IT4InDev 2015





## Information Technology for Inclusive Development

Conference Proceedings Volume 26

### **Edited by:**

Professor Charles UWADIA Professor Adesola ADEROUNMU Dr. Adesina SODIYA

ISSN: 2141-9663



# 12<sup>th</sup>

### International Conference

#### ACKNOWLEDGEMENT

It is with great pleasure that the Nigeria Computer Society (NCS) acknowledges the immense and revered support received from the following organisaions:

- CBC Networks Limited
- Central Security Clearing Systems (CSCS)
- Chams Plc
- Computer Professionals (Registration Council of Nigeria) (CPN)
- Computer Warehouse Group
- Data Sciences Nigeria limited
- MainOne Cable Limited
- Management Edge

- National Information Technology Development Agency (NITDA)
- Phoenix Ditribution
- Power Soft Limited
- Resort Savings & Loans
- Sidmach Technologies Limited
- University of Lagos
- Zinox Technologies Limited
- Zippro Systems

We also recognize the laudable efforts of all others in cash or in kind towards the success of the 12<sup>th</sup> International Conference

#### **REVIEWERS:**

Prof. G. A. ADEROUNMU Dr. A.S. SODIYA Prof. J. O. A. AYENI Prof. L. O. KEHINDE Prof. G. M. M. OBI Prof. 'Dele OLUWADE Prof. Vincent E. ASOR Prof. O. S. ADEWALE Prof. Peter K. ORIOGUN Dr. D. K. ALESE Dr. ' Sesan ADEYEMO Dr. (Mrs.) S.A. ONASHOGA Dr. A. P. ADEWOLE Dr. A. O. AJAYI Dr. S.E ADEWUMI Dr. I. K. OYEYINKA Dr. S. O. OLABIYISI

Dr. O. B. AJAYI Dr. O. FOLORUNSO Dr. (Mrs.) O.S. ONAOLAPO Dr. A. I. OLUWARANTI Dr. A. T. AKINWALE Dr. E. O. OLAJUBU Dr. O. A. OJESANMI Dr. E. ESSIEN Dr. E. G. ESEYIN Dr. (Mrs.) O. R. VINCENT Dr. Olumide B. LONGE Dr. A. A. O. OBINIYI Dr. F. T. IBRAHALU Dr. O. T. AROGUNDADE Dr. O. R. VINCENT Dr. S. A. AREKETE Dr. O. J. OYELADE

#### **Publication Office:**

Nigeria Computer Society (NCS): Plot 10, Otunba Jobi Fele Way, Central Business District, Behind MKO Abiola Garden, Alausa, Ikeja – Lagos, Nigeria P.M.B. 4800 Surulere, Lagos, Nigeria. Phone: 01-7744600, 08097744600 E-mail: publication@ncs.org.ng, ncs@ncs.org.ng Website: www.ncs.org.ng

© All rights reserved. No part of this publication may be reproduced in whole or in part, in any form or by any means, electronically or mechanically without the written permission of the **Nigeria Computer Society (NCS)**.



## 12<sup>th</sup> Intern

### **International Conference**

#### Forward

It is my great pleasure and delight to welcome all of us to the 12<sup>th</sup> International Conference of the Nigeria Computer Society which is holding at the beautiful city of Akure, Ondo State, Nigeria from July 22 to-24, 2015. The theme of this year's conference is "Information Technology for Inclusive Development". This year's conference is intended to provide a forum for policy makers, public and private sector, IT practitioners and academia to keep up to date with technology trends and to discover opportunities for the private sector in the process of inclusive development and avenues for public-private partnerships towards realizing the IT objectives within the context of Millennium Development Goals (MDG). The Conference will provides opportunities for the delegates to exchange new ideas, establish business or research relations, and find global partners for future collaborations.

In this year conference, professionals from government circle, industry, research institutes and academia have submitted insightful papers in E-Business, Broadband & Information Security, Data Analytics, Local Content and Entrepreneurship Development. In addition to the above, there are three special sessions: Social Media Master-Class, ICT Policy for States in Nigeria, and Keynote Session. This year Research Consortium on Information Technology Innovations (ReCITI) will focus more on Big Data Analytics Using Hadoop (theory and practical).

Going by the quality of the papers and the personalities presenting lead papers on well researched and challenging issues, I am persuaded that an extremely rich cross-fertilization of ideas of experts from across the globe is guaranteed.

The organizers of the conference owe special thanks to His Excellency, the Governor of Ondo State, and Dr. Olusegun Mimiko for his support towards the success of this conference. **The** Permanent Secretary, Federal Minister of Communications Technology, Dr Tunji Olaopa has also been pillar of support in putting up this conference. We are indeed very grateful to the Vice-Chancellor, Federal University of Technology, Akure, Professor Adebiyi Daramola for the total support we received from the Federal University of Technology, Akure.

I would like to appreciate our national and international keynote/lead paper presenters including Dr. Anu Gokhale (USA), Dr. Iyilade Johnson (Canada), Mr Austin Okere (Nigeria), Mr Denzil Kentebe (Nigeria), Mr Martin Eigbike (Nigeria), Mr Tope Aladenusi (Nigeria), and Mr. Biyi Dosunmu (Nigeria)

Let me also use this opportunity to thank all our sponsors including CBC Networks Limited, Computer Warehouse Group, Data Sciences Nigeria limited, MainOne Cable Limited, National Information Technology Development Agency (NITDA), Sidmach Technologies Limited and Zinox Technologies Limited who have generally sponsored this year's conference. It is my prayer that together, we will move Nigeria to a greater height. I wish you all, very exciting and resourceful deliberations and Journey mercies back to your destinations at the close of this conference.

The Chairman, Local Organizing Committee, Mr Ebezer Adejuyigbe and all members of his committee have taken the organization of this Conference as task that must be accomplished. The National Executive Council of the Nigeria Computer Society is immensely indebted to the dynamic Conference Planning Committee, which worked assiduously, even against odds, to make today a reality.

Thank you and God bless you all.



#### TABLE OF CONTENTES

	1.	Acknowledgement	I
	2.	Forward	П
	3.	Table of Contents	111
Se	ssio	nl	
50	1.	Harnessing FPGA Processor Cores in Evolving Cloud Based Datacenter Network Design – K.C. Okafor, G.N. Ezeh, I.E. Achumba, F.N. Ugwoke, C.C Okezie and U.H.Diala	1
	2.	Reusable Components in the Political Domain - Computational Algorithms for Enhancing Participatory Social Accountability – E. Okewu and J. Okewu	15
	3.	Introducing a Model to Improve Recent Sub-Saharan Africa Networked Readiness Index – P. K. Oriogun, K. K. Agbele, K. D. Aruleba and A. O. Agho	31
	4.	<b>A New Vector Space Model for Grammaticality Grading in Handcrafted Grammar</b> – M. K. Aregbesola, R. A. Ganiyu , S. O. Olabiyisi, E. O. Omidiora and O. O. Alo	39
	5.	Privacy- Enhancing Big Data Computing in Financial Institutions – C. Oloruntoba and A.S. Sodiya	53
So	scio		
Je	1.	Facilitating SMEs for Inclusive Development: The Role of Service-Oriented Software Engineering (SOSE) – E. Okewu	63
	2.	Smart City: Concept, Applications and Architectural – N.F. Efozia, T. Immanuel, F. Odebunmi, O. Giwa and T.Balogun	74
	3.	A User-Centric Design Towards Performance Development for Adaptive e-Learning Systems – K. Okoye	87
	4.	Secure Automated Teller Machine (ATM) Using Fingerprint Authentication and Short-Code Message in A cashless Society – J. B. Awotunde, R. G. Jimoh and O. E. Matiluko	99
	5.	Object Classification in Videos Using Neural Network: A Case Study of Vehicles and Pedestrians Classification for Surveillance systems – G. Akintola, O.C. Akinyokun, O. Olabode and G.B. Iwasokun	112
Se	ssio	n III	
	1.	Experimental Formulation of a Vector Space for Grammaticality Grading in Handcrafted Grammar – M. K. Aregbesola, R. A. Ganiyu, S. O. Olabiyisi, E. O. Omidiora, and O. O. Alo	121
	2.	A Modified Expectation Confirmation Theory with Perceived Trust on Internet Cafes: A Conceptual Framework – A. Lawan, B. S. Galadanci and A. A. Abdallah	134
	3.	<b>Performance Evaluation of Apache Web Server on Cloud Computing Environments</b> – S. Lawal and M. Aminu	141



