk and T. Rabin. Robust and efficient sharing of RSA functions. Proceedings of Crypto'96, LNCS 1109, Springer-Verlag, pp. 157{172 (1996).

[17] R. Gennaro, S. Jarecki, H. Krawczyk and T. Rabin. Secure distributed key generation for discrete-log based cryptosystems. Proceedings of Eurocrypt'99, LNCS 1592, Springer-Verlag, pp. 295{310 (1999).

[18] R. Gennaro, S. Jarecki, H. Krawczyk and T. Rabin. Robust Threshold DSS signatures. Proceedings of Eurocrypt'96, LNCS 1070, Springer-Verlag, pp. 354{371 (1996).

[19] J. Herranz and G. Saez. Verifiable secret sharing for general access structures, with application to fully distributed proxy signatures. Proceedings of Financial Cryptography Conference 2003, LNCS 2742, Springer-Verlag, pp. 286{302 (2003).

[20] J. Herranz and G. Saez. Revisiting fully distributed proxy signature schemes. Proceedings of Indocrypt'04, LNCS 3348, Springer-Verlag, pp. 356{370 (2004).

[21] Erwin Kreyszig. Advanced Engineering Mathematics, 8th Edition. John Wiley & Sons, Inc., 1999.

[22] V. Kolesnikov, A. Sadeghi, and T. Schneider. Modular design of efficient secure function evaluation protocols. Cryptology ePrint Archive, Report 2010/079, 2010.

[23] C. L. Liu. Introduction to Combinatorial Mathematics. McGraw-Hill, 1968.

[24] Y. Lindell and B. Pinkas. Secure two-party computation via cut-and-choose oblivious transfer. In TCC, 2011.

[25] R. Ostrovsky and M. Yung. How to withstand mobile virus attacks. In *Proc. of the 10th Ann. ACM Symp. on Principles of Distributed Computing*, pp. 51–59. Aug. 1991.

[26] T. Rabin. A simplified approach to threshold and proactive RSA. In *Proc. of CRYPTO 1998, the 18th Ann. Intl. Cryptology Conf.*, vol. 1462 of *Lecture Notes in Computer Science*, pp. 89–104. Aug. 1998.

[27] T. Rabin and M. Ben-Or. Verifiable secret sharing and multiparty protocols with honest majority. In *Proc. of the 21st Symp. on the Theory of Computing*, pp. 73–85. May 1989.

[28] Kenneth H. Rosen. Discrete Mathematics and Its Applications, Fifth Edition. McGraw-Hill, 2003.

[29] Bruce Schneier. Applied Cryptography - Protocols, Algorithms, and Source Code in C, Second Edition. John Wiley & Sons, Inc., 1996.

[30] A. Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, Nov. 1979.

[31] V. Shoup. Practical threshold signatures. Proceedings of Eurocrypt'00, LNCS 1807, Springer-Verlag, pp. 207{220 (2000).

[32] D.R.Stinson and R. Strobl. Provably secure distributed Schnorr signatures and a (t; n) threshold scheme for implicit certi¯cates. Proceedings of ACISP'01, LNCS 2119, Springer-Verlag, pp. 417{434 (2001).

[33] Wade Trappe and Lawrence Washington. Introduction to Cryptography with Coding Theory, Second Edition. Pearson Prentice Hall, 2006.

[34] Andrew C. Yao. Protocols for secure computations. In SFCS'82: Proceedings of the 23rd Annual Symposium on Foundations of Computer Science, pages 160–164, Washington, DC, USA,1982. IEEE Computer Society.

[35] L. Zhou, F. B. Schneider, and R. van Renesse. COCA: A secure distributed on-line certification authority. Tech. Rep. TR2000-1828, Dept. of Computer Science, Cornell University, Ithaca, NY 14853, Dec. 2000.

Combating Phishing Attacks in Enterprise Information Networks

Durosinmi A. E.
ICT Unit, Federal Medical Centre
Abeokuta, Ogun State, Nigeria.
cunlexie@hotmail.com

Adekusibe K. G.
Department of Computer Engineering,
Ogun State Institute of Technology, Igbesa, Ogun State
Nigeria.
kusibus@yahoo.com

Ayanlowo K. A.
Department of Computer Science
Moshood Abiola Polyetchnic,
Abeokuta, Ogun State
Nigeria.
blessedkay@yahoo.com

Longe O. B.
Department of Computer Science & Mathematics
Adeleke University, Ede,
Osun State, Nigeria.
longeolumide@fulbrightmail.com

Oladejo R.A
Department of Computer Science,
Ogun State Institute of Technology, Igbesa,
Ogun State. Nigeria.
funrack2002@yahoo.com

***Abstract*-** Phishing attacks cause substantial damages to individuals and corporations. We analyze these attacks, and identified that most of them exploit the fact that victims are not sufficiently aware of the attacks. We present the challenges associated with phishing and proposed security measures that service providers can take to prevent and manage phishing attacks.

Keywords: Phishing, fraud, detection, E-mail.

I. INTRODUCTION

Phishing is form of identity theft that combines social engineering techniques and sophisticated attack vectors to harvest financial information from unsuspecting consumers. Often a phisher tries to lure her victim into clicking a URL pointing to a rogue page. We describe several features that can be used to distinguish a phishing URL from a benign one. These features are used to model a logistic regression filter that is efficient and has a high accuracy. We use this filter to perform thorough measurements on several million URLs and quantify the prevalence of phishing on the Internet today [13]. Most Internet users have encountered phishing in the form of emails purporting to come from a bank or other business, but in fact originating from a malicious source and designed to persuade the recipient to hand over personal information such as credit card details[1]. Phishing scams normally occur via emails, websites, text messages and phone calls that can delude recipients' to think that Christmas came early. Cybercriminals will often pose as your bank or financial institution, your employer, or any other entity that you normally trust with your information. Only when the email phishing process and characteristics are fully understood can effective measures be designed against phishing attacks. Successful phishing attacks are based on a form of copying, or reengineering, a website's design and layout in order to pass themselves off as a genuine (targeted) website. A malicious website is crafted which looks and feels like the original site, convincing unsuspecting users that they are giving personal information to a trusted organization**.** Users are frequently drawn to the sites by forged emails designed to look like legitimate correspondence and may even copy the body from real email, but when the user clicks a link to visit the website, they will be directed to the malicious site instead. The more convincing a phishing attack appears - or rather, the more genuine a malicious website looks - the more success the attack will have in extracting personal information. Some phishing attacks go so far as to create faux websites for which there is no legitimate counterpart; e.g. a page prompting users for personal information the organization wouldn't have otherwise asked for. According to Anti-Phishing Working Group (APWG), phishing activities have been increasing and most phishing websites are hosted in the US. In 2012, an average of over 25,000 unique phishing email reports were reported to the APWG.

Plus, the number of unique phishing sites detected exceeded 45,000 per month [2].

STEPS IN PHISHING ATTACK

All phishing attacks fit into the same general information flow.

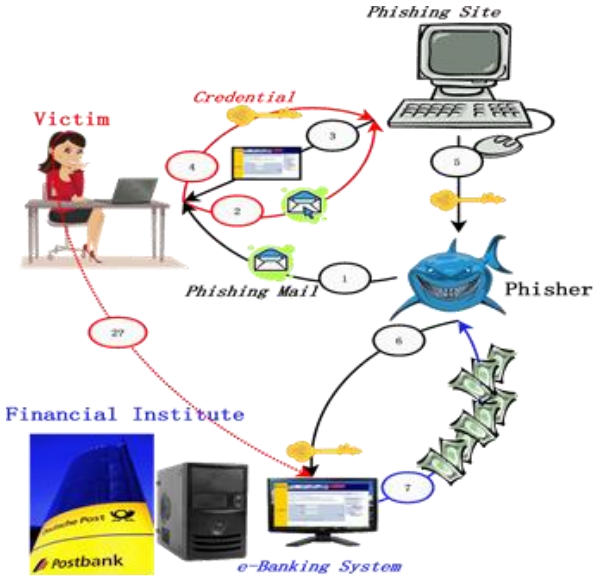


Fig. 1: Typical Phishing Scenario

http://www.hooklee.com/default.asp?t=Honeypots4AntiPhishing

The steps are:

Step 1: blocking the information flow from phishers to potential victims phishing email detection and filtering, email authentication, anti-malware software, cousin domain rejection, and so on.

Step 2-4: avoiding credential leakage user education, phishing site warning, inconsistent DNS information detection, cross-site/injected script rejection, mutual authentication, trusted path between user and web browser, delayed password disclosure, and so on.

Step 5: preventing phishers from getting stolen credentials early detection and quick takedown of phishing sites, fake credential submission, password rescue, and so on.

Step 6: making stolen credentials useless two-factor user authentication, password hashing, transaction monitoring and reconfirmation, and so on.

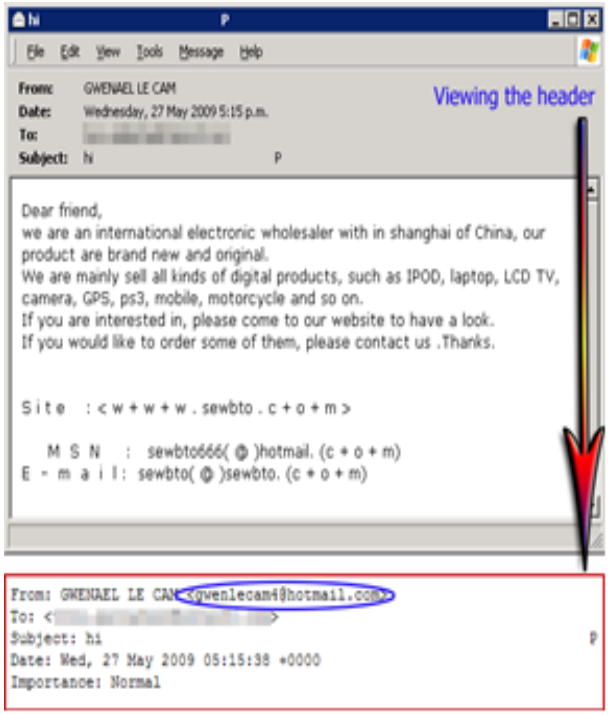
Step 7: preventing phishers from getting the stolen money, or catching phishers transaction authentication, intentional transaction delay, law enforcement, and so on[3].

II. PHISHING TECHNIQUES

There are a number of different phishing techniques used to obtain personal information from users. As technology becomes more advanced, the phishing techniques being used are also more advanced. To prevent Internet phishing, users should have knowledge of various types of phishing techniques and they should also be aware of anti-phishing techniques to protect themselves from getting phished. Let's look at some of these phishing techniques.

A. EMAIL PHISHING

Phishers may send the same email to millions of users, requesting them to fill in personal details. These details will be used by the phishers for their illegal activities. Phishing with email and spam is a very common phishing scam. Most of the messages have an urgent note which requires the user to enter credentials to update account information, change details, and verify accounts. Sometimes, they may be asked to fill out a form to access a new service through a link which is provided in the email. This

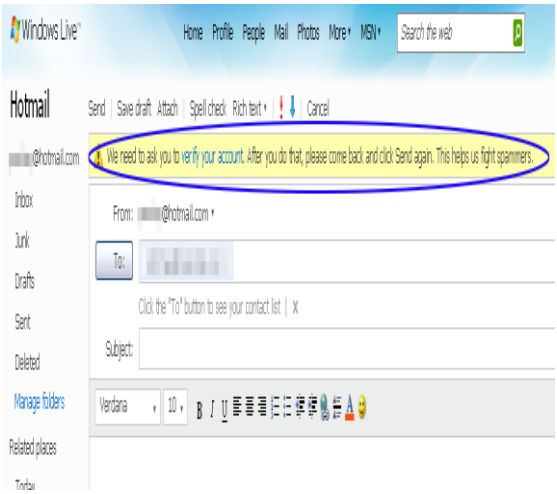


scams usually introduces a website based in China selling cheap and branded electronics. If you weren't already suspicious from reading the email, you should be after viewing the website. It reeks of a scam, especially because of the payment method options when buying goods from them. The websites prefer payment through Western Union, Bank Transfer or MoneyGram. These types of payment methods are not recommended especially when dealing with people you don't know and trust [4].