4.	An Investigation on the Performance of some Nigerian Mobile Broadband Network Services – S. Sirajou, T. A.H. Tambuwal and A. Mohammed	147
5.	<b>Evaluation of Full Text Search Retrieval System</b> – K.D. Aruleba, R.D. Aremu, P.K. Oriogun, K.K. Agbele and A.O. Agho	154
Sessio	n IV	
1.	An Enhanced Hybrid Algorithm for Cellular Data Offloading – A. A. Bichi and S. M. Zahid	160
2.	Multi-factor Authentication for Social Media – O. O. Ayannuga, O.N. Lawal, O. Oladipo and W.A. Salau	168
3.	A Cryptographic Anti-Phishing Scheme – A. A. Orunsolu, A.S Sodiya, I.M Fadeyi and S.A Onashoga	176
4.	Design and Development of a Multi-Modal Biometric System with De-Duplication Functionality – C.C. Nwannenna	186
5.	A Smart Home SIM300 GSM Module Controller - D.N. Nnakwuzie, C.I. Akobundu and M.U. Onuh	193
Sessio	n V:	
1.	Securing Data in the Cloud Using RSA Homomorphic Encryption – O. B. Alaba, W. O. Adesanya, A. S. Ogunbanwo and O.O. Olusanya	204
2.	Towards an adaptive and scalable access control modelfor a Cloud-Based environment - N.A Azeez; O.B Okunoye, F.AOladeji, E.O. Edafeadjeke	214
3.	<b>Prediction of Peadiatric Hiv/Aids Survival n Nigeria Using Naive Baye's Approach</b> – O. Agbelusi, T. A. Aladekomo and P. A. Idowu	224



### HARNESSING FPGA PROCESSOR CORES IN EVOLVING CLOUD BASED DATACENTER NETWORK DESIGNS, (DCCN)

#### K.C.Okafor<sup>1</sup>; G.N. Ezeh<sup>1</sup>; I.E. Achumba<sup>1</sup>; F.N. Ugwoke<sup>2</sup>; C.C Okezie<sup>3</sup> and U.H.Diala<sup>4</sup>

<sup>1</sup>Dept. of Electrical Electronic Engineering, Federal University of Technology, Owerri, Nigeria.
 <sup>2</sup>Dept. of Computer Science, Michael Okpara University of Agriculture, Umudike, Nigeria.
 <sup>3</sup>Dept. of Electronic and Computer Engineering, NnamdiAzikiwe University, Awka, Nigeria
 <sup>4</sup>Dept. of Automatic Control and Systems Engineering, University of Sheffield, UK

<sup>1</sup>ugoezeh2002@yahoo.com,<sup>2</sup>kennedy.okafor@futo.edu.ng,<sup>3</sup>ifeyinwaeucharia@yahoo.com<sup>4</sup> ndidi.ugwoke@gmail.com,<sup>5</sup>christianaobioma@yahoo.com, <sup>6</sup>uhdiala1@sheffield.ac.uk,

#### ABSTRACT

Contemporarily, the processors in backend server domain havemultiple cores, but the bandwidth to external memory and other interfaces is not keeping pace with the increase in computing power and the required Quality of Service (QoS). Many of these servers are running at average utilization rates and perform well under peak processing power. These servers are good candidates for Field Programmable gate Array (FPGA) processors in cloud based datacenters. In this research, hardware acceleration through FPGAs is shown to be an attractive alternative to replacing traditional processors. This paper, explains whycurrent IT network infrastructure, processing bandwidth of DCNs must leverage the FPGA capabilities to support critical business processes. The paper analysed FPGAs, characterised a simplified process model for behavioural networkdescription and demonstrated a typical cloud based DCN for evaluation purposes. From the performance evaluation, it was concluded that FPGA based platforms, *e.g.*, NetFPGA, or OpenFlow switches, etc. can effectively implement DCN designs including topology design, control plane and routing, and congestion control.

Keywords: Processor Cores, Cloud Datacenters, FPGA, OpenFlow Switches, Networking, Infrastructure

#### 1. INTRODUCTION

Datacenters have been built around the world for various cloud computing services,Lu, et al (2011). Servers in datacenters are interconnected using datacenter networks. A large datacenter network may connect hundreds of thousands of servers. Due to the rise of cloud computing, datacenter networking (DCN) is becoming an important area of research, Lu, et al (2011). Several contribution and studies on various aspects of DCN have been carried out. For instance,topology design and routing, Chuanxiong et al (2008),AI-Fares, et al (2008) ,Guo, et al (2009),Greenberg, et al (2011),Mysore, et al (2009), Okafor et al., (2014),flow scheduling and congestion control,Okafor and Nwaodo



(2012), Alizadeh et al (2008), Fares, et al (2010), virtualization Okafor, et al (2015), Guo, et al (2010), application support, Shieh, et al (2010),Abu-Libdeh, et al (2010), have been studied.

Since DCN is a relatively new exploration area in cloud based networks, many of the designsincluding Libdeh et al (2010)have departed from the traditional Ethernet/IP/TCP based packet processing. In cloud based scenario, the type of switch or even server processor cores can contribute to congestion delays. For example, Portland performs longest prefix matching (LPM) on destination MAC address, BCube advocates source routing, and Quantized Congestion Notification (QCN), Alizadeh et al (2008) uses rate based congestion control.

Hence, current Ethernet switches and IP routers therefore cannot be used to implement high performance datacenter designs. To implement these cloud DCN designs, rich programmability is absolutely required in the cloud DCN service processors. There are approaches that provide thisprogrammability: software-based, pure Kohler, et al (2000), Dobrescu, et al (2009) or FPGA-based systems (e.g., NetFPGA, Naous, et al (2008). Software-based systems can provide full programmability as verified byDobrescu, et al (2009), Han, et al (2010) and may provide a reasonable packet forwarding rate. But their forwarding rate is still not comparable to commodity switching FPGA Application Specific Integrated Circuit (ASICs). The batch processing used in existing server switches and Softwarebased switches vield optimization that introduces highlatency. This may be critical for various control plane functions such as signalling and congestion control.

Also, the packet forwarding procedure in DCN are generally simple and hence arebetter implemented in silicon semiconductor for cost and power savings.

Since,FPGA-based systems are fully programmable, Yang, et al.,(2013), one could virtualize the FPGA by reconfiguring it at runtime to support more functions than could fit into a single device.

There is need to explore FPGA service boxes in developing DCNs for cloud based services (such the Enterprise Energy Tracking Analytic Cloud Portal, EETACP, social networks, etc.).

The rest of this paper is organized as follows: Section II presented related works, while section III discussed the proposed FPGA-based network model. Section IV outlined the implementation via simulation. Section V shows the results and conclusion.

#### 2. RELATED WORKS

#### 2.1 Datacenter Networking

The datacenter architectures such as DCell, Chuanxiong G. et al (2008), BCube, C.Guo, et al (2009),Portland, Mysore, et al (2009),etc. have been studied byUdeze, et al (2014), andOkafor, (2015).According to Okafor, (2015), and Zhangxi, (2013), there are many practical issues scaling existing commercial off-the-shelf Ethernet switches, especially at a high link speed, such as 10 gigabit per second (Gbps). These practical issues are:

a. Current networks are extremely complex, particularly the switch software.

b. Existing networks have many different failure modes. Occasionally, correlated failures are found in replicated million-dollar units.

c. Existing large commercial switches and routers command high margins and charge a great deal for features that are rarely used in datacenter. Therefore, they are very difficult to scale out to larger configurations without complete redesign. d. Some datacenters require a large number of ports at aggregate or datacenter-level switches at extremely high link bandwidth. But such switches do not exist on the market currently, A. Ganesan, et al (2011). For instance, Google Gscale network is running on custom built 10Gbps switches with 128-ports, with plans for 40Gbps systems supporting as many as 1,000 ports.

Therefore, many researchers have proposed novel datacenter network architectures,Chuanxiong et al (2008),Greenberg, et al (2011), Guo, et al (2009),Mysore, et al (2009),Joseph, et al (2008),Tavakoli, et al (2009), Thacker (2007) with most of them focusing on new switch designs. There are also several new network products emphasizing low latency and simple switch designs.

Zhangxi, et al (2013) established from figure 1that datacenter network infrastructure accounts up to 18% monthly cost (3rd largest cost). In this



## 12<sup>th</sup>

### International Conference

case, large switches/routers are expensive and unreliable.