Another suspicious thing is that, if this website was legitimate they would not use free email services like Hotmail to adverstise their wares. They could have at least used an email address using their own domain.
Meanwhile, we noticed recently that Hotmail had actually moved to stop these spammers from abusing their free email service. A "Verify your Account" feature, which involve CAPTCHA verification was added to help prevent botnets from automatically sending spam. That's a welcome step. But in this case, it seems the spammers have found a way around it. Perhaps accounts are being manually verified prior to being used as spam conduits [4].

B. FRUIT SUCKER

When another attacker breaks into a hacked, vulnerable website with an existing phishing page and changes the email address. Often since the



attacker can see where the data is being emailed to he/she will keep the original email addresses intact and merely add his/her own email address in the BCC field. This is a very easy way for an attacker to make extra money. All passwords and account numbers keyed in reach his inbox directly without tipping off the original phishers. If the phishers are rookies and lack automated money transferring scripts, are too lazy to keep a watchful eye on their victims or are situated in a different time zone, then these advantages can help the fruit sucker withdraw a large amount of data (or other assets) from the compromised accounts before they do [6].

C. SPEAR SUCKER

An attack against the original crackers. For example, a good guy who breaks into the phishing websites and changes the email address to the NEDBANK's CSO's or CEO's email address. After this he contacts the bank to make them aware of the security breach.

D. HAXTORTIONIST

When an attacker patches the system, pulls down the phishing page and emails the attackers threatening them that he/she will report the crime and inform NEDBANK of their malicious activity. Reporting such abuse to email servers hosted by Google, Yahoo and similarly large companies in the United States is easy. In this way the attacker may extort a small share of money from the original crackers in return for keeping silent.

E. Robin-HAT

Here the attacker, after collecting a lot of passwords, changes the recipient's email address for the purpose of redistributing wealth. He/She withdraws money from the accounts and donates a significant portion to charity. Such individuals cannot be called grey hats because they are criminals robbing from other criminals. They are Robin-HATS, those who steal from rich victims and their attackers and redistribute the wealth to the poor and needy.
Another version on this above type of attack: the Robin-HATs uniformly redistributed the assets from the richer compromised accounts to the compromised accounts which had lower funds; especially if particular attention was paid to those accounts with low balances for a prolonged period of time.

F. MALWARE-BASED PHISHING

Malware-based phishing involves running malicious software on the user's machine. The malware can be introduced as an email attachment or as a downloadable file exploiting security vulnerabilities. This is a particular threat for small and medium businesses (SMBs) who fails to update their software applications.

G. SESSION HIJACKING PHISHING

Session Hijacking is a kind of phishing attack where user's activities are monitored clearly until they log into a target account like the bank account and

establish their credentials. At that point, the malicious software takes control and can undertake unauthorized actions, such as transferring funds, without the knowledge of the user.

H. HOSTS FILE POISONING PHISHING

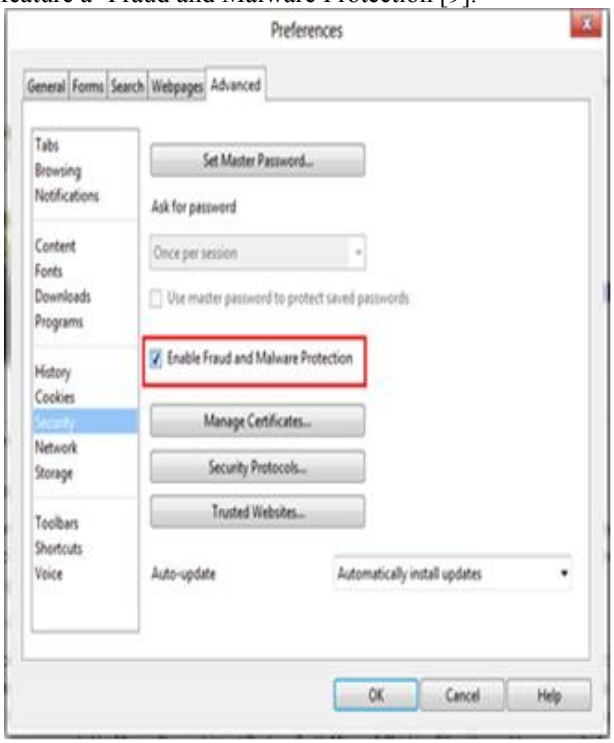
When a user types a URL of a website it is first translated into an IP address before it's transmitted over the Internet. The majority of user's PCs running a Microsoft Windows operating system first look up these "host names" in their "hosts" file before undertaking a Domain Name System (DNS) lookup. Phishers steal information by "poisoning" the hosts file. They transmit a bogus address, taking the user unwittingly to a fake "look alike" website.

I. SYSTEM RECONFIGURATION ATTACKS

This is a kind of phishing attack where the settings on a user's PC are modified with bad intentions. For example: URLs in a favorites file might be modified to direct users to bogus websites that look alike. For example: a financial institution's website URL may be changed from "bankofxyz.com" to "bancofxyz.com".

III. MITIGATING PHISHING

The following information applies to functions available in browsers such as Opera, Internet Explorer, Chrome and Microsoft Firefox.
Opera relies on blacklists from Netcraft and calls the feature a 'Fraud and Malware Protection [9].



Every URL you visit first goes to sitecheck2.opera.com (one of the noted IP is 91.203.99.45). Microsoft Internet Explorer calls its feature to detect and report malicious websites a 'SmartScreen Filter.'

It can be enabled as shown in the screenshot above or go to Tools / Internet Options/ <Navigate to the 'Security level for this zone' area> / Click on Custom level… / Miscellaneous/ Use SmartScreen Filter / Enabled. Similar checks are done by Microsoft Internet Explorer, Chrome and Firefox.
Alternatively you can use OpenDNS as your DNS on your Ethernet and wireless cards IP configuration. Although from personal observations, using OpenDNS takes you to a search engine (possibly powered by Google) on every website which fails to get resolved – and it might miss resolving a website hosting a phishing page.

Still it is quite good at maintaining uncorrupted, un-poisoned DNS records. It's reliable when you suspect your ISP of injecting iframes, redirecting you and tracking your visits.

OpenDNS can be set by filing the DNS as 208.67.222.222 and 208.67.220.220 (as an alternative or vice-versa)

Such features greatly reduce the chances of a person falling prey to phishing websites if it has been reported. On an average it takes no more than 8 hours of hosting a phishing page before somebody reports it. This is a very long gap when you consider that in Asia banks are provided phishing alerts by third-parties who sign an SLA stating that they have to inform the bank within 5 minutes of a phishing page popping up on the Internet. The third-party will report the page to the banks and maybe even inform PhishTank, Netcraft, Haute Security, etc.

a. Pay attention to Grammar & Spelling. Many scammers are not native English speakers and this can be very apparent in their messages. Official businesses, on the other hand, usually take special care to use correct grammar and spelling.

b. If you can determine a message is sent from a legitimate organization but you don't want to receive additional messages, you can often use an 'Unsubscribe' from a link at the bottom of the message. However you should not 'Unsubscribe' from scam messages. Many scam messages include this to make them appear more legitimate and use this information to identify 'active' email accounts or possibly obtain sensitive information from users.

c. As an act of retaliation, some users will attempt to respond in anger or string spammers along to waste their time. However, these actions could provoke spammers and attackers into attempting to send more advanced, targeted messages to the user or find other ways to outright attack the user. The best course of action when dealing with spam is to delete and ignore the messages entirely.

IV. CONCLUDING REMARKS

Phishing will always exist, because there will always be ways to trick people. The phenomenon of phishing is growing and the number of variations of techniques implemented demonstrates the high interest in these types of attacks by cybercrime. It's easy to look down upon the victims as being stupid, but often the people who fall for the tricks simply lack proper education about computers. The phenomena must be carefully studied. Fundamental is training people in the secure use of computer tools, the cyber threat that is looming, and how the user can recognize threats in order to avoid serious problems.

REFERENCES

[1] Phishing and Pharming: A Guide to understanding and mitigating the risks. Retrieved from http://www.cpni.gov.uk/documents/publications/2010/2010019phishing_pharming_guide.pdf?epslanguage=en-gb

[2] Phishing Attacks and How To Prevent From Being Hooked. Retrieved from http://www.hongkiat.com/blog/phishing-reports-prevention/

[3] A Novel Anti-Phishing Framework Based on Honeypots: In Proceedings of 4th Annual APWG eCrime Researchers Summit 2009 (APWG eCrime/eCRS 2009, Tacoma, WA, USA, October 20 & 21, 2009), IEEE. Retrieved from http://www.hooklee.com/default.asp?t=Honeypots4AntiPhishing

[4] A Little Spam with Your Bagle? Retrieved from http://labs.m86security.com/2009/06/page/2/

[5] Phishing: A Very Dangerous Cyber Threat. Retrieved from http://resources.infosecinstitute.com/phishing-dangerous-cyber-threat/

[6] Attacking the Phishers: An Autopsy on Compromised Phishing Websites. Retrieved from

http://resources.infosecinstitute.com/attacking-the-phishers/

[7] Phishing. Retrieved from http://labs.m86security.com/tag/phishing/

[8] What Exactly Is Phishing & What Techniques Are Scammers Using. Retrieved from http://www.makeuseof.com/tag/phishing-techniques-scammers/

[9] Security: Types of Phishing Scams & How to Recognize Them. Retrieved from http://grok.lsu.edu/article.aspx?articleid=16680

[10] What are the Different Types of Phishing Attacks. Retrieved from http://www.innovateus.net/science/what-are-different-types-phishing-attacks

[11] Cut the Line on Phishing Scams. Retrieved from http://www.visa.ca/en/personal/securewithvisa/phishing.jsp

[12] A Framework for Detection and Measurement of Phishing Attacks. Retrieved from http://mmnet.iis.sinica.edu.tw/botnet/file/20100524/20100524_1.pdf

Interrogating the Nexus between Globalization and Terrorism in Nigeria

Michael I. Ogu
Department of Political Science and Public Administration,
Babcock University, Ilishan Remo, Ogun State.
ogumyk@gmail.com

Emmanuel C. Ogu
Department of Computer Science,
Babcock University, Ilishan Remo, Ogun State.
ecoxd1@yahoo.com

***Abstract*–** Globalization arguably has a link with terrorism, but the extent to which these two are linked has been a subject of continued research and debate across time and space. The subject of terrorism has been top in the agenda of the United Nations and other regional and even national governments, including Nigeria, since the end of the Cold War, and particularly since the beginning of the millennium. This article explored again the nexus between globalization and terrorism particularly in Nigeria. Qualitatively, this study identified some computer applications and processes that may be helpful in ensuring that the merits of such nexus outweigh its demerits, since globalization has arguably become a part of the human existence. The study concludes by restating the growing indispensable nature of globalization, re-emphasizes the seeming nexus between globalization and terrorism in Nigeria in particular and Africa in general. Policy makers and state actors are advices to proactively explore the merits of the artifacts of globalization in fighting this terrible enemy of development – terrorism, at all levels.

***Key Words*: Nexus, Globalization, Terrorism, Development**

I. INTRODUCTION

The realities of terrorism have come to be accepted globally, regionally and even locally, in the recent past, as opposed to the end of the Cold War and particularly the beginning of the millennium. More nations have been badly hit by the activities of terrorists groups, some of whom are sponsored and armed by influential individuals, governments and organizations across the globe, to the extent that these groups possess much more sophisticated weapons than the state, and even the capacity to completely destroy their target if careful and decisive efforts are not made to curtail the activities of such groups. Although there are several reasons that have been propagated as causes of terrorism, most of which are related to government actions and inactions – policy –, it is no doubt that many groups have emerged globally, regionally and even in Nigeria, whose activities have had devastating consequences on the population as well as economy and development of the state.