Figure 1: DCN Distribution, **(source:** James 2009) The work then proposed DIABLO cluster Prototype shown in figure2which shows an implementation of high-throughput datacenter using FPGAs. The DIABLO model abstractions and the FPGA implementations for three major datacenter components: server, switch, and network interface card were carried out.



### Figure 2: DIABLO cluster Prototype, Zhangxi (2013)

Andrew etal.,(2014) described a reconfigurable fabric (Catapult) designed to balance someperformance concerns. The Catapult fabric is embedded into each half-rack of 48 servers in the form of a small board with a medium-sized FPGA and local DRAM attached to each server.

As depicted in Fig 2. FPGAs are directly wired to each other in a 6x8 two-dimensional torus, allowing services to allocate groups of FPGAs to provide the necessary area to implement the desired functionality.The work evaluated the Catapult fabric by offloading a significant fraction of Microsoft Bing's ranking stack onto groups of eight FPGAs to support each instance of this service, Andrew et.al, (2014).

#### a. An FPGA-Based Multiprocessor Architecture

The new processor architectures based on a FPGA-device programmable have several advantages. It allows for scalability on demand and loosely coupled system designs, Joost, and Salomon (2009) showed that fieldprogrammable gate arrays (FPGAs) are suited to fulfil for most industrial applications. It was established that FPGAs are very powerful, relatively inexpensive, and adaptable, since their configuration is specified in an abstract hardware description language. FPGA-based systems combine many advantages of Digital Signal Processors (DSPs) and application specific integrated circuits (ASICs). This includes, rapid development cycles, high flexibility and reusability, moderate costs, easy upgrading (due to the usage of abstract Hardware Description Languages (HDLs)), and feature extension (as long as the FPGA is not exhausted).

Furthermore, current FPGAs allow for the integration of soft-core processors. An FPGA scalable architecture was proposed by Tenca, et al (2000)with relevant evaluations.Xilinx FPGA comparison showing optimal configuration for VirtexUltraScale device has been enumerated in http://www.xilinx.com/products/silicon-

devices/fpga.html. In context, the logic Cells (K), UltraRAM (Mb),Block RAM (Mb),DSP Slices, Transceiver Count, Maximum Transceiver Speed (Gb/s), Total Transceiver Bandwidth (full duplex) (Gb/s),Memory Interface (DDR3 ),Memory Interface (DDR4),PCI Express, Configuration AES, I/O Pins and I/O Voltages were all compared against other device architecture variants showing VirtexUltraScale device as the most preferred choice. This was adopted in the proposed DCN design described in section 3.

In DCN design context, FPGA based system implementations have the following characteristics:

 FPGAsdoes not implement complex combinatorial logic well, such wide multiplexers. Any logic is mapped with a number of multi-input Lookup Tables (LUTs).Actually, complex combinatorial logic occupies a large number of lookup tables in the device. Although modern FPGAs have plenty of logic resources,



- routing resources remain to be a critical component. The more logic levels the more routing delays. This observation will lead to simplified DCN service process micro-architectural design that will avoid complex combinatorial logic.
- FPGAs have a lot of RAM. This observation combined with the lack of bypass paths, led to a multithreaded design of large modules.

This work used the RAMs on FPGAs to store simulation thread state, and dynamically switch threads to keep simple module pipelines saturated. This strategy is referred as host-multithreading.

- Modern FPGAs have hard-wired DSP blocks. Execution units, especially Floating Point Units (FPUs), dominate Look Up Tables (LUT) resource consumption when implementing a processor on an FPGA. By mapping functional units to DSP blocks rather than just LUTs, more resources can be devoted to execution timing.
- DRAM accesses are relatively fast on FPGAs. Logic in FPGAs often runs slower than DRAM because of on-chip routing delays. This insight will greatly simplify host memory system, as large, associative caches are not needed for high performance. This observation will led to modelling large target buffers, such as packet-switch port buffers in DRAM, with minimum simulation performance impact.

Given that the proposed DCN design is a throughput optimized system with a feed through data path design, one could easily insert pipeline stages to boost host FPGA clock frequency. Longer pipeline stages can be filled with more hardware threads. In addition, it is possible to double-clock FPGA primitives such as DSP to have shorter access latency and more read/write ports on the primitives.

From literature studies, little work has been done on FPGA based DCNs. Also, most proposed DCN designs have only been tested with a very small test bed running unrealistic micro-benchmarks. Now, it is very difficult to evaluate such network architecture innovations at scale. DCNs that are not built on FPGA core will yield different observations but there is need for a sound simulation methodology to evaluate these DCNs based FPGA cores. The system model is discussed below.

#### 3. PROPOSEDFPGA-BASED NETWORKMODEL

#### 3.1 Mapping to FPGAs

The goal of the FPGA based server model is to have a credible workload generator that drives more detailed networking models. Essentially, building highly-accurate analytical models for the DCN workload is a less practical approach. Therefore, the server model must be capable of running complex server application software with minimum modifications.

In this proposal, an FPGA functional service model is responsible for executing the target procedure (router, switch or server CPU) correctly and maintaining the device architectural state. The benefits of this functional/timing split are:

- Simplified mapping of the functional FPGA model. The separation allows complex operations to take multiple host cycles. For example, a highly-ported register file can be mapped to a block RAM and accessed in multiple host cycles, avoiding a large, slow mapping to FPGA registers, multiplexers, etc.
- Improved modelling flexibility and reuse. The timing model can be changed without modifying the functional model, reducing modelling complexity and amortizing the functional model's design effort. For instance, it is possible to use the same switch functional model to simulate both 10Gbps switches and 100Gbps switches just by changing the timing model only.
- It enables highly-configurable а abstracted timing model. Splitting timing from function allows the timing model to be more abstract. For example, a timing model might only contain target cache metadata. Different cache sizes could then be simulated without resynthesis bv changing how the metadata RAMs are indexed and masked at runtime.



When looking closely at the FPGA characteristics for network architectures, this work identified a wide variety of design choices in almost every aspect of the design space, such as switch designs, network topology, protocols, and applications. See fig 7.

#### 3.2 FPGA Cloud Datacenter Design

#### Specifications.

The proposed FPGA datacenter is referred to as Distributed Cloud Computing Datacenter Design (DCCN).Considering the DCCN architecture discussed by Okafor, K.C. (2015), the network design will focus on the two layers functional areas, viz-a-viz Remote user access and Hybrid speed redundancy layer. The gateway load balancer (GLB)/ speed redundancy layer will be used interchangeably with the service processor or integrated service Openflow load balancer (ISOLB)in this work.

In the proposed design, the FPGA based ISOLB is the major component in the hybrid speed redundancy layer. This with the interconnected server subnet clusters are the major components of the hybrid speed redundancy layer. The DCCN port architectural model for the FPGA based ISOLB is shown in fig 4. This will be described to facilitate a clear understanding of the model specifications discussed below.

- Let  $DCCN_{lb}$  be an acronym chosen for the DCCN server cluster managed by the FPGA based ISOLB controller.  $DCCN_{lb}$  is designed to have four subnets for its clusters (subnet 1-4) referred to as  $DCCN_{sa}$ ,  $DCCN_{sb}$ ,  $DCCN_{sc}$ ,  $DCCN_{sd}$  interconnected as shown in Fig 3, where s is a subnet factor such that s > 0.
- Each cluster (DCCNs) uses High Performance Computing (HPC) servers with the ISOLB

controller layered in linearly defined architecture. Since our designing of datacenter network is for efficient server load balancing and application integration, this requires one (4-port) from FPGA based ISOLB controller and few servers, hence, the choice of four subnets. Virtual server instances running on the HPC servers made up for further need of hardware servers in the network.

- Cisco router (Cisco\_Cat6000 [Cs\_4000\_3s\_e6\_fr2\_sl2\_tr2] will be used as the user gateway interface into the FPGA based ISOLB.
- Servers in DCCN cluster are connected to FPGA based ISOLB port of the load balancer corresponding to it, and owing to the running virtual instances Vi, a commodity 4-port switching/routing device with 40GB/s per port serve the design purpose. Also, each of the DCCNsis interconnected to each other through the ISOLB switch ports.
- The virtualized server used in this work has two ports for redundancy (in Gigabytes). Each server is assigned a 2tuple  $[a_n, a_o]$  in consonance with its ports  $(a_n, a_o$  are the redundant factors) together with an OpenFlow VLAN id from(1 to 1005).
- Emulated NEC IP8800 OpenFlow controller was the FPGA based ISOLB used in this work, hence, the number 1005 is the maximum number of OpenFlow-VLAN that can be created in it. The load balancer switch is a multilayer commodity switch that has a load balancing capability. This capability together with its OpenFlow-VLAN capability was leveraged upon to improve the overall DCCN performance.





Figure 3: Proposed FPGA based Load Balancer Interfacing in DCCN Subnet Cluster

 Each server has its interface links in DCCN<sub>s.</sub> One connects to an FPGA based ISOLB, and other servers connect as well but all segmented within their subnets via OpenFlow-VLAN segmentation, as shown in figure 3. OpenFlow-DCN<sub>s</sub> servers have virtual instances running on it and are fully connected with every other virtual node in the architecture. The composite model, Okafor, (2015) was be used to expand the details of figure 3.

#### 3.3DCCN Composite Process Model

Figure4 shows the composite process model for resource allocation, scheduling and load balancing. From the model, the DCCN comprises of the remote users  $U_i$ , terminal gateways  $G_g$ , load managers (OpenFlow Controller), and the resource pool of virtualized processors, storages, etc. (DCCN clustered resources or resource pools).

In the design, while the core layer (resource pool)

addresses issues of resource control and processing center jobs, the load balancer OpenFlow effects services integration while offering fault tolerance.

Users  $U_i$ , gets services from internet via the gateways  $G_{gusing}$  their terminal devices  $T_{ds}$ . All jobs or tasks sent by the Users  $U_i$ , represents a request to network resources  $N_r$  in the DCCN.

The resource pools in the DCCN cluster server's processes all user requests via the Internet, but the DCCN OpenFlow load balance and the Vm processors are connected via high speed interconnection links  $H_{il.}$ 

The controller at the hybrid layer allocates the jobs/requests received from Users  $U_i$ , to the processors (DCCN<sub>sa</sub>, DCCN<sub>sb</sub>, DCCN<sub>sc</sub>, DCCN<sub>sd</sub>) they manage in their clusters (DCCN<sub>lb</sub>). The FPGA based Vm processors execute the jobs received from the load balancer and send it back to the Users  $U_i$ . Fig 5simplifies the Vm Server cluster integrations via the FPGA based OpenFlow controller.



**12<sup>th</sup>** 



Figure4: Composite process model for DCCN Figure5: Virtualization Design for DCCN Server Cluster

From figure 4, let the FPGA based OpenFlow Controller manager  $L_m$  (hybrid layer) and its server processors (resource pool) represent a cluster C. It literally depicts an M/M/1 Queuing System.In the model, there is an interplay between the access layer  $E_{U}$ , the Hybrid layer  $L_m$  and the server resource pool. Equation1 gives the composite model description.

$$\begin{array}{c} D \\ c \\ = \sum_{i=0}^{n} \langle E_{U1} + E_{U} + E_{U} \\ + \sum_{i=0}^{n+1} \langle L_{m_{1}} + L_{m_{2}} \\ + L_{m_{3}} \\ + L_{m_{1}} \\ + l \\ - l \\ + l \\ - l \\ + l \\ - l \\ + l \\$$

**Work-Work:** users With fight generation rates  $K_r$  represents users accessing DCCN resources via the front end application from the OpenFlow load manager  $L_m$  in the cloud domain C at a pay-as- you (PAGO) price  $P_{m}$ .

A user can connect to the computing resources of the application servers from the distributed load managers/controllers. The load manager  $L_m$  with job arrival rates  $R_j$  or  $\mu_j$  connects the users at the  $P_{R}$  and facilitates the  $V_{m_{n+1}}$  processors as soon as it receives user requests  $U_r$ . Each  $V_m$ processor only executes jobs allocated to it and never dispatches again to another processor's  $V_m$ . In this case, each  $V_m$  processor maintains a queue (M/M/1) that holds the jobs to be executed based on First-In-First-Out (FIFO) pattern and sends feedback to the user  $U_{i}$ . From figures 4 and 5, a derivation of the job allocation to OpenFlow load controller given by

$$\vartheta_{J_{\alpha}} = \sum_{i=1}^{n} L_{m} \beta_{U_{j}} = \sum_{i=1}^{n} \langle L_{m} \beta_{U_{3}} + L_{m} \beta_{U_{2}} + L_{m} \beta_{U_{3}} + \dots + L_{m} \beta_{U_{n+1}} \rangle$$
(2)
(2)
(2)

 $L_m \beta_{U_{n+1}}$  represents a percentage of the workload (jobs) that user  $U_i$  sends to OpenFlow load controller  $L_m$ .

For an effective model, the average job arrival rate  $R_j$  of the  $L_m$  should be less than the total average processing rate of the cluster  $L_m$  resource pool.

Hence, the Job arrival rate is given by

 $\begin{aligned} R_{L_m} &= \sum_{i=1}^{\mu} (L_m \beta_{U_{n+1}}) * K_r < P_{V \quad L_m}(3) \\ \text{But } L_m \beta_{U_{n+1}} \text{ must be constrained such that Equ 4} \\ \text{holds} \end{aligned}$ 

$$0 \leq (L_m \beta_{U_{n+1}} \leq 1$$
(4)
Such that Equ 5 validates this scenario
$$\sum_{i=1}^{M} \langle L_m \beta_{U_{n+1}} \rangle = 1$$

Where  $K_{\rm F}$  = Job generation rate of user  $U_{\rm i}$ 

 $P_{V L_{IM}}$  = Processing capability of  $L_{IM}$  cluster branch and server instance

(5)

From Fig4, recall that the OpenFlow load manager



 $L_{m}$  dispatches the user  $U_{ij}$  obs to the V processors it manages immediately it is received. This will result in zero waiting queue state (ZWQS) at the load balancer controller  $L_{m}$ . In the DCCN model, this results in excellent service availability condition and it is a measurable QoS parameter.

Consequently, in the DCCN, the work considered a resource pool server cluster V managed by a cluster load controller  $L_m$  (OpenFlow Controller) as a High Performance Processor (HPP) with an average job processing rate given by

 $P_{V \ L_{m}} = \sum_{k=1}^{mL_{m}} [P_{V \ L_{m}} * K](6)$ Where  $K_{T}$  = Job generation rate of user  $U_{i}$   $P_{V \ L_{m}} * K$  = Processing capability of V processor K managed by the OpenFlow load controller  $L_{m}$ .

#### 3.4 DCCN Job Completion Delay Model

In proposed DCCN architecture, job completion time is described as the transaction time between the user  $U_i$ via the load manager  $L_m$  to the servers, in addition to possible waiting time at the queue and the processing time at both the  $L_m$  and server cluster.

Let user Uitransaction time be given  $asT_u$ . This influenced by the average size of the job, the distance between the user and the DCCN<sub>Ib</sub> the bandwidth available to users, etc.

At each cluster DCCN<sub>Ib</sub>, the server chain is viewed as an M/M/1 queuing system (Poisson arrivals and the exponentially distributed times), Okafor, K.C. (2015).

Recall that this type of queuing system is a single server queuing system with Poisson inputs, exponential service times and unlimited number of waiting positions  $(\infty)$  as shown in Fig 4.

Now, let the system response time (delay time) of a job workload sent by user *Ui* at cluster C be given by

$$T_{L_{m}}(\mathbf{t}) = \left[ Tu + \frac{1}{\left[ P_{V} \quad L_{m} - R_{L_{m}} \right]} + \frac{1}{\left[ P_{V} \quad L_{m} * K - R_{L_{m}} * K \right]} \right] (7)$$
Where  $R$  — is a striked rate of the pressure of the

Where,  $R_{L_m^*K}$  = job arrival rate of the processor K managed by the load balancer  $L_m$ 

The system delay time as depicted in Equ.7 comprises of the load balancer delay time, the user transaction time and the server processing queuing times.

If a user  $U_i$ sends job to the load controller  $L_m$  via  $D_g$ , it will require  $T_{L_m}(t)$  time in seconds from  $L_m$  and server cluster at a price  $P_{r_{L_m}}$ . The  $T_{L_m}(t)$  is the time required to communicate to a user request

per time. Therefore, the functional cost model for Uiwith its job allocation  $\vartheta_{I_{cl}}$  is given by

$$C_{U}\left(\vartheta_{J_{\Omega}}\right) = \sum_{L_{m}=1}^{n} \langle L_{m}\beta_{U_{n+1}}\rangle * P_{T_{L_{m}}} * T_{L_{m}}(t)$$
(8)

From the foregoing, the DCCN Piossion Server Queuing System Model (User Job Tasks) and the DCCN service availability models have been addressed for FPGA based processor cores.