Globalization is another phenomenon that has arguably come to be closely connected with terrorism and terrorist activities. The opportunities provided by globalization have served as very effective vehicles for terrorist activities to drive; from transportation, to communication, and even weapons manufacture, terrorist activities have increasingly become much more effective and its consequences much more devastating, as a result of the opportunities and even artifacts provided for by globalization. Hence, this article attempts to interrogate the link between terrorism in Africa generally, and Nigeria in particular, as well as identify ways by which globalization can also be more effectively used in combating terrorism in Nigeria.

II. CONCEPTUAL DISCUSS

It is of important that we discuss a few of the concepts that will be recurrent in this article, so as to afford the reader a more apt understanding of such concepts, especially within the context in which they are used in this study.

Terrorism

The concept of terrorism has been vastly interrogated in the literature and by several scholars as well. [1], citing [2], defined terrorism as; "an anxiety inspiring method of repeated violent action, employed by semi clandestine individuals, groups or state actors for idiosyncratic, criminal or political reasons…" [1]. Terrorism has also increasingly been described as a political phenomenon in contemporary times, especially because it is related to power and influence [3]. [4], also argued that terrorism in its most contemporary understanding is inherently political. [5], identified terrorism as being essentially sectarian in nature, usually associated with the activities of a fringe and ideologically misguided sub-group of a larger religious or political movement.

For this paper, however, our definition of terrorism is not so much about the reason for the violence – idiosyncratic, criminal or political – but the nature and effect of the violence itself. Hence, the definition of terrorism, which will best suffice as our operational definition in this study, is provided by the United States Department of Defence, cited in [6]; "unlawful use of force or violence against individuals or property to coerce end intimidate government to accept political, religious or ideological objectives" [6].

Globalization

Globalization is one concept, among others, that has been laden by several definitions in the literature. [7], speaking on the controversy of defining globalization, observed that;

> "Some claim that it cannot be done, others claim that it would constrain the meaning to do so, and still others have defied these two beliefs and have constructed a working definition. Despite differing opinions about developing a definition, all authors agree on one thing: that defining this term is anything but easy" [7].

Some of the definitions of globalization are discussed below, especially those that might be particularly relevant to the discussion in this paper. [8], defined globalization as ""is the process of world shrinkage, of distances getting shorter, things moving closer. It pertains to the increasing ease with which somebody on one side of the world can interact, to mutual benefit, with somebody on the other side of the world" [8]. This definition is more concerned with the consequences and effects of globalization. Although some equally important aspects of globalization, such as the course, process and equipment (artifacts), have been left out of this definition, it provides very basic and concise information on the very essence of the concept. Perhaps a more detailed definition would include those by [9], [10] cited in [11], among others provided respectively below;

> "the characteristics of the globalization trend include the internationalizing of production, the new international division of labor, new migratory movements from South to North, the new competitive environment that accelerates these processes, and the internationalizing of the state… making states into agencies of the globalizing world" [9].

> "Multiplicity of linkages and interconnections that transcend the nation states (and by implication the societies) which make up the modern world system. It defines a process through which events, decisions and activities in one part of the world can come to have a significant consequence for individuals and communities in quite distant parts of the globe" [11].

> "Globalization is a process that encompasses the causes, course, and consequences of transnational and transcultural integration of human and non-human activities" [7].

[7], further observe that;

> "Globalization involves economic integration; the transfer of policies across borders; the transmission of knowledge; cultural stability; the reproduction, relations, and discourses of power; it is a global process, a concept, a revolution, and "an establishment of the global market free from sociopolitical control". Globalization encompasses all of these things. It is a concept that has been defined variously over the years, with some connotations referring to progress, development and stability, integration and cooperation, and others referring to regression, colonialism, and destabilization. Despite these challenges, this term brings with it a multitude of hidden agendas" [7]

Globalization, no doubt, will not be free of the arguments surrounding its definition in the nearest future. Judging from the various definitions identified above, it may be difficult to relate with the concept in this paper, if an operational definition is not provided. Hence, globalization in this study, will is defined as "the expansion of social, economic, political and technological opportunities and processes, capable of resulting in the progress and/or decline in development within a state".

III. THEORETICAL FRAME

In order to better pose the central theme of this study, consider the diagrammatic summary presented in figure 1:

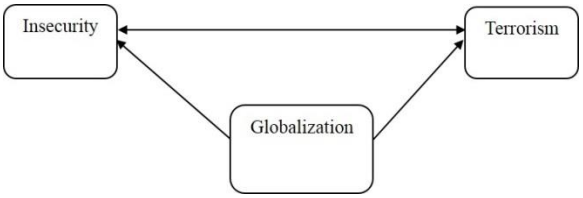


Fig 1: Nexus between globalization, terrorism and insecurity [Source: authors]

It is arguably impossible to discuss terrorism without some emphasis on insecurity, and globalization without emphasis on development. It is already been established in the literature that there appears to be a nexus between globalization and terrorism. The nature of these links is one major objective of this article. From the figure above, it is observed that terrorism and insecurity have a mutually-reinforcing relationship. While globalization influences terrorism and insecurity individually both positively and negatively, it can be argued that these variables may not have any direct influence on globalization.

Using the grievance theory to explain the diagram above, [12] argues that there is a link between horizontal inequalities (insecurity) and armed violence (including terrorism). When these two variables exist in an environment open to globalization, the diagram above may better explains the nature of violence in such a society. [13], also sees armed violence (including terrorism) is a response to relative deprivation or exclusion. When a population becomes politically conscious and aware of their political right which they have been denied for long, it is likely that such a group or population would demand inclusion and political right, either peacefully, but in most cases violently.

IV. TERRORISM AND GLOBALIZATION: THE GENESIS

The genesis of what is now referred to as terrorism is said to date back to at least 2000 years ago [14]. History records that the first acts of terrorism, at least with similar features as terrorism today, were recorded in the activities of the radical Jewish sect – Zealots – active in Judea during the 1st century AD. [14], reported that the Zealots resisted the rule of the Roman Empire's (what is today known as Israel) through a determined violent campaign primarily involving assassination. In these attacks, the Zealot fighters employed the use of primitive daggers – sica – to attack their enemies in broad daylight, making sure that there were people to witness the violent attacks. As a result of these acts of terror, many Roman soldiers and properties were destroyed. Between 1090 and 1272, it is recorded that an Islamic movement known as the Assassins also used similar tactics in their struggle against the Christian Crusaders who had invaded their territory, which is today part of Syria. The Assassins embraced the same notions of self-sacrifice and suicidal martyrdom evident in some Islamic terrorist groups today. According to [15], violence was to them a sacramental or divine act that ensured its perpetrators would ascend to a glorious heaven should they perish during the task. Religion continued to be the major motivation for terrorism globally until the French Revolution (1789-1799) up until the 1800s, when other issues such as nationalism, anarchism, Marxism and other secular political movements emerged to challenge divine right of monarchs to rule.

It can be observed from the above analysis that globalization was arguably not always a component of terrorism as it is evident today. Hence it is important that we also consider the origin of globalization, so as to better identify the point of contact of these two phenomena.

Although historians are likely to describe a longer process of globalization, other practitioners, academics and economists in particular, are likely to be refined to a shorter timeframe. Further, although historians and social scientists have varied starting points for the history of globalization, the most proposed starting points of globalization center on one of three historical events: the industrial revolution about 1800, Western maritime expansion around 1500, or the integration of Asia around 1000 or earlier [16]. Some other arguments place the beginnings of globalization in very recent times; Tom Feiedman, in his book 'The Lexus and the Olive Tree' argues globalization began in the early 1990s resulting from the end of the Cold War and the explosion in the use of the Internet. In another book, 'The Earth is Flat', Friedman refines the beginning of globalization to be the beginning of the 21st century, although he also attests to the fact that the larger process of globalization dates back to 1945, when the world began economic and political integration efforts, which began in the early 1800s and were disrupted by the world wars and the Great Depression in the 1900s [17].

The influence of globalization on terrorism in the recent past cannot be overemphasized. Globalization, it has been argued has definitely changed the nature and dynamics of terrorist activities and networks, globally. Between 1968 and 1989, [18] observed that, the incidents of terrorism were rare, approximately 1.673 per year, as opposed to the period between 1990 and 1996 when there was an increase of 162% compared to the Cold War; 4.389 per year. This

percentage increase in terrorist acts seems to be approaching 200% or even more taking into account the alarming growth of global conflicts in recent times. [19], cited in [18] has observed a more alarming trend in recent times where new acts of terrorism are directed against citizens and not against governments anymore. Globalization has not only redefined terrorism, but encourages religious fundamentalism [18], especially in regions where there are diverse religious and cultural inclinations. [18], went further to observe that 40 years ago there was no single religious terrorist movement worldwide, and in fact, in the 80's only two of 64 known terrorist organizations in the world were animated by religious faith. However, the world today, has continued to witness the proliferation of religious fundamentalism at various levels and in various countries of the world, also engaging in activities that not only devastate human existence, but potent enough to limit, to a large extent, state capacity to protect lives and properties within its territory. Further asserting the growing influence of globalization on terrorism, [18] also reported that "the assassins responsible for the death of the Afghan weapons trafficker, Ahmed Shah Masood were Algerians with Belgian passports, who had visa to enter Pakistan, issued in London. Masood's death shows how radical Islam does not reject globalization" [18].

Globalization is widely understood to be a positive tool for social, economic and political advancement of a society. According to [18], globalizations took the place of the old Cold War system, and it has its own laws and logic, which might influence, directly or indirectly, politics, environment, geopolitics and the economy of any country. Globalization, he continued, has its own technologies, which involves computerization, miniaturization, digitization, satellite communication, optic fiber and Internet, which reinforces the defining perspective of the concept. [20], referred to globalization as "the continuation of modernization and a force of progress, increased wealth, freedom, democracy, and happiness. Its champions present globalization as beneficial, generating fresh economic opportunities, political democratization, cultural diversity, and the opening to an exciting new world" [20].

The mindless terror on the economic, political and military symbols of the United State on the 11th of September 2001 (9/11) and the subsequent global war on terror seems to bring to the fore the double-faced consequences of globalization. [20] observed that;

> "The terrorist acts on the United States on September 11 and subsequent Terror War dramatically disclose the downsides of globalization, the ways that global flows of technology, goods, information, ideologies, and people can have destructive as well as productive effects. The disclosure of powerful anti-Western terrorist networks shows that globalization divides the world as it unifies, that it produces enemies as it incorporates participants. Globalization links people together and brings new commonalties into experience just as it differentiates them and produces new inequalities. Likewise, while it connects and brings into global networks parts of the world that were isolated and cut-off, it ignores and bypasses other regions. The events disclose explosive contradictions and conflicts at the heart of globalization and that the technologies of information, communication, and transportation that facilitate globalization can also be used to undermine and attack it, and generate instruments of destruction as well as production [20].

We have observed from the above analysis that terrorism globally, is largely influenced by globalization, whether directly or indirectly. Several opportunities provided for by globalization, perhaps for the development of society and advancement of the economy, have turn out to be mediums of terror and war on the state and unfortunately on civilians too. This growing global trend has also increased the level of insecurity in some of the countries that have been badly affected by these terrorist activities; millions of individuals have been forced to migrate from their countries, posing significant security challenges for their host countries, many other individuals who have managed to remain in these terrorized countries have continued to live in fear of losing their lives, while also enduring the scourge of hunger, deprivation, abuse of human rights among other forms of insecurity. Nigeria is one of the countries in Africa that is waging an ongoing war on the Boko Haram terrorist group, and the next section will pay particular interest to the nature of terrorism in Nigeria.