#### 3.5 DCCN Switch Models

There are two categories of datacenter switches: connectionless packet switching, also known as datagram switching, and connection-oriented virtual circuit switching. In the first case, each packet includes complete routing information, and is routed by network devices individually. The second case requires a pre-allocated virtual circuit path before transferring any packet. To demonstrate the flexibility of the proposed approach, this work builta simplified FPGA cloud DCN model for only connection-oriented virtual circuit switching.

In order to provide more predicted latencies and take advantage of new high-speed switching FPGA technologies, fordatacenters, the connectionoriented virtual circuit switching is proposed to be directly implemented on FPGAs core. Therefore, it is straightforward to build highly-accurate models for these circuit-switching switches. Here are the abstractions employed in the FPGA based DCCN, Okafor, K.C. (2015):

- Ethernet QoS related features (e.g. support of IEEE 802.3 class of service (CoS))
- Use of a simplified source routing. This is used to simplify modelling of packet routing, and it was noted that source routing is actually a component of many datacenter-switch research proposals. The DCCN switch model implement large flow tables.
- Abstract packet processors: Commercial datacenter switches include many pipelined packet processors that handle different tasks such as MAC address learning, VLAN membership, and so on. The processing time of each stage is relatively constant regardless of packet size, and the time can be as short as a few hundred nanoseconds to a few microseconds. The work simply



- employedFIFO with runtimeconfigurable delays to model packet processing. Virtual-output-queue switch and common scheduling algorithms are to be used in developing the mode in Fig 4 and 5.
- High link bandwidth and cut through switching fabrics for low port-to-port latencies. These are essential features to evaluate FPGA based high performance datacenter switches, which are very hard to deploy at scale in real world because of cost issues.

While there has been an increasing research interest in the area of Software Defined Network (SDN), using Openflow-capable switches in datacenter, simplifies the switch control performance for packet switching networks.

It is known that the basic architecture of existing commercial datacenter switches fall into two categories: output-queue and combined inputoutput queue with virtual output queue. The former has no on-chip congestion, and minimum buffering latencies. Therefore, it is the ideal memory architecture for FPGA based switches. However, this architecture requires that all switch ports can simultaneously read/write into the shared buffer memory, which demands enormous bandwidth. The combined input-output queue architecture uses separate egress and ingress memory structures to reduce the memory bandwidth requirement at the cost of a more complex internal switching arbitration design. The input queues employ Virtual-output-Queue (VoQ) like structures to eliminate the issue of Head-of-Line (HOL) blocking. The proposed FPGA based DCCN switches will adopt this architecture in the implementation. Combined input-output queue switches have lower performance compared to an ideal output-queue switch. Nevertheless, the basic queuing architectures of either input or output queues are similar.

#### 4 SYSTEM IMPLEMENTATION

#### A. Network Simulation

Firstly, this work built the FPGA server process models on top of the Riverbed modeller academic edition 17.5 (https://splash.riverbed.com/community/productlines/steelcentral/university-supportcenter/blog/2014/06/11/riverbed-modeler-academicedition-release) with a heavily modified host-cache design. The server model supports the full 32-bit OS. The VirtexUltraScale was emulated into the service processors shown in Fig 7.The FPGA service processor cores for gateways, firewall, and servers at was built for low cost. The work utilized programmable FPGA devices in order to achieve high performance in the datacenter

network design. Now, from figure 7, the components introduced in this test environment includes: the server farm- firewall router, emulated NEC IP8800 OpenFlow controller, application and profile configuration centres. This testcenter configuration emulates all the web servers, such as DCCN server1, server2, server3, server4, server5,.....N and six location A-F with active users. Table 1 shows the simulation design specifications. Here, the DCCN load balancer (emulated as NEC IP8800 OpenFlow-enabled switch or load balancer) simulates firewall. The service processes connects user locations to the cloud internet which processes services concurrently. The applications (http service) runs on the OpenFlow controller whose job is to dispatch the requests to the server clusters.

The work used the testCenterconfigs to emulate clients that possess authorized public IP addresses (class A). On the other hand, servers are placed on the DCCN as clusters. The Network Address Translation (NAT) function translatesthe server's private IP addresses (192.168.1.0/24) to the public IP address (100.1.1.100). As a result, clients can access the DCCN servers successfully once authorised. Figure 7 shows that the HTTP cloud servers will equally receive request packets from the clients. From the design, OpenFlow switch is capable of providing Firewall, switching functions.

	rubie 1. binanation besign parameters for been			
S/N	Design Parameters/Specifications	Values		
1	Server link Connection	40GB Ethernet		
2	No of Servers	2(Intel D850EMVR-3GH) [ <u>VirtexUltraScale</u> ]		
3	Virtualization Type	Full/Active		

 Table 2: Simulation Design parameters for DCCN





4	Load balancer	1(Random Http)[ <u>VirtexUltraScale]</u>
5	Load balancer Address	Auto Configured
6	IP Routing Protocols	Enabled
7	IP Gateway Function	Enabled
8	Server farm Firewall	Etherent2_Slip8_Firewall [ <u>VirtexUltraScale</u> ]
9	Internet Cloud	Enabled
10	Client gateway	75xx_router [ <u>VirtexUltraScale</u> ]
11	Packet Discard Ratio (PDR)	0.0%
12	Packet latency (Secs)	None
13	Number of Client Location	6 Subnets(Min)
14	Failure Recovery Modelling (FRM)	Enabled
15	Profile Configuration	Web User (http Service[EETACP]), VirtexUltraScale
16	Client Address	Auto Assigned



Figure 7: Riverbed Modeller DCCN Service Process Model

Fig 8 shows the FPGA service process throughput (Bits/Secs). High-throughput architectures for maximizing the number of bits per second that can be processed by the design. A high throughput design is one that is concerned with the steady-state data ratebut less concerned about the time any specific piece of data requires to propagatethrough the design (latency). The idea with a high-throughput design is that an FPGA pipeline design conceptually allowsvariousstages of manipulation and processing. This Pipeline concept found in FPGA makes facilitates very-high-performance in the

proposed DCCN. The throughput yield in Fig 8 is over 90% while maintaining stability over a long period of time. This behaviour literally accommodates scalable cloud services.

In Fig 9, a low latency scenario was observed. A low-latency design is one that passes the data from the input to the output as quickly as possible by minimizing the intermediate processing delays. Oftentimes, a low-latency design will require parallelisms, removal of pipelining, and logical short cuts that may reduce the throughput or the max clock speed in a design. As showed in the plot, with less than



0.002 secs at a simulation time of less than 50secs, the system already switches the loads. This is also very acceptable for high performance workloads.

From the plots in figures 8 and 9, IT network infrastructure that processes bandwidth intensive applications will scale optimally by using FPGA based devices. This is because; their capabilities can support critical business processes that depend on speed and scalability. The data networking and telecommunication market segments can explore FPGA capabilities to drive their business processes further.



Figure 8: FPGA Service Process Throughput (Bits/Secs)





Figure 9: FPGA Service Process latency/delay Response (Secs)

#### 5 RESULTS AND CONCLUSION

FPGA's are increasingly used in the datacenter for CPU data access as well as to improve the overall performance of the network. Cloud datacenter built on FPGA core fabrics will offer excellent throughput and low latency as deduced from the results of this study. It is evident that most hardware architects in multiple markets are looking for alternatives to ASIC, and solutions that can address their bandwidth, performance, integration, and power requirements. This research validates the proposal of FPGA cores in cloud DCNs. This paper proposes FPGA-based hardware acceleration architecture for the processing of high-throughput datacenter feeds, providing a solution able to operate up to the maximum data-rate of the network connection, while offering a very low latency path from the network interface to the consuming process, irrespective of network load. Various mathematical models on the FPGA based DCN was discussed while analysing the obtained results from the Riverbed modeller version 17.5. This work have extended the capabilities of the Smart green Energy management system discussed in an earlier workwhile showing that FPGAs

can enhance the performance of computing systems. Future work will derive a comparison between FPGA based DCCN proposal and the traditional DCN cores. Also, various varying parameters will be investigated in future studies.

#### REFERENCES

- Ganesan, A. et al (2011). What is the impact of cloud computing on the datacenter interconnect? In Hot Interconnects.
- Fares, A. et al **(2010).** Hedera: Dynamic Flow Scheduling for Datacenter Networks, In NSDI.
- Greenberg, A., Hamilton, J.R., Jain, N. Kandula, S., Kim,C., Lahiri, P., Maltz, D.A. Patel, P., and Sengupta, S. (2011): "VL2: a scalable and flexible data center network",In *Communications of the ACM*, March.Vol.54, No.3.doi :10.1145/1897852.1897877.
- A.Shieh, et al **(2010)**. Building programmable Datacenter Networks without programmable switches. In ACM HotNets.

Abu-Libdeh, et al **(2010)**. A.Symbiotic Routing in Future Data Centers. In ACM SIGCOMM.

A. Libdeh, et al (2010).Symbiotic Routing in Future Data centers. In ACM SIGCOMM, 2010.



- A.F. Tenca, et al **(2000).** A Scalable and Unified Multiplier Architecture for Finite Fields *GF(p) andGF(2m),*" in *Cryptographic Hardware and Embedded Systems*, Ed. 2000, Lecture Notes in Computer Science, pp. 94-108, Springer, Berlin, Germany.
- Andrew P. et.al, (2014). A Reconfigurable Fabric for Accelerating Large-Scale Datacenter Services. In IEEE Conf.
- Chuanxiong et al., (**2008).** DCell: A Scalable and Fault-Tolerant Network Structure for Data Centers", In Proc. of SIGGCOM Seattle, Washington, USA, Aug.17<sup>th</sup> - 22<sup>nd</sup>.

Tavakoli, et al **(2009)**. Applying NOX to the datacenter.In HotNets.

- Guo, et al (**2008).** DCell: A Scalable and Fault-Tolerant Network Structure for Data Centers, In *Proc. of SIGGCOM Seattle*, Washington, USA, Aug.17–22<sup>nd</sup>.
- Udeze, C.C., et al (2014). Performance Analysis of R-DCN Architecture for Next Generation Web Application Integration", In Proc. of the 6th ICAST Conference, Covenant University Otta, 19th-31<sup>st</sup>, pp.1-12
- Thacker, C.(2007). Rethinking data centers. October 2007.
- Guo, C. et al (2009). BCube: a high performance, server-centric network architecture for modular data centers", In Proc. Of the ACM SIGCOMM Conference on Data communication, (in SIGCOMM '09), pp.63–74, New York, NY, USA.
- Joseph, D.A et al (2008). A policy-aware switching layer for data centers. In SIGCOMM '08, Pp.51–62, New York, NY, USA, 2008. ACM.
- Kohler, E. et al (2000). The Click Modular Router.In ACM Transactions on Computer Systems, Aug, 2000, Pp.263-297.
- Guo, C. et al **(2010)**. Second Net: A Data Center Network Virtualization Architecture with Bandwidth Guarantees. In ACM CoNext
- Lu, G. et al (2011). ServerSwitch: A Programmable and High Performance Platform for DataCenter Networks, In NSDI'11 Proc. of 8th USENIX conference on Networked systems design and implementation, Pp.15-28.
- Google G-Scale Network, online at http://www.eetimes.com/electronicsnews/ 4371179/Google-describes-its-OpenFlownetwork, **2012**.

https://splash.riverbed.com/community/productlines/steelcentral/university-support-

Center/blog/2014/06/11/riverbed-modeleracademic-edition-release

- Naous, J. et al (2008). NetFPGA: Reusable Router Architecture for Experimental research, In PRSTO.
- Al-Fares, M. et al (**2008).** A scalable, commodity data center network architecture," In Proceedings of the ACM SIGCOMM 2008 conference on Data communication. ACM,Pp. 63–74.
- Alizadeh, M. et al **(2008).** Datacenter Transport Mechanisms: Congestion Control Theory and IEEE Standardization. In 46<sup>th</sup> Annual Allerton Conf. on Communication, Control, and Computing.
- Codrescu, M. et al., (2009).RouteBricks: Exploiting Parallelism to scale Software Routers. In ACM SOSP,.
- Okafor, K.C.(2015). A Model for Smart Green Energy Management Using Distributed Cloud Computing Network", Ph.D. Thesis, University of Nigeria Nsukka, April.
- Okafor, K.C, et al (2015). Characterization of Distributed Cloud Computing Network (DCCN) Server Resource Pool Using Virtualization Dynamics", In International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE), India, Vol4.No.2, Pp.280-294.
- Okafor.K.C, et al **(2014).**Gateway Load Balancing Service In Cloud Data Centre Environments Using Throughput Metric Index", In American Journal of Computation, Communication and Control, (AJCCC), AASCIT, USA; 1(1), pp.8-17.
- Okafor.K.C, et al **(2014).** Performance Evaluation of Discrete Event Process Algorithms for a Two-Tier High Performance Cloud Computing Network Architecture (DCCN)", in International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE), India, Vol 4.No.2, Feb.2015, Pp.295-313.
- Okafor,K.C.and Nwaodo T.N(2012):"A Synthesis VLAN Approach to Congestion Management in Datacenter internet Networks",



International Journal of Electronics and Telecommunication System Research, Vol 5, No 6, May Pp.86–92.

- Yang, R., et al(2013).FPGA Implementation of a Parameterized FourierSynthesizer", IEEE International Conference on Electronics, Circuits, and Systems, Abu Dhabi ÉmiratsArabes Unis.
- Joost, R., and Salomon, R. "Advantages of FPGA-Based Multiprocessor Systems in Industrial Applications"
- Mysore, R.M., et al **(2009)**.PortLand: A Scalable Fault-Tolerant Layer 2 Data Center Network Fabric, In ACM SIGCOMM, August 17<sup>th</sup>–21<sup>st</sup>, Barcelona, Spain.
- Sun DatacenterInfiniBand Switch 648, http://www.sun.com/products/networking/inf iniband.jsp,2009.
- Han, S. Et al **(2010)**, PacketShader: A GPU-Accelerated Sotware Router. In ACM SIGCOMM,
- Switching Architectures for Cloud Network Designs.Online:
  - http://www.aristanetworks.com/en/Switching Architecture wp.pdf, 2009.
- Zhangxi, (2013). Using FPGAs to Simulate Novel Datacenter Network Architectures At Scale", A PhD thesis- University Of California, Berkele.



### Reusable Components in the Political Domain -Computational Algorithms for Enhancing Participatory Social Accountability

#### E. Okewu<sup>1</sup> and J. Okewu<sup>2</sup>

<sup>1</sup>Centre for Information Technology and Systems, University of Lagos, Lagos, Nigeria <sup>2</sup>Department of Visual and Creative Arts, Federal University Lafia, Nassarawa State, Nigeria

<sup>1</sup>okewue@yahoo.com, <sup>2</sup>jonathan.okewu@gmail.com

#### ABSTRACT

Public confidence in the political process is normally enhanced by access to information. Both the government and the governed are entitled to voice concerns on policies, programmes and projects that could transform the lives of citizenry through prudent management of state resources. However, in recent memory, fiscal indiscipline has characterized the African political domain with vested interests served above and beyond the collective good of the people. In the aftermath of monumental corruption, the African landscape is branded by weak institutions, poor infrastructure and low human capital development index and as such, African countries are less competitive in the comity of nations. To address this downward spiral, some measures have been suggested, one of which is application of Information and Communication Technologies (ICTs). In this study, the researchers explored the use of ICTs in promoting participatory social accountability against the background that corruption thrives in the absence of transparency and accountability. We applied design and software engineering skills to achieve the objective. The work designed and implemented a multi-tier online real-time e-Democracy system using Componentbased Software Engineering (CBSE) approach. Requirements were gathered, proposed system modelled using Universal Modelling Language, and development done on Microsoft SharePoint platform. The proposed e-Democracy system will drive conversation between leaders and the led in an African political setting in such a fashion that the masses can voice concerns over policies, programmes, projects and services that political office holders are offering. This background check will ensure that fiscal budgets are judiciously utilized for the greatest good of the greater number of people. The paper reports outcome of simulated government-people interactions using the e-Democracy prototype system and indicates that optimal use of reusable components in the political domain could significantly promote participatory social accountability. It equally brings to the fore the role of computational algorithms in driving healthy political dialogue for inclusive development.