V. TERRORISM: THE NIGERIAN EXPERIENCE

Africa's largest economy, housing over 170 million people, who account for Africa's largest market and a major attraction of foreign individual and government investment, and occupying a land area of about 925,000 square kilometers over 356,700 sq.

meters, Nigeria has, in recent times, been under serious terror attacks. Although many argue that the activities of various terrorist groups in Nigeria do not come as a surprise, owing to the colonial configuration of the territory and the nature of governance. The colonial amalgamation of over 200 unique and hitherto independent nations into one Nigeria in 1914 has continued to throw up peculiar issues – class, ethnicity, religion, exclusion, among others – which have come to shape the nature of political, social, economic and even religious activities across the country.

[21], argued that signs and symptoms of terrorism are not a new occurrence in Nigeria, as they have been for a long time before now, because the causal factors of such terror have been a part of the Nigerian state for a very long time before now. [22], also corroborated this view by asserting that "any society that seeks to achieve adequate security against the background of acute food shortage, population explosion, low level of productivity and per capita income, low technological development, inadequate and insufficient public utilities and chronic problems of unemployment; (religious intolerance and criminal politicking) has a false sense of security" [22]. These conditions identified as background for a false sense of security have been the defining characteristics of the Nigerian state for so many decades, and unfortunately, they appear to remain for a very long time to come.

[23], advanced the same argument by observing that although Nigeria remained a relatively peaceful and terror free country in the early 1960s and 1970s, there were certain incidents in the 80s and 90s that heralded the activities of terrorist groups in Nigeria. Some of these incidents include: the killing of Mr Dele Giwa, by a 'Letter Bomb' in October 1986, which marked the beginning of violent killing and use of Improvised Explosive Devices (IEDs) in Nigeria; the hijack of a Nigeria Airways aircraft in October 1993 by the Movement for the Advancement of Democracy, following the annulment of Chief M.K.O.Abiola's presidential election in June 1993; and a vicious bomb blast that ripped Shed 6 of Ilorin Stadium in August 1994; the bomb attack on the car of the then Chief Security Officer of Federal Aviation Authority of Nigeria, Dr Omoshola; the attack on the convoy of the former Military Administrator of Lagos State, Brig Gen MB Marwa (Rtd) between 1996 and 1998, among others.

Since the beginning of the millennium, many of the terror attacks in Nigeria have been traced to the Jama'atu Ahlis Sunna Lidda'awati wal-Jihad, meaning "People committed to the propagation of the Prophet's teachings and jihad", popularly called Boko Haram. Believed to have been established in the 1960s, the group, whose mission is reflected in its name, began to gain so much attention in and out of Nigeria in the year 2002. Although the Nigerian Police is believed to have begun official investigation of the group in July 2009, the activities of the group, prior to and since 2009, have claimed a lot of innocent lives in Nigeria, loss of properties and led to other manifestations of insecurity in the country – migration, arms trafficking, human rights abuse, human trafficking, kidnapping, among others. In 2014, following the attack and abduction of over 200 school girls from Chibok village in Borno State, the United State officially included the Boko Haram group on the list of terrorist groups across the globe. Some of the activities of the Boko Haram terrorist group in the 2nd and 3rd quarter of 2014 alone have been summarized in the table below;

Table 1: Boko Haram Terror Activities from April – September, 2014.

Date	Location	Dead	Event/Type of Violence
17-Sep	Kano	**13**	Two suicide bombers blew up a school, killing 13 students and lecturers in addition to the two bombers.
15-Sep	Michika	**16**	Militants riding on horses rode into the town and shot indiscriminately. At least 50 women and an unknown number of children were also abducted.
9-Sep	Adawama	**24**	Boko Haram shot and wounded ex-President Obasanjo's son. Also killed 24 soldiers.
1-Sep	Bama		Heavy casualities on both side (CHECK BACK ON THIS)
August			
30-Aug	Borno State	**14**	Boko Haram invaded four villages in Borno State, killing 14.
21-Aug	Buni Yadi, Yobe		Boko Haram has taken control of Buni Yadi.

Date	Location	Dead	Event/Type of Violence
20-Aug	Borno State		35 Police officers are missing after two attacks on police training camps.
17-Aug	Maforo	**4**	Suspected Boko Haram militants stormed the town of Maforo and shot 4 dead.
17-Aug	Durwa	**6**	Suspected Boko Haram militants stormed the town of Durwa and shot 6 dead.
14-Aug	Borno State	**28**	At least 97 men and boys were kidnapped, while 28 people were killed and 25 injured during a string of attacks over a course of a week,
9-Aug	Limankara	**5**	Pastor Musa Gaiyo of Church of Christ in Nations, his father, and three of his children were killed by Boko Haram.
5-Aug	Gwoza, Borno	**100**	Militants raided the town, and reportedly kidnapped or killed the Muslim traditional leader. Many were also Christians.
2-Aug	Borno State	**1**	Soldier killed by Boko Haram while on a military contingent in Borno.
July			
30-Jul	Borno State		5 Churches bombed.
29-Jul	Yobe	**13**	13 died and many others injured when 2 mosque bombings occurred.
29-Jul	Yobe	**8**	A bridged was bombed killed 8 people. The soldiers guarding the bridge were outnumbered by Boko Haram members.
28-Jul	Kano		A female suicide bomber killed herself across from a Shoprite supermarket. No one else was injured.
28-Jul	Kano	**3**	A female suicide bomber killed herself and 3 others while waiting to buy kerosene.
28-Jul	Katarko, Yobe		A Muslim cleric, his wife, and two others were kidnapped by Boko Haram.
27-Jul	Kolofata, Cameroon	**15**	Members of Boko Haram broke into the house of Prime Minister Amadou Ali, kidnapping his wife and killing an unknown number. They also kidnapped the mayor of the town and six of his family members. At least 15 people were killed between the raids of the two homes.
27-Jul	Shafa	**2**	Suspected Boko Haram members killed two residents who had just fled Damboa.
27-Jul	Hong	**30**	Suspected Boko Haram members rode into Hong and shot haphazardly, killing at least 30.
27-Jul	Kano	**5**	Suspected Boko Haram members bombed a church, killing 5.
27-Jul	Kano		A 15-year-old female suicide bomber set off a bomb near a temporary university site. No one else was injured.
26-Jul	Shaffa District, Hawul	**2**	Boko Haram members beheaded the victims.
24-Jul	Kano	**1**	A busy bus station was bombed, killing one and injuring eight.
23-Jul	Kaduna	**44**	A double bombing occurred in Kaduna, targeting the former head of state, and a prominent moderate Muslim cleric. Both escaped unharmed.
23-Jul	Borno State	**12**	Suspected Boko Haram members dressed in camoflauge raided a village, killing the traditional leader and 11 others.
18-Jul	Damboa, Borno		Raids on Damboa caused as many as 15,000 residents to flee.
18-Jul	Damboa, Borno	**21**	Islamic militants raided the village in the early morning, killing at least 21.
17-Jul	Gombi, Adamawa		German aid worker and two children of a Muslim cleric have been kidnapped.
14-Jul	Dille, Borno	**38**	Militants raided a Christian village firing indiscrimately, throwing grenedes, burning the whole town and three churches. Deaths included the pastor of Redeemed Church of God, Taiwo Dokun.
5-Jul	Damboa, Borno	**20**	Suspected Boko Haram members attacked a military base in Damboa killed no fewer than 12 soldiers, 4 police officers, and 4 civilians.
1-Jul	Maiduguri, Borno	**56**	A bomb planted in a van carrying charcoal exploded in a busy market, killing 56 people.
June			
29-Jun	Kwada, Ngurojina, Karagau, Kautikari	**54**	Four villages were stormed by gunmen hurling explosives into churches, torching buildings, and shooting at worshippers.

Date	Location	Dead	Event/Type of Violence
25-Jun	Abuja	**21**	A blast occurred right as people were about to watch the Nigeria vs Argentina soccer game, killing 22 people and wounding dozens.
24-Jun	Bulamburin, Borno	**12**	Militants attacked a military checkpoint, killing 12 soldiers but also losing 25 of their own.
23-Jun	Kano	**8**	A bomb went off around 1 pm at the School of Hygiene, killing 8 and injuring a dozen more.
18-Jun	Kummabza, Borno	**30**	Gunmen dressed as soldiers invaded the village killing 30 men and abducting 60 females. The village was burned down, while food was also looted. The women escaped July 3rd and 4th
15-Jun	Gumsa, Yobe		Gunmen stormed a mosque during prayer and abducted the village chief.
15-Jun	Daku	**15**	Gunmen opened fire on market vendors, killing 15. They also burned down homes and a church.
8-Jun	Abuja	**1**	Female suicide bombers made attack on Nigerian barracks. One officer was killed as the bomb went off as he was searching the woman.
5-Jun	Borno State		Gunmen arrived in the remote village around noon and forced 20 women into a vehicle at gunpoint. Three men were also taken when they tried to stop the kidnapping.
5-Jun	Madagali, Adamawa	**2**	2 killed in gunfire exchange.
4-Jun	Maidugur, Borno	**45**	Members of Boko Haram posing as preachers killed 45 people as they surrounded the insurgents to hear them preach.
2-Jun	Attagara, Mandara Mountains	**250**	Men in military uniforms arrived in the village, annoucing that they were there to assess the security situation. When people gathered, the insurgents opened fire.
1-Jun	Mubi, Adamawa	**18**	Suspected Boko Haram members bombed the predominantly Christian area of Mubi. Victims included individuals that were watching soccer at a nearby bar, as well as people at a nearby soccer game. Witnesses say that the bomb was planted in a three-wheeled vehicle outside of the bar.
1-Jun	Borno State	9	Suspected Boko Haram gunmen invaded a church during worship service and opened fire around 9:30 am, killing 9 people.
May			
15-May - 12-Jun	Kunde	7	7 people were killed, 2,500 were forced to flee to Cameroon and the hills, and all of the communities' houses and churches were burnt.
15-May - 12-Jun	Baladigavuraza	3	3 people were killed, while 2,000 people fled to Cameroon and 500 fled to various parts of Nigeria.
15-May - 12-Jun	Angurva		An unknown number of people were killed, while nearly 4,000 have been forced to flee to Cameroon.
15-May - 12-Jun	Kaghum	20	Boko Haram killed more than 20 people, and more than 3,000 were forced to flee to Cameroon and other parts of Nigeria.
15-May - 12-Jun	Pege	5	5 people were killed while 2,000 were forced to flee to Cameroon and other parts of Nigeria.
15-May - 12-Jun	Barawa	20	20 people were killed and all of the community's houses and churches were burned. Survivors fled to Cameroon.
15-May - 12-Jun	Chikide	2	2 people were killed, while about 1,500 were forced to flee to Cameroon.
15-May - 12-Jun	Arboko	18	18 people were killed, all houses and churched were burned to the ground, and about 2,000 have fled to Cameroon.
15-May - 12-Jun	Chinene	21	21 people were killed, while about 1,500 were forced to flee to Cameroon and other parts of Nigeria. 1,000 are still in the mountains.
15-May - 12-Jun	Kirawa		The whole village was taken over by Boko Haram; villagers were forced to flee to Cameroon.
15-May - 12-Jun	Pulka	30	30 people were killed, including the District Head. Many members have fled to Cameroon.
15-May - 12-Jun	Kwadale		An undetermined number of people were killed, while about 2,000 fled to Cameroon.

Date	Location	Dead	Event/Type of Violence
15-May - 12-Jun	Vale	3	The whole community was chased out by Boko Haram, their village was looted and 3 people were killed.
15-May - 12-Jun	Ashigashiya	10	10 people were attacked and killed, while the rest of the community fled to Cameroon.
15-May - 12-Jun	Zamgba	13	13 people were killed and about 1,700 were forced to flee to Cameroon.
15-May - 12-Jun	Jibrili		An unknown number of people were killed, while Christian houses and churches were torched. Survivors have fled to Cameroon.
15-May - 12-Jun	Ganjara	46	46 were killed and more than 2,500 fled to Cameroon.
15-May - 12-Jun	Agapalawa	15	15 were killed while more than 4,000 fled to Cameroon.
15-May - 12-Jun	Halaghwa		An undetermined number of people were killed, while close to 2,500 were forced to flee to Cameroon and other Nigerian states.
15-May - 12-Jun	Amuda	17	17 people were killed and more than 4,000 were forced to flee to other places.
15-May - 12-Jun	Gava	26	26 were killed and about 5,000 fled to Cameroon and other Nigerian states. All the houses and churches were torched.
15-May - 12-Jun	Ngoshe, Borno	46	46 were killed and nearly 6,000 had to flee to Cameroon. All the houses and churches were torched.
15-May - 12-Jun	Bokko Timta, Borno		Approximately 1,500 Christians had to flee to Cameroon after Boko Haram raided their community.
15-May - 12-Jun	Kugyly		2,500 villagers had to flee to Cameroon after Boko Haram visited.
15-May - 12-Jun	Bokko Wizhe	4	Insurgents opened fire indiscriminately in a predominately Christian village. About 2,000 survivors have fled to Cameroon.
31-May	Kala Balge, Borno	40	Boko Haram militants raided a community, killing 40 people, after the residents had repelled the militant group a week prior.
30-May	Zhur, Borno	3	Nigerian traditional leader, Shehu Mustapha Idris Timta, Emir of Gwoza, was shot dead in his car while on his way to a funeral. Two of the police escorts were also killed.
28-May	Borno State	35	Members of Boko Haram raided three villages in Borno State in the early morning. The insurgents arrived on motorcycles and in all-terrain vans. They shot down residents and burned down homes, killing at least 35 people.
26-May	Buni Yadi, Yobe	31	Gunmen arrived in an armored personnel carrier and six Toyota pickup trucks, and raided a military base and adjacent police barracks. The militants fired rocket propelled grenades at the two bases, killing 17 soldiers and 14 police officers.
26-May	Gowza, Borno	8	Boko Haram rebels burned down seven churches, numerous houses, and killed 8 Christians.
25-May	Kumuta, Borno	24	Dozens of militants riding motorcycles stormed the village early Sunday morning and opened fire as the residents headed to the weekly market.
25-May	Waga, Adamawa	20	Millitants stormed the Christian village riding in trucks and on motorcycles. They burnt several houses and opened fire on residents, killing 20 people.
25-May	Gwoza, Borno	21	Boko Haram members killed 21 members of the congregation of Church of Christ in Nations during the worship service.
24-May	Jos, Plateau State		Only the bomber was killed.
21-May	Chukku Nguddoa, Borno	29	Suspected Boko Haram militants shot dead 29 farmers as they tilled their fields, also wounding 10 others.
20-May	Northeastern Nigeria, near Chibok	48	Militants raided three villages near Chibok Tuesday overnight into Wednesday. The extremists set the village homes on fire and killed 48 residents.
20-May	Jos, Plateau State	118	Two bombs exploded during the local rush hour killing mostly women and children. One bomb was planted in the back of a cargo truck and went off in the Terminus market. The second explosion involved a minibus. The first explosion killed 7 Christian students. The market is

Date	Location	Dead	Event/Type of Violence
			predominantly Christian women selling fruits, vegetables, and various items.
18-May	Kano, Kano	4	Suicide bomber in vehicle had sought directions from Christians to local churches. When Christians became suspicious of the bomber's intentions, the bomb detonated.
14-May	Chibok, Borno	12	Boko Haram fighters ambushed Nigerian troops, which killed 12 soldiers.
9-May	Yola, Adamawa	30	Islamic extremists blew up a bridge, abducted the wife and two children of a retired police officer, and killed at least 30 people.
5-May	Gamboru Ngala	332	Dozens of militants wearing fatigues entered the town firing indiscriminately at people and torching houses. 17 of those killed were officers.
4-May	Warabe		Militants kidnapped at least 8 girls between the ages of 12 and 15.
2-Mar	Mafa, Borno	33	Gunmen stormed Mafa village in Borno, around 50 km (30 miles) east of the state capital Maiduguri, at around 8 p.m. on Sunday, shooting fleeing civilians and throwing explosives at occupied houses. The insurgents were dressed in military fatigues and the whole town was razed. 20 people were also injured. An unknown number of observers were killed during nightly prayer when explosives were thrown into the mosque.
April			
21-Apr	Borno State	2	Suspected Boko Haram gunmen on motorcycles set up a road block near Mafa town. They ambushed the vehicle of two APC leaders and shot them.
16-Apr	Borno State	18	Around 7 am, gunmen attacked Wala village, leaving 18 dead and many injured.
15-Apr	Borno State	2	Gunmen attacked Sabon-Kasuawa village and killed a local monarch (district head) and his guard at his palace.
14-Apr	Borno State		Suspected Boko Haram members abducted 276 secondary school girls in the town of Chibok (in a Christian region). Some sources suggest that the girls may be used as war brides or human shields. Government and civilian search parties are still combing the Sambisa forest.
14-Apr	Nassarawa State	200	Police reported the blast from a bomb at an Abuja bus station affected 16 buses plus smaller commercial vehicles. There were over 100 victims in all. The number of dead rose from 71 to 75 to 200, and the wounded from 123 to 141.
14-Apr	Abuja	75	A bomb blast killed 71 individuals and injured 124 at the crowded bus station of Nyanya Motor Park. The bomb also destroyed more than 30 vehicles, which caused secondary explosions.
10-Apr	Borno State	60	Gunmen attacked Kala Balge town and killed three villagers, allegedly dressed in military attire. There were reportedly 60 people left dead after the attack.
10-Apr	Borno State	8	One of the attacks targeted a Teachers' College in Dikwa in which 7 people were killed and the school library and administrative block were burned.
10-Apr	Borno State	207	Between April 10 and 11, at least 207 people were killed in Borno, including JAMB candidates on their way to exam centers.
6-Apr	Yobe State	17	Boko Haram militants attacked a northern Nigerian village, leaving 17 dead and many houses and cars in flames. Muslim worshippers at the village mosque were among the dead, according to the governor's spokesman.

Source: http://factsnigeriaviolence.org/spreadsheet/2013-2/ accessed Sept. 29, 2014 [24].

VI. FIGHTING TERRORISM IN NIGERIA

The literature on counter-terrorism strategies has continued to grow, providing government nationally, regionally and globally, with adequate information and knowledge on how to curb the challenge of terrorism and minimize its effects on development and progress within states. Before we go forward to explore any globalization induced counter-terrorism strategies, we will first attempt an overview of the popular strategies that exist in the literature.

[25], cited in [21] proposed what he regarded as the four **"Ds**" of fighting terrorism: 'Defeat terrorists and their organizations; Deny sponsorship, support and sanctuary to terrorists; Diminish the underlying conditions that terrorist seek to exploit, and Defend citizens and interests at home and abroad'. [21], argued that, considering the complexities of Nigeria's security situation, counter terrorism should be more of an interconnected and interdependent strategy which, should be built around community resilience against terrorism, enhance cooperation among law enforcement agencies and strengthening judicial institutions. Some counter terrorism strategies for Nigeria are suggested subsequently;

1. **The Meghalaya Model:** This model, according to [26], is a multi-sector approach which involves government, law enforcement, judiciary and civil society organizations that creates a comprehensive framework for combating trafficking in North East Asia in 1999 under the **five "Ps"** Prevention, Protection, Policing, Press and Prosecution. Through the process of the **five "Rs"** Reporting, Rescue, Rehabilitation, Repatriation and Re-education, the model has been used to track and rescue trafficked children in North East Asia; to facilitate rehabilitation, and making sure that survivors do not get re-trafficked. [21], argued that through a collaborative and far-reaching security network of stakeholders in counter terrorism and human trafficking (Security and Intelligence Agencies, NAPTIP, the media, Civil Society and NGOs), terrorists, mercenaries and terror suspects are more likely to be arrested, and will face prosecution.

2. **Public Private Partnerships (PPP) Model:** Also emphasizing cooperation and interrelation, the PPP model provides for a private sector participation in the counter terrorism efforts, especially where state resources and expertise are limited, although combating terrorism is primarily a responsibility of the state. Despite the growing need of a partnership in security, Jonathan Lucas, Director of the United Nations Interregional Crime and Justice Research Institute (UNICRI), has identified two major limitations to such partnerships; first is a lack of legislation to facilitate information exchange between the public and private sectors, and second is the absence of incentives to encourage the private sector to invest in the protection of vulnerable targets, thereby considering funds for security PPP initiatives as costs rather than an investment [27].

3. **Youth Empowerment Plus:** It is evident in the literature that youth unemployment is a veritable tool for influencing armed conflicts, and sponsors of violence leverage on youth unemployment, while those of development target job creation as a major antidote for armed conflicts. [28], [29], [30], among others have argued that donors globally have identified job creation as an urgent priority, both in the field of peace-building and in efforts to foster economic development. [21], however argues that job creation is only a part of an integrated or comprehensive armed conflict or Anti Violence Response (AVR) strategy in conflict zones. Such youth strategies "should be based on 'holistic', 'comprehensive' and 'integrated' approaches that go beyond simply addressing a lack of economic opportunities, but seek to address the more complex array of factors (like belief system, religious, political and ethnic fundamentalism) that make youth entrepreneurs of violence rather than peace-seeds" [21]. Hence, the initiative of 'Peace Ambassadors' and other national recognitions and awards might prove a useful counter terrorism strategy under this model.

4. **Applications and Infrastructure:** [31], investigated the role of ICT in the national security issues facing developing and underdeveloped countries, with the Nigerian scenario as a case study; they opined that the national security challenges of most developing and underdeveloped countries are traceable to poorly guarded and managed cyberspace boarders, as well as inefficient monitoring of the applications and infrastructure that feature in this cyber environment. ICT, being one of the principal tools that has fostered globalization over the past decade, with impacts that have generated global ripples, has been proven to

be a force for good as well as evil; hence, it is important that every nation takes upon herself the responsibility of guarding their cyberspaces as well as the applications, transmissions and infrastructure that participate in this evasive arena. The need for an eye within every national cyberspace, for monitoring and security, is brought to the fore.

VII. CONCLUSION

Globalization is a phenomenon that has come to stay globally and considering the ease and dynamics it has brought to processes and procedures, it is not likely to go away anytime soon, at least not in the near future. We have established from this study the link between globalization and the growing realities of terrorism globally and in Nigeria in particular. Although terrorist activities have begun from as early as the 17th century, the emergence of globalization and it's assistant tools has redefined the practices and processes of terrorism and has, in fact, increased the magnitude of the devastating effect of terrorist activities, especially in developing nations in Africa. The complex nature of political, economic and social processes, as well as constantly widening gaps in policy issues, in many of these African countries are identified as the major causes of terrorist activities, which the state needs to combat, sometimes assisted by globalization. The non-availability or inadequacy of technological infrastructures and applications in most of these states, like Nigeria, has also further exacerbated the challenge of combating terrorism. Hence, a way forward in mitigating these challenges may well lie in the need for a comprehensive shift in paradigm with a more pragmatic focus that would touch on all the processes of political, economic, social and policy endeavors in these nations.

REFERENCES

[1] Barga, T. (2012), "Towards a theology of peace: A panacea to terrorism and violence". Jos Studies, Vol. 20, pp. 1-24

[2] Schmid, A.P. (1998), **Thesaurus and glossary of early warning and conflict prevention terms**, Forum on Early Warning and Early Response.; Synthesis Foundation Publisher: London, UK :

[3] Okoli, A. C. & Iortyer, P. (2014), Terrorism and Humanitarian Crisis in Nigeria: Insights from Boko Haram Insurgency, Global Journal of Human-Social Science, Volume 14 Issue 1 Version 1.0 Year 2014

[4] Hoffman, B (1998), Inside terrorism, Columbia University Press, London

[5] Nchi, S.I (2013), Religion and Politics in Nigeria: The Constitutional Issues, Jos: Green-world Publishing Co. Ltd.

[6] Eze, C.M. (2013). "Boko Haram Insurgency: A Northern Agenda for Regime Change and Islamization of Nigeria, 2007-2013". Global Journal of Human Social Sciences: F (Political Science), Vol 13(5:1.0), pp.87-98.