### KEYWORDS: COMPONENT-BASED SOFTWARE ENGINEERING, DEMOCRACY, E-GOVERNMENT, INCLUSIVE DEVELOPMENT, PARTICIPATORY SOCIAL ACCOUNTABILITY

1. INTRODUCTION

The failure of democracies across Africa to deliver dividends of governance has generated heated



debates. Some schools of thought are of the view that devolution of power at the centre and empowerment of confederating units for competitiveness is the solution. This theory does not resonate with other schools of thought, though there is a general consensus that good governance is key to survival. Meanwhile, religious bigotry, ethno-religious cleansing, tribalism, nepotism, terrorism, kidnapping have taken root in the wake of governance failure and pervasive corruption. Little did Africa's founding fathers know that years down the line the continent would be gasping for breadth for socio-economic development and growth, giving its abundant natural and human resources. Their optimism of a great Africa and a future economic frontier began to fade soon after many countries gained independence in the 1960s corruption fuelled by poor leadership had contributed to this downward spiral. Studies have shown that corruption thrived mainly because transparency and accountability were lacking in governance - lack of participatory social accountability. Leaders were more egocentric than people-centric. And in this scenario, state resources were plundered for vested interest leaving sour legacies of poor infrastructure, weak institutions, and low-skilled human capital. Clearly, such a business landscape is unattractive to investors let alone becoming a global economic frontier.

To realize the dreams of the founding fathers, the need to put systems in place for managing corruption in Africa cannot be overemphasized.

As expected, countries within the continent share a lot in common in terms of identity, culture, custom and history, but still have their peculiarities. Equally, there are similarities and peculiarities across governments ministries, departments and agencies (MDAs) coupled with complex administrative and operational bureaucracies. This scenario suggests that a reuse-based approach -(Component-based software engineering (CBSE)) to modelling and developing software that will be relevant to several countries and MDAs would yield considerable benefits. The application of CBSE for developing the e-Democracy systems is promising because of the reasons stated in (Pressman, 2009; Debayan, 2011; Sommerville, 2011; Crnkovic and Larsson, 2003).

This paper reports a study of the use of CBSE for developing an e-Democracy system for the political domain in Nigeria. It empirically investigates the claims of endemic corruption in Africa that some authors have alluded to in the literature (Bamidele, 2013; Tom and Attai, 2014; Raymond, 2008). Specifically, we proposed an e-Democracy system that implements a transparency process in Africa's democracies.

The remaining segment of this paper comprises the following: section 2 gives the background of study and related work; section 3 presents the methodology and the selected case study; section 4 focuses on results and discussions; and finally, the paper is concluded in section 5.

#### 2. BACKGROUND AND RELATED WORK

### 2.1 Corruption Incidence and Measurement in Africa

In order to manage corruption successfully in Africa, there is need to understand its depth and breadth. The menace is real and of monumental threat to the socio-political and socio-economic development of the continent. Against the backdrop that measuring corruption will enable us to manage it more effectively and efficiently, we measured corruption incidence in Africa for a period of 3 years (2012 - 2014) relying on data from the global corruption perception index by the global corruption watchdog, Transparency International. Our findings indicate that Africa is the poster child for corruption and poor governance. Of the 175 countries measured for the 3-year period, data clearly indicates that vast majority of African countries were at the bottom of the table, a segment classified as highly corrupt. To corroborate these statistics and global perspective with ground-level perspective, virtually all institutions in Africa ranging from legislature to judiciary are under the yoke of corruption and mismanagement of public resources. A case in point: despite the pervasive poverty in African countries like Nigeria and Kenya, the cost of governance is high. In a comparative study, Tom and Attai (2014) provided statistical evidence (Table 1) comparing the emoluments of legislators and their minimum wages in six countries, Nigeria and Kenya inclusive.



Table 1: Comparison of legislators' pay in six countries					
Country	Legislators' pay monthly	Legislators' pay annually	Minimum wage monthly	Minimum wage annually	% of legislators' pay that is minimum wage
Nigeria	Senate N15.2m Reps N10.6m (\$69,533)	Senate N182m Reps N127m (\$834,402)	N18,000 (\$118.15)	N234,000 (\$1,536) inclusive of 13th month salary	0.13% 0.18%
India	N305,058 (\$1,999)	N3.7m (\$23,988)	Varies from state to state, sector to sector	-	-
US	N2.2m (\$14,500)	N26.5m (\$174,000)	N191,667 (\$1,257)	N2.3m (\$15,080)	8.6%
UK	N1.3m (\$8,686)	N15.9m (\$104,228)	N283,333 (\$1,883)	N3.4m (\$22,597)	21.68%
Sweden	N1.2m	N14.1m	Set by annual collective bargaining deal	-	-
France	N1.02m (\$6,754)	N12.3m (\$81,951)	N275,433 (\$1,805)	N3.3m (\$21,664)	26.73%
Kenya	N2.2m (\$14,543)	N26.7m (\$175,000)	N10,534 (\$6,917)	N126,413 (\$830)	0.4%

It is alarming, if the excerpt above is anything to go by, that politicians in Africa (Nigeria and Kenya) compared to their counterparts in developed societies have positioned themselves to get stinkingly rich while the masses get impoverished by poverty. It is instructive to note that only 0.47% and 0.13% (0.18%) of legislators' pay constitute minimum wage in Kenya and Nigeria respectively while it is 8.6%, 21.68% and 26.73% in US, UK and France respectively. Since these pay structures don't reflect the economic realities of these African countries, it is safe to say that politics in Africa favours political officials to the detriment of the masses. Little wonder then that would-be political office holders would do anything, including corrupt practices, to secure position at all cost. An inclusive dialogue platform that allows the masses to air their concerns will certainly checkmate these fiscal excesses.

To substantiate the sentiment in some quarters that corruption has assumed the status of a culture in Africa, both the highly and lowly placed citizens engage in the practice with impunity. Overtime, the weakening of institutions has made things worse as prosecution of offenders is now a mirage. Nonetheless, the judicial system, home and abroad, has been instrumental in bringing to book some high profile corruption cases in Africa. This situation calls for urgent measures to get Africa out of the woods. One of such measures we proposed in this study is the technology approach (e-Democracy) which provides a technique of enhancing participatory social accountability in the political process using e-Government, e-Governance and e-Administration.

#### 2.2 Information and Communication Technologies (ICTs)

Information and communication technologies (ICTs) include any communication device encompassing radio, television, cellular phones, computer and network hardware and software, satellite systems and so on, as well as the various services and applications associated with them, such as videoconferencing and distance learning. This research study revealed that the application of ICTs in the political domain has popularized the concept of e-Government as both a developmental



agenda and academic discipline. The pervasive nature of ICTs and its use for all-round development has equally popularized the concept of ICT for Inclusive Development (ICT4ID). Consequently, governments, NGOs, international organizations and individuals have advocated the use of ICTs for improving living standards, particularly in developing economies where the underprivileged are shut out of mainstream economy.

### 2.3 Inclusive Innovation, Social Inclusion and Inclusive Development

Irrespective of race, social status, religion, gender, all hands have to be on deck for growth and development. Inclusive innovation defines the characteristic of new goods and services that are created for those who are denied access from the development mainstream—most especially lowest incomes and/or the poverty line. These new technologies are for the lowest ladder in the social hierarchy - which includes: Information and Communication (like mobile phones, mobile services and telecentres); Agriculture (better seed varieties); Healthcare (vaccines); and etc. Quiet a number of countries are increasingly using inclusive innovation in various sectors and fields, like China, Thailand, and other national India, Indonesia, governments. Some African governments have engaged ICTs to directly empower the less privileged across Africa (Matous et al., 2014, Bhavni et al., 2008, Masiero 2013 ) just as Okewu (2013) reports that a number of social security schemes have been put in place in Nigeria, targeting the vulnerable groups. However, as a result of corrupt practices in the social security schemes, the study advocated the integration of biometric system. Biometrics ensured that only genuine beneficiaries were handed social benefits.

Inclusion is an organizational practice and goal aimed at bringing together different groups or individuals having different backgrounds like origin, age, race and ethnicity, religion, gender, sexual orientation and gender identity in a culturally and socially accepted fashion for equal treatment towards the attainment of set common objectives (Miller and Katz, 2002; Gasorek, 1998; Hyter and Turnock, 2006; Roberson, 2006). Apparently, governance is meant to drive set objectives of growth and development. Online real-time participatory social accountability will harness the vocal skills, experience and knowledge of even citizens at the bottom of the economic ladder (Arnstein, 1969) for inclusive development. The need to coordinate effort and share information dynamically among all players in the polity regardless of gender, ethnicity, age, sexual orientation, disability or poverty is key to the success of governance. Inclusive development in the context of participatory social accountability galvanise the electorate to use their voices to dictate the pace of growth and development by prioritizing initiatives that focus on productive employment as a means of increasing the incomes of the poor and excluded groups and raising their standards of living (lanchovichina and Lundstrom, 2009).