[7] [7] Al-Rodhan, N.R.F. and Stoudmann, G. (2006), Definitions of Globalization: A Comprehensive Overview and a Proposed Definition, Program on the Geopolitical Implications of Globalization and Transnational Security, Geneva Centre for Security Policy, June 19, 2006

[8] Larsson, T. (2001), The Race to the Top: The Real Story of Globalization, U.S.: Cato Institute, 2001, p. 9.

[9] Cox, R. (1994), "Multilateralism and the Democratization of World Order", paper for International Symposium on Sources of Innovation in Multilateralism, Lausanne, May 26-28, 1994

[10] McGrew, A. (1990), "A Global Society" in Stuart Hall, David Held, and Anthony McGrew, Modernity and Its Futures, Cambridge: Polity Press, 1990.

[11] Reich, S. (1998), What Is Globalization? Four Possible Answers, Working Paper #261 – December 1998. P. 5

[12] Abimbola, J.O, and Adesote, S.A, (2012), Domestic Terrorism and Boko Haram Insurgency in Nigeria, Issues and Trends: A Historical Discourse, Journal of Arts and Contemporary Society Vol. 4, September 2012

[13] Rapport, D. (1984), Fear and Trembling Terrorism in Three Religious Traditions, American Political Science Review, 78 (3), 658-677.

[14] Northrup, D. (2003) Globalization in Historical Perspectives, World System History, Encyclopedia of Life Support Sysytems (EOLSS)

[15] Friedman, L. T. (1999), The Lexus and Olive Tree – Understanding Globalization, Anchor Books, New York

[16] http://gseis.ucla.edu/faculty/kellner.

[17] http://www.eolss.net/sample-chapters/c04/e6-94-16.pdf

[18] Stibli, F. (2010), Terrorism in the context of globalization, AARMS Vol. 9, No. 1 (2010) 1–7

[19] Coker, C. (2002): Globalization and terrorism, http://www.g8.utoronto.ca/conferences/2002/tokyo/cocker.pdf; September 30, 2014.

[20] Kellner, D. 2005. "Globalization, Terrorism, and Democracy: 9/11 and its Aftermath."

[21] Omale, D.J.O. (2013), Terrorism and Counter Terrorism in Nigeria: Theoretical Paradigms and Lessons for Public Policy, Canadian Social Science, Vol. 9, No. 3, 2013, pp. 96-103

[22] McNamara, R. S. (1990), The essence of security: Reflections in office, NY: Harper and Row.

[23] Obene, W. R. (2012).'Home-grown terrorism: An emerging challenge to Nigeria's national security. Presentation to Haske Biyu 2012 participants at the Armed Forces Command and Staff College, Jaji, Kaduna, 03/04/2012.

[24] The Facts On Nigeria Violence, Incidents, Reports, Statistics & Links, http://factsnigeriaviolence.org/spreadsheet/2013-2/

[25] Piombo, J. R. (2007), Terrorism and U. S. Counter-Terrorism Programs in Africa: An Overview, Strategic Insights, 6(1).

[26] Kharbhih, H. (2010). Meghalaya Model -Combating Human Trafficking through a Comprehensive Framework that Engages Stakeholders, Impulse NGO Network

[27] Lucas, J. (2012), Promoting public private partnerships, The Beam UN CTITF (Vol. 5), p.8.

[28] World Bank (2008), Youth and employment in Africa: The potential, the problem, the promise. Africa

development indicators, http://www.siteresources.worldbank.org/INTSTATINAFR/Resources/ADI-200809-essay-EN.pdf.

[29] International Labor Organization (2010), Socio-economic Reintegration of Ex-combatants. Guidelines, programme for crisis response and reconstruction. Rerieved from: http://bit.ly/a0iDxx.

[30] United Nations (2009), United Nations Policy for Post-conflict Employment Creation, Income Generation and Reintegration, United Nations, Geneva, Rerieved from http://bit.ly/bN4K65

[31] Ogu, E.C. & Oyerinde O.D. (2014), ICT and National Security in Developing and Underdeveloped Countries – The Good, the Bad, and the Ugly: A Case Study of Nigeria's Cyberspace, International Journal of Computer Science and Information Technologies, Vol. 5 (4), 2014, 5625-5633

Robustness and Security Issues in Digital Audio Watermarking

Agbaje M. Olugbenga
Department of Computer Science
Babcock University, Nigeria

Adebayo A.Olawale
Department of Computer Science
Babcock University, Nigeria

Abstract- Digital watermarking is the process of embedding information into digital media such as image, audio or video data to easily identify the copyright ownership of the media. Such information is for different purposes, such as for broadcast monitoring, copyright protection, source tracking, and piracy deterrence and many others. It is usually embedded in a way that makes it difficult to be removed. There are extensive literatures about watermarking algorithms and methods as well as possible attack techniques. This work used extensive literature review about watermarking to have a high-level overview on new watermarking trends security and robustness issues in order to be able to establish a clear frontier between them. It concludes by discussing possible future research directions.

Keywords- Robustness, Security, Digital watermarking.

I. INTRODUCTION

Digital watermarking is the process of embedding information into digital media such as image, audio or video data to easily identify the copyright ownership of the media. Such information is embedded for many different purposes, such as broadcast monitoring, copyright protection, source tracking, and piracy deterrence, amongst others. It is usually embedded in a way that makes it difficult to be removed. It is a competitive technology for monitoring content consumption and detecting fraudulent use of copyrighted material. It embeds metadata in multimedia content in a robust and imperceptible way [1][2][3].

Digital watermarking, traditionally modelled as communication with side information, is generally considered to have important potential applications in various scenarios such as digital rights managements. However, the current literature mainly focuses on robustness, capacity and imperceptibility, which lacks systematic formal approach in tackling secure issues of watermarking [4]. Papers presenting a complete analysis of watermarking security are extremely rare [5][6].

This paper considers security and robustness of digital watermarking as an effective instrument against piracy and also metrics for the measurement of the two terms. Digital watermarking studies have always been driven by the improvement of robustness. Robust watermarking is defined as a communication channel multiplexed into original content [7]. On the contrary, security has received little attention in the watermarking community.

The metrics involved in robustness measurement are quite developed and well known. The metrics for watermarking security is still at the developmental stage. In this study, information theory is used to provide both theoretical and practical analysis of security offered by watermarking on the theoretical side, the measure of security is estimated by means of information theory tools for all possible attacks. On the practical side, simulations assess the security levels derived in the theoretical part. The first difficulty is that security and robustness are neighbouring concepts, which are hardly perceived as different. Indeed, there are a number of desirable characteristics that a watermarking technique should exhibit.

The methodology of the research is through literature search and case study. The rest of the paper presents a brief overview of the digital watermarking and the value of metrics determination. It is followed by concept of robustness. We describe common security concepts and analysis and finally conclude with the relationship between security and robustness and direction for future works.

A. Digital watermarking

A digital watermark is a transparent, invisible information pattern that is inserted into a suitable component of the data source by using a specific computer algorithm [8][9].

The main advantages of digital watermarking are the following:

- Database is not needed. Message hidden should be self sufficient to screen the contents. This allows placing watermark decoders in many different points in the content distribution chain. For instance, the Bluray disc players look for watermarks in the audio channels to spot illegal camcording of theatre movies [2].
- Two versions of a Work can be watermarked with two different messages. This allows identifying, for instance, which distribution chain a piece of content went through (IPTV, DVD or theatre for example)[10]

Its main drawbacks are as follows:

- The content is modified by the watermark embedder. It resorts to complex human perception models to shape the embedding distortion under the just noticeable difference to remain invisible or inaudible to humans.
- Placing watermark decoders here and there raises the issue of keeping the secret key secret.

From the viewpoint of a software engineer, digital content is no more than a 0's and 1's bit string. Anyone with a handy editing tool can manipulate the digital content freely and easily. The modified content can be disseminated rapidly without any limitation. To protect copyright and/or content authentication/integrity, cryptography primitives such as hash functions and encryption/decryption algorithms are conventional tool[11].

There are three classes of watermarking methods:

- A **fragile watermarking scheme,** which detects any manipulation in order to guarantee the content integrity.
- A **semi-fragile scheme,** which detects unacceptable image manipulations, while allowing lossy compression or other acceptable modifications.
- A **robust watermarking scheme,** which resists content modification to the highest degree. This method prevents an attacker from removing the watermark unless the quality of the content is reduced greatly [12].

Multimedia information hiding (MIH) techniques have aimed to help preserve the values of multimedia information such as text, digital-audio, images, and video, help to hide imperceptible marks such as copyright notice into them, or even help to prevent their unauthorized copying. MIH techniques composed of content protection of multimedia information schemes such as watermarking and steganography the hiding of multimedia information in other multimedia information. As it is possible to use MIH techniques together with cryptographic techniques, they are applicable for secure content authentication such as fingerprint [13].

Typical applications based on MIH techniques have recently been attracted as state-of- the-art techniques for copyright protection and these have been realized as digital watermarking methods. In particular, there have recently been serious social issues involved in protecting the copyright of all digital-audio content by preventing it from being illegally copied and distributed on the Internet. Digital-audio watermarking has been focused on as a state-of-the-art technique enabling copyright protection, as shown in Figure[13].

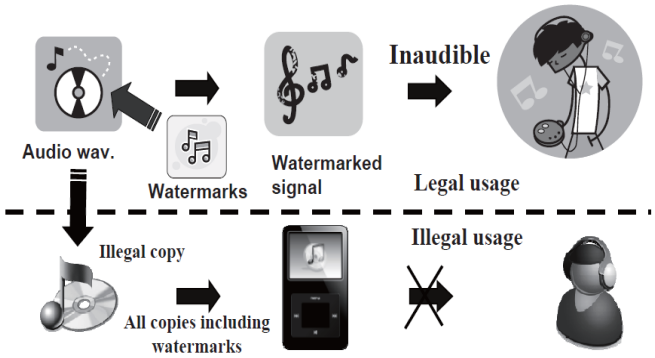


Figure 1: Schematic illustration of digital-audio watermarking [13].

The aim is to embed codes to protect the copyright in audio content that is inaudible to and inseparable by users, and to detect embedded codes from watermarked signals. However, in contrast with watermarking techniques for image/video contents, there seems to be no complete or successful method for digital audio contents in industrial applications [13][6].

In general, audio watermarking methods must satisfy three requirements to provide a useful and reliable form of copyright protection:
(a) Inaudibility (inaudible to humans with no sound distortion caused by the embedded data).
(b) Confidentiality (secure and undetectable concealment of embedded data),
(c) Robustness (not affected when subjected to techniques such as data compression)

The first requirement (inaudibility) is the most important in the method of audio watermarking because this must not affect the sound quality of the original audio. If the sound quality of the original is degraded, the original content may lose its commercial value. The second requirement (confidentiality) is important to conceal watermarks to protect copyright, and it is important that users do not know whether the audio content contains watermarking or not. The last requirement (robustness) is important to ensure the watermarking methods are tamper-proof to resist any manipulations by illegal users.

B. Value of Metrics

The metrics involved in robustness measurement are quite developed and well known. The metrics for watermarking security is still at the developmental stage. In this study, information theory is used to provide both theoretical and practical analysis of security offered by watermarking. On the theoretical side, the measure of security is estimated by means of information theory tools for all possible attacks.

The requirement to measure information security performance is driven by regulatory, financial and organizational reasons. A number of existing laws, rules, and regulation cite information performance measurement in general and information security performance measurement in particular, as requirements [14]. From literature we have found out that the robustness of watermarking system has a well defined set of metrics, measurements of performance for security is not as developed. Hence the paper takes a look at robustness measure and security as it is today.

C. Value of Security Metrics

A widely accepted management principle is that an activity cannot be managed if it cannot be measured. Security falls under this rubric. Measurement of performance can be an effective tool for security managers to discern the effectiveness of various components of their security programs, the security of a specific system, product or process, and the ability of staff or departments within an organization to address security issues for which they are responsible. Measurement of performance can also help identify the level of risk in not taking a given action, and in that way provide guidance in prioritizing corrective actions [14]. For example towards the development of an improved watermarking system .contest are usually given out to see if anybody can break a watermarking system.

D. Why Metrics Generation Is Difficult

Many in the security industry will agree that the number of successful security attacks an organization has experienced is not necessarily an indication of how secure that organization is. Security manager needs to look beyond the organization's security incident record for indicators of security strength. There are further complications they need to keep in mind, however, in their search for meaningful metrics [14].

E. Secure Watermark and Robust Watermark

Although a great amount of the watermarking and data-hiding literature deals with the problem of robustness, little has been said about security, and even in this time of relative maturity of watermarking research no consensus has been reached about its definition, and robustness and security continue to be often seen as overlapping concepts.[15]. The purpose of next section is to give an overview of the evolution of research on watermarking robustness and security.[16] make a distinction between robust and secure watermarks. A secure watermark is robust, but also secure in a sense that it can resist intentional tampering. A robust watermark, in its turn, is designed to be resilient. The watermark must withstand attacks that are aimed directly at the embedded information. It must not be possible for an attacker to delete the watermark without rendering the multimedia data unusable. Especially it must not be possible to retrieve or even modify the watermark without knowledge of the secret watermark key[17].

II. CONCEPTS OF ROBUSTNESS

Watermarking robustness deals with the performance of a watermarking scheme against common processing operations for example re-compression, transcoding, editing operations and so on.[18].

Robustness attacks

Pirates can creates quasi-copies by applying some standard modifications such as rotation, shearing, resampling, resizing, cropping, filtering, loss compression .and their combination. These modifications can be found in many image processing softwares such as Adobe Photoshop, Microsoft Paint. Another tool to create quasi-copy is Stirmark .Originally, StirMark is a generic tool for benchmarking the robustness of still image watermarking algorithms. Basically, StirMark simulates a resampling process, i.e. it introduces the same kind of distortions than one would expect if we print an image on a high-quality printer and then scan the document with a high-quality scanner[2].
The StirMark Benchmark for Audio (SMBA) tool contains currently a set of 39 single geometric attacks. The motivation for all attacks in SMBA is to destroy or weaken the embedded watermark signal. Audio attacks examples include compression and Digital- to- Analogue (D/A) and Analogue-to-Digital (A/D), temporal filtering, recording on audio tapes among others.

Robustness Metrics

The challenge in an audio watermarking scheme is the trade-off between robustness and imperceptibility. To assess robustness, we used the normalized correlation (NC) and Bite Error rate (BER) of the original and extracted watermark data, i.e. binary images.

1. Normalized Correlation

The robustness of the scheme is measured by calculating the similarities between the extracted watermark binary image, w′ and original watermark binary image, w. NC can be defined as:

$$NC(w,w') = \frac{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} w(i,j)\, w'(i,j)}{\sqrt{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} w^2(i,j)} * \sqrt{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1} w'^2(i,j)}}$$

Where m and n are the number of rows and columns of the binary image. The value of NC is between 0 and 1. A higher NC indicates more similarity between w and w′.

2. Bit Error Rate

To find the percentage of errors between the embedded and extracted watermark, bit error rate is used. BER is given by

$$BER = \frac{100 * \sum_{n=1}^{K} \begin{cases} 0, & w(n) = w'(n) \\ 1, & w(n) \neq w'(n) \end{cases}}{K}$$

where, w and w ' are the binary data formed from embedded and extracted watermark images, respectivel.[19].

Practical Evaluation of robustness against attacks

1. Robustness without Attack

When no process is applied, scheme compare watermark image to the original binary image.

2. Robustness against Low-pass Filtering Attacks

A low-pass filter with cut off frequencies of 4 KHz, 8 KHz and 11.3 KHz will be applied to the watermarked audio signals

3. Robustness against AWGN

Here white Gaussian noise is added to the watermarked audio signals so the resulting signals.

4. Robustness against MP3 Compression

MPEG-1 layer 3 (MP3) compressions is the most popular compression technique for audio signals. MP3 compression is applied to the WAV format of the watermarked audio signals at a rate of 64 kbps, 128 kbps and 256 kbps. Finally, the MP3 format is converted back to WAV format.

5. Robustness against Re-quantization

Each 16-bit watermarked audio signal was re-quantized to 8-bit and also up to 32-bit, and then again converted to 16-bit.

6. Robustness against Resampling

Watermarked audio originally sampled at 44,100 Hz is resampled to 22050/3200/9600 Hz, then restored by resampling again to 44,100 Hz.

7. Robustness against Amplitude Modification Attacks

Various amplitude modifications (+3 dB boost, -3 dB cut, +10 dB boost, -10 dB cut, +5% increases in amplitude and +10% increases in amplitude) on audio signals are done[19].

[9] employed Adobe Audition CS 6 to apply attacks on the watermarked audio. MATLAB was also used to apply amplitude modification of 5% and 10%, and additive white Gaussian noise (AWGN). StirMark for Audio (SMBA) is also used to test robustness to attack of watermarked audio files. StirMark also has the version for watermarked digital image.

III. CONCEPTS OF SECURITY

Security may be defined as the ability to resist hostile attacks.Security is always the major issue while transferring data over the network. Watermarking technique provides a secure data transmission over the network. The security is quantified in an information-theoretic sense by means of the information leakage between the watermarked signals seen by the attacker and the secret key used in the embedding process [21]**.** High security and capacity of Watermark data are very important issue to be tackled in a watermarking system [20].

The main target of the security analysis is to determine whether the watermarking scheme conceals properly the secret key. The security level of the scheme is defined as the number of observations needed to achieve an estimate of the secret key up to certain accuracy. The security level is measured by means of the mutual information (a.k.a. information leakage) between the watermarked signals and the secret key, which is related to the variance of the key estimation error.

Some common security attacks are given below:

- Unauthorised removal: Eliminating attacks, Masking attack, Collusion attacks
- Unauthorised embedding: Embedding forged watermarks into works that should not contain watermarks for example fragile watermarks for authentication.
- Unauthorised detection: Unauthorised reading[24] .

General model for evaluating security

The purpose of this first section is to give an overview of the evolution of research on watermarking robustness and security.

First, the notation and a general model for the evaluation of watermarking security will be introduced. The model is depicted in Figures 2-a and 2-b: a message **M** will be embedded in an original document **X** (the host), yielding a watermarked vector **Y**. The embedding stage is parameterized by the embedding key Θ_e, and the resulting

watermark is **W**. In the detection/decoding stage, the detection key Θ **d** is needed; $\hat{M}$ denotes the estimated message in the case of decoding, and the decision whether the received signal is watermarked or not in the case of detection. Capital letters denote random variables, and bold letters denote vectors[21].

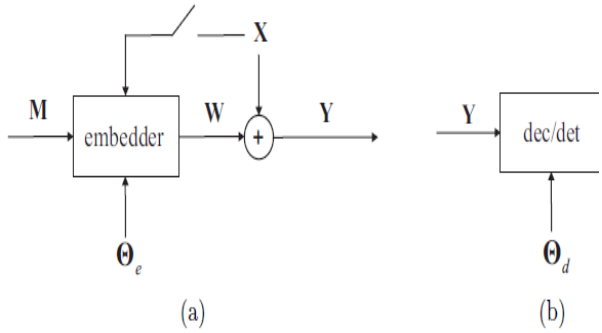


Figure. 2: General model for security analysis: embedding (a) and decoding/detection (b) Source[21].

Many literatures at the initial period deal with the issue of robustness and security with the notion of intentional /non intentional attacks. [7] shed some light on the concept of security in watermarking, giving definitions for robust watermarking and security, but perhaps they have the problem of being too general.

Another framework for watermarking security was proposed modelling watermarking as a game with some rules that determine which information (parameters of the algorithm, the algorithm itself, etc.) is public. According to these rules, attacks are classified as fair (the attacker only exploits publicly available information) or unfair (the attacker tries to access all the possible information which can be of help for him/her). The authors also define the security level as "the amount of observation, the complexity, the amount of time, or the work that the attacker needs to gather in order to hack a system"[21]

One of the most pioneer paper dealing with security by Cayre, Fontaine & Furon(2005) even had the difficulty of distinguishing between security and robustness. Kerckhoff's principle is also translated from cryptography to watermarking (it was introduced for the first time in [22]: all functions (encoding/embedding, decoding/detection) should be declared as public except for a parameter called the secret key. An important contribution of [23] is the proposal of a security measure based on Fisher's Information Matrix .

Fundamental definitions

[21] proposes some concept of watermarking security and some definitions are proposed in order to establish a clear line between robustness and security.

Definition 1. *Attacks to robustness are those whose target is to increase the probability of error of the data-hiding channel.*

Definition 2. *Attacks to security are those aimed at gaining knowledge about the secrets of the system (e.g. the embedding and/or detection keys).*

About the distinction between security and robustness : a watermarking scheme can be extremely secure, in the sense that it is (almost) impossible for an attacker to estimate the secret key(s), but this does not necessarily affect the robustness of the system. Therefore, higher security does not imply higher robustness. About the measure of security itself: security must be measured separately from robustness.

*Security attacks***.**

Security of a watermarking scheme is addressed whenever an adversary is part of the game and tries to remove the watermark[25]. There are two important families of security attacks:

Sensitivity attacks aim at removing the watermark by using the watermark detector as an oracle [15].

Information leakage attacks aim at estimating the secret key analysing contents watermarked with the same key[23]

If the algorithm is known, two neighbouring concepts are defined by the digital watermarking community:

- The **worst case attack,** which is the most efficient attack when the pirate knows the algorithm but not the secret key. Efficiency is typically measured by the loss of quality versus the increase in probability of watermarking decoding errors.
- The **security hack,** which is different in the sense that the pirate first takes time to analyze observations in order to estimate the missing information such as the secret key, and then leads an attack based on this stolen knowledge [2].

Oracle attacks

In cryptanalysis and digital watermarking security assessments, there is a class of attack named oracle attacks. The pirate has an unlimited access to a piece of software that is part of the system under scrutiny - this piece is said to be an oracle. It may even be a sealed black box process. What matters is the access to the output and the total freedom to run that software on any arbitrary, yet well chosen, input.

Kerckhoffs' Principle

Generally speaking, there are two broad approaches to keep things secret:

The first approach maintains security of the system through obscurity.

The second approach maintains security through the use of cryptographic-based techniques defining a secret key.

1.Security through obscurity relies on the assumption that pirates may have very hard time to precisely know what are the algorithms used in the system, what the implementation details are, what are the parameters and what can be their values. In other words, the security through obscurity assumes the pirates are unlikely to find the security flaws due to the great complexity of the system they are attacking. It has been demonstrated that solely relying on obscurity is not reliable. Pirates look for any piece of information in publications, patents, standards or by social engineering. It is impossible to empirically assess how difficult it is to disclose information about a system. The second drawback of security through obscurity is the high cost of changing the algorithm if it gets disclosed. Designers need to re-implement another obscure algorithm, likely way different, with heavy testing phases and a high burden for deploying the algorithm on sites. These problems have been reported for a long time and Kerckhoffs came up with several design principles still applicable today [2].

2. Secret Key

A secret key is a piece of information that determines the output of a particular algorithm: the same algorithm does not produce the same output if it is fed with the same data but with different keys. This secret key solely guarantees the security of the system. As it is much easier to protect a small piece of information (the secret key) than a complete system by obscurity, secure systems using secret keys are more reliable. A secret key is one or more very large random number.

Finding its value by an exhaustive search is almost impossible as the key space is very large. If the secret key is discovered, then creating a new key is easy and fast. Real-world secure systems typically include • secret keys and/or elements of obscurity [2].

IV. SECURITY ANALYSIS

Security of watermarking schemes can be assessed either with a theoretical analysis or with a practical evaluation. Theoretical security analysis consists in calculating the information leakage occurring when observing several watermarked contents by means of information theoretic measures such as equivocation or mutual information between the secret key and the observations (Bas & Doerr, 2007).

These measurements prove whether or not there is some information leakage that might be exploited to estimate the secret key. However, they do not give any clue about the tools that could be used to perform this estimation. The security level *s* can be defined as the number of watermarked contents necessary to perform a successful Worst Case Attack(WCA) [25].

Many of Watermarking techniques claim to be secure just because a secret key is used at the encoder. In reality, they present many weaknesses and the watermarking protection can be broken. The security of a watermarking concerns its capability to resist intentional forgeries.[23] proposed the foundations of the watermarking security measure. Based on Shannon theory on cryptographic security presented in the famous paper "Communication Theory of Secrecy Systems" ,they evaluate the security of watermarking system by measuring the information leakages about the secret key, used to prevent the unauthorized users reading the embedded message, from the observations available to the attacker.

[23] provide watermarking security against estimation attacks. The cryptographic key must remain secret in such a way that no one without authority should be able to break the secrecy. The watermarking security is the ability to make the secrecy recovering difficult for the attacker. [23] assess this difficulties by the size of the observations available to the attackers. They gave theoretical tools from the information theory in order to evaluate the security level. The two most important measures of the security level are summarised below:

1. Measure of security proposed based on Shannon's approach

The measure of security proposed here is a direct translation of Shannon's approach to the case of continuous random variables, which was already hinted for watermarking by [26]. Furthermore, we will take into account Kerckhoff's principle namely that the secrecy of a system must depend only on the secret keys. Security can be evaluated in the two scenarios of Figure 2.

For the scenario depicted in Figure 2a, security is measured by the mutual information between the observations Y and the secret key Θ

$$I(Y^1,Y^2,\ldots,Y^{No};\Theta) = h(Y^1,Y^2,\ldots,Y^{No}) - h(Y^1,Y^2,\ldots,Y^{No}\mid\Theta) = h(\Theta) - h(\theta|Y^1,Y^2,\ldots,Y^{No}), \quad (1)$$

Where, $h(\cdot)$ stands for differential entropy, and Y^n denotes the n-th observation.

Equivocation is defined as the remaining uncertainty about the key after the observations:

$$h(\Theta|Y^1,Y^2,\ldots,Y^{No}) = h(\theta) - I(Y^1,Y^2,\ldots,Y^{No};\Theta). \quad (2)$$

This scenario encompasses attacks concerning the observation of watermarked signals, where it is possible that additional parameters like the embedded message **M** or the host **X** are also known by the attacker. The model is valid for both side-informed and non-side-informed watermarking/data-hiding schemes [15].

- The scenario depicted in Figure 2b covers the so-called oracle attacks. In this case, the attacker tries to gain knowledge about the secret key Θ by observing the outputs $\hat{M}$ of the detector/decoder corresponding to some selected inputs Y, so the information leakage is measured by

$$I(\hat{M}^1,\ldots,\hat{M}^{No},Y^1,\ldots,Y^{No};\Theta), \quad (3)$$

where, in this case, Y^n are not necessarily watermarked objects but any arbitrary signal, for instance the result of the iterations of an attacking algorithm.

The translation of Shannon's approach to the continuous case is straightforward; we only must be careful with the concept of differential entropies, in order to redefine properly the unicity distance for continuous random variables:

in this case, an attacker will have perfect knowledge of the key when

$$h(\Theta|Y^1,Y^2,\ldots,YNo) = -\infty. \quad (4)$$

Hence, the security level is the number No of observations required to reach the unicity distance. However, since this number is ∞ in general, the security level could be measured by the growth-rate of mutual information with the number of observations No. Another possibility is the establishment of a threshold in the value of the equivocation, which is directly related to the minimum error variance in the estimation of the key:

$$\sigma_E^2 \geq \frac{1}{2\pi e} e^{2h(\Theta|Y)}$$

Where, σ_E^2 is the estimation error variance. For an attack based on the key estimate, its probability of success is given by the variance of the estimation error[15].

However, it is impossible to use the Shannons' measure with a continuous parameter since the entropy does not still a measure of information [27]

Fisher's measure

[23] proposed to use the fisher information in order to measure the security of the watermarking systems instead of the equivocation. It replaces the Shannon measure when the secret parameter is continuous and the joint probability density function, between the available information O and the secret parameter K, is derivable. In statistics ,Fisher was one of the first to introduce the measure of the amount of information supplied by the observations about an unknown parameter to be estimated. Suppose observation O is a random variable with a probability distribution function depending on a parameter vector θ.

The *Fisher Information Matrix* (FIM) concerning ɵ is defined as

$\mathrm{FIM}(\theta) = E\varphi\varphi^{\mathbf{T}}$ with $\varphi = \nabla \log pO(o;\theta)$

where E is the mathematical expectation operator and ∇θ is the gradient vector operator defined by

$$\nabla\theta = (\partial/\partial\theta[1]),\ldots,(\partial/\partial\theta[N\theta])^{T}$$

The Cramer-Rao theorem gives a lower bound of the covariance matrix of an unbiaised estimator of parameter vector θ whenever the FIM is invertible:

$$\boldsymbol{R\theta} \geq \mathrm{FIM}(\theta)^{-1}$$

in the sense of non-negative definiteness of the difference matrix. In this framework, the parameter vector can be the watermark signal or the secret key. Rθ provides a physical interpretation i.e the bigger the information leakage is, the more accurate the estimation of the secret parameter is [23]

Theoretical evaluation of security

Theoretical measures about the residual entropy will be presented later. The notation Nv will denote the length of the vectors (number of samples in each observation), No the number of observations, and Nc the number of carriers (or hidden symbols). After some modifications in the nomenclature the following attacks will be analyzed under: Known Message Attack (KMA), Watermarked Only Attack (WOA) and Original Attack (KOA).

Security level is assessed for a given context given below:

• the **Watermarked Only Attack (WOA),** in which the opponent only has access to N_o watermarked vectors y^{No} ;

• the **Known Message Attack (KMA),** in which the opponent only has access to N_o watermarked vectors and the associate messages $(\mathbf{y},\mathbf{m})^{No}$;

• the **Known Original Attack (KOA),** in which the opponent only has access to No watermarked vectors and the corresponding original ones $(\mathbf{y}, \mathbf{x})^{No}$.

V. Conclusion

To formulate the security problems and to design techniques to tackle them is a tricky business in digital watermarking. Many previous approaches are

more or less heuristic, which often lead to schemes that are later proved insecure. Hence, we are going to need more rigorous approaches to better assess the security. There are currently two main categories of approaches. One is application and attack specific, but easier to be applied. The other is more high level and theoretical, but allows more general results to be proved. To achieve security in watermarking in general, we need a blend of these two approaches, where the results are rigorous and general, yet allow detailed analysis of security in practice. Many authors consider that for normal signal processing operations we talk of robustness and when it become intention we talk about security. Also higher security does not translate to higher robustness and belies that both quantities should be separately measured. Future works will be on the investigation of further relationships between the two watermarking system properties.

REFERENCES

[1] Agbaje, M.O (2008): Effect of Blocksizes on watermarked digital images using discrete cosine transform (DCT).M.Sc Thesis, FUNAAB, Nigeria.

[2] DO Thanh-Toan (2012),Phd thesis Security analysis of image copy detection systems based on SIFT descriptors. UNIVERSIT´E DE RENNES.

[3] Yadav Anupma, Yadav Anju (2014) (2014), Comparison of SVD-Watermarking and LSB-Watermarking Techniques, *IJCSMC, Vol. 3, Issue. 5, May 2014, pg.495 – 499*, ISSN 2320–088X.

[4] Li Qiming and Memon Nasir (2007),Security Models of Digital Watermarking MCAM, LNCS 4577, pp. 60–64, 2007. Springer-Verlag Berlin Heidelberg

[5] Fazli, Amir reza Asli Mohammad Eghbali and Sarbisheie Ghazale (2012), Security Evaluation of DWT Based Watermarking Schemes and its Improvement Using RS Coding, Research journal of Applied Sciences, Engineering and Technology 4(17): 3031-3039, 2012 ISSN: 2040-7467, Maxwell Scientific Organization, 2012.

[6] Dymarski P. and Markiewicz R.(2014), Robust Audio Watermarks in Frequency Domain, Journal of Telecommunications and information Technology.

[7] Kalker Ton 2001), "Considerations on watermarking security," in *Proc. MMSP*, Cannes, France, pp. 201–206.

[8] Hartung & Kutter, 1999 Hartung, F., & Kutter, M. (1999). Multimedia watermarking techniques. Proceedings of the IEEE, 87(7), 1079–1107.

[9] Katzenbeisser, S., & Petitcolas, F.A.P. (2000). Information hiding techniques for steganography and digital watermarking. Boston: Artech House

[10] Althos(2009),IPTV Industry Dictionary.

[11] Wu Yongdong, Bao Feng and Xu Changsheng,(2003),The Security Flaws in some authentication, Watermarking Schemes, IEEE II ICME ,pg493-496

[12] Harjito, Bambang. 2013. Copyright protection of scalar and multimedia sensor network data using digital watermarking. Ph.D. Curtin University, School of Information Systems, Curtin Business School

[13] Unoki, Kuniaki Imabeppu, Daiki Hamada, Atsushi Haniu, and Ryota Miyauchi(2011), Embedding Limitations with Digital-audio Watermarking Method Based on Cochlear Delay Characteristics, Journal of Information Hiding and Multimedia Signal Processing ©2011 ISSN 2073-4212Ubiquitous International Volume 2, Number 1, January 2011

[14] Payne Shirley C. (2006), A Guide to Security Metrics, SANS Security Essentials GSEC Practical Assignment Version 1.2e.

[15] Comesa˜na P., P´erez Freire L., and P´erez-Gonz´alez F. (2006).Blind Newton Sensitivity Attack. IEE Proceedings on Information Security, 153(3):115–125.

[16] Cox, I.J., Miller, M.L., & Bloom, J.A. (2002). *Digital watermarking*. San Francisco: Morgan Kaufmann.

[17] Mitra Abbasfard(2009) Digital Image Watermarking Robustness: A Comparative Study

[18] Bas Patrick, Westfeld Andreas (2009) **:**Two Key Estimation Techniques for the Broken-Arrows Watermarking Scheme.

[19] Chowdhury Alok Kumar and Khan Md. Ibrahim (2013), A Tutorial for Audio Watermarking in the Cepstrum Domain Smart Computing Review, vol. 3, no. 5, DOI: 10.6029/smartcr.2013.05.003

[20] Verma Monika, Praveen (2013)International Journal of Advanced Research in Computer and Communication Engineering,Vol. 2, Issue 7, July 2013, pg 2913-2917,Capacity and Security analysis of watermark image truly imperceptible, ISSN (Print) : 2319-5940,ISSN (Online) : 2278-1021.

[21] P´erez-Freire Luis and P´erez-Gonz´alez Fernando (2007), Exploiting Security Holes in Lattice Data Hiding , Signal Theory and Communications Department.

[22] Furon, T., et al(2000).: Security Analysis. European Project IST-1999-10987 CERTI-MARK, Deliverable D.5.5.

[23] Cayre Fracois, Caroline Fontaine, and Teddy Furon(2005), Watermarking Security: Theory and Practice,IEEE Transactions on signal processing.

[24] Cheung S.(2004),Multimedia Information System,EE 639,Lecture 23,Digital watermarking I.

[25] Bas Patrick and Doerr Gwenael (2007),Practical Security Analysis of Dirty Paper Trellis Watermarking

[26] Hern´andez, J.R., P´erez-Gonz´alez, F(1998).: Throwing more light on image watermarks.In Aucsmith, D., ed.: 2nd Int. Workshop on Information Hiding, IH'98. Volume 1525 of Lecture Notes in Computer Science., Portland, OR, USA, Springer Verlag, 191–207

[27] Braci S., Boyer R., Delpha C. (2009), "Security evaluation of informed watermarking schemes", 16th IEEE International Conference on Image Processing (ICIP), Nov. 2009, pp. 117 - 120