Languepin (2010) reports that efforts to promote digital inclusion has made the use of mobile phones predominant among vulnerable people and proven to be a success as the rapid distribution of mobile telephony has made it possible for poor people to have easy access to useful and interactive information, thereby increasing the momentum of political discourse.

The United Nations Development Programme (UNDP) opines that many people are excluded from development because of their gender, ethnicity, age, sexual orientation, disability or poverty. Development can be inclusive - and reduce poverty - only if all groups of people contribute to creating opportunities, share the benefits of development and participate in decision-making.

#### 2.5 Related Work

Some of the previous efforts that are related to this study in the literature are presented as follows.

Raymond (2008) warned of the negative impact of corruption on foreign investment in Africa. He perceived corruption as a dent that could dwindle investors confidence and therefore derail the dream of Africa emerging as the new economic frontier of the world. As a curative measure, he advocated sustained investment in institutions, legal structures and civil service reform. Though he drew a link between corruption and the downward



investment trajectory in Africa as well as suggested curative measures, his approach did not include the use of ICTs as a tool and technique for instituting participatory social accountability for inclusive development, the main motivation for this work.

Kalika (2000) studied e-Management, a variant of e-Administration and postulated that the emergence of e-Management shows that the integration of ICTs in organizations makes it possible to generate new practices and services and to reduce the cost and time of the services rendered. The researchers observed that the introduction of technology and innovation into the political domain represents a new way of conducting government business. We equally observed that cost-reduction and efficiencyinjecting sentiments, though not the chief objectives of this work, are hallmarks of an online real-time dialogue platform like e-Democracy. The work did not indicate whether the e-Management application could be used to enhance participatory social accountability for the purpose of inclusive development. Neither was there a mention of the software approach used.

Michel (2005) articulated e-Administration, e-Government, e-Governance and the Learning City as a typology of citizenship management in France. This two-year research study opined citizenship implies a certain model of relationship between citizens and their government. This type of relationship can be conceived in several ways. Citizenship can be presented in the form of an object to be governed in various ways. Citizens can also be positioned as either passive or active agents in governance. By and large, the researchers characterised citizens as source of ideas and initiatives that provides a mutual enrichment in a democracy and would at the same time be actors and determinants of the rules. The study concluded that role of government officials and the ICT tools in governance remains to be imagined. Though citizenship management is a cardinal component of a democratic system, the work did not elaborate on how participatory democracy serves as a tool for consolidating inclusive development in Africa. The author also fell short of mentioning whether software applications were implemented let alone highlighting the software approach used.

Okewu (2014) applied component-based software engineering (CBSE) approach to the development of a university e-Administration system. The study compared CBSE approach to traditional software development approach, and based on the outcome of the experiment in an educational domain, concluded that CBSE has gains over traditional approach. Another outcome of the study was a ntier enterprise application that added practical and commercial value to the university community. It is worth mentioning that this study focused on an industrial experience in the education sector while our research seeks to explore reusable components in the political domain for galvanising participatory social accountability with a view to mitigating corruption and promoting inclusive development.

ICT4D in Africa (2014) and The Transformation (2014) focused on harnessing the power of ICTs in Africa. They hinted that over the last decade, the ICT access in Africa has increased immensely. As access increases, opportunities arise to leverage ICT to extend timely information and services to previously underserved populations, and to increase productivity and innovation in the public and private sectors. Examples of this are the increase in the number of people who are able to acquire mobile phone service, improved disease monitoring and vaccination planning and mbanking services using the mobile to extend access financial services to populations that never before had a bank account. Nonetheless, they observed that despite dramatic ICT improvements made, significant access gaps are still there. They however were shy of mentioning that ICTs could be used to inclusive development strengthen through participatory social accountability in Africa.

Other instances of ICT implementations in Africa and other continents that have been mentioned in the literature include Heeks and Molla (2012), Development Informatics (2014), Transparency International (2014). Although most of these efforts highlighted the benefits, costs, challenges, and implementation scenarios of ICTs in the African context, the emphasis were not the use of ICTs for enhancing participatory social accountability.

Summarily, it was observed from the literature that none of the previous studies had focused on the implementation of an e-Democracy system in an



African context using CBSE approach with a view to promoting participatory social accountability for inclusive development, which is the main motivation for this work. We are also motivated by the fact that transforming the African political and business landscapes through judicious utilization of state resources as a consequence of fiscal discipline will make the continent investors' delight and pave the way for its emergence as the new global economic frontier.

### 3. METHODOLOGY - CBSE FOR E-DEMOCRACY SYSTEM

In order to identify the requirements for the e-Democracy system and get a sense of current developments in the democratic space, relevant literature were consulted, interviews held and several observations made in the Nigerian polity as a reasonable representation of Africa.

Governance process and procedures were modelled using the Unified Modeling Language - in particular: use cases, collaboration diagrams, sequence diagrams, class diagram and deployment diagram.

The study used the CBSE approach for the design and development of the proposed solution on the Microsoft SharePoint development platform. This platform supports four components namely: Document Library, Custom List, Task and Site. Microsoft SharePoint was used because it supports the doctrine of component reusability with COM+ as its component model. It is also a web-based platform that advocates distributed computing.

A number of controlled experiments were conducted using simulated data. The participants in the experiments were given the opportunity to interact with the system. Thereafter, their perspectives of the system were sought. Figure 1 shows visual version of the activity workflow of the methodology used for the work.



Figure 1: Activity Workflow of the Research Methodology

We used Nigeria as a case study amid established concerns that political developments in Nigeria represent in microcosm happenings in Sub-Saharan Africa. It is made up of a Federal Government, 36 State Governments (including the Federal Capital Territory, Abuja) and 774 Local Government Areas with a population of over 170 million people (Figure 2).





Figure 2: The geopolitical space called Nigeria (Source: Google). Downloaded December, 2014

Using the objective-methodology mapping in the Table 2, we embarked on the CBSE lifecycle activities to actualize the proposed e-Democracy system as a measure for promoting participatory social accountability and inclusive development in the polity.

#### Table 2: Objective-methodology mapping

SN	Objective	Methodology
1.	To provide online real-time information on proposed public policies, programmes, projects and services by government for approval and feedback of electorate	Design and implement an e- Government sub- system
2.	To provide online real-time information on proposals by citizens on people- oriented policies, programmes, projects and services for government's consideration	Design and implement an e- Governance sub- system
3.	To provide online real-time information on services offered by government for the wellbeing of citizens	Design and implement an e- Administration sub-system
4.	To provide integrated information system	Integrate e- Government, e- Governance and e-

for transparency and	Administration
accountability in	sub-systems to get
governance for	e-Democracy
purposes of inclusive	system
development in the	
African body politic	

#### 3.1 Requirements Analysis and Specification

In this section, the requirements for the e-Democracy system are analysed. The requirements were gathered by interview, observation and studying existing processes and systems. The functional requirements (Table 3) include add information, access information, edit information, and delete information while the non-functional requirements include quality requirements that span performance, security, usability, aesthetics, availability, reliability, scalability, fault tolerance, modifiability, portability and interoperability. The e-Democracy system incorporates mechanisms that respond to these requirements. The mechanisms include components of e-Democracy: e-Government, e-Governance and e-Administration.

#### **Table 3: Functional requirements**

Requirement ID	Requirement	Brief Description
Ro1	Add Information	The system shall allow every user to add proposals on public policies, programmes, projects and services
Ro2	Access Information	The system shall allow every user to retrieve and view information on public policies, programmes, projects and services
Ro3	Edit Information	The system shall allow users to edit information on public policies, programmes, projects and



		services
Ro4	Delete Information	The system shall allow users to delete information on public policies, programmes, projects and services

Use Case modelling was used (Figure 3) to galvanise requirements analysis in a bid to comprehend the core functionalities and usage scenarios associated with the identified requirements. Use case diagram captures the functional aspects of a system by visually representing what transpires when an actor interacts with the system (Aggarwal and Singh, 2008).



Figure 3: Use Cases for e-Democracy System

The use cases empower both government officials and the citizenry as end-users to articulate information on strategic tools (policies, programmes, projects) and social services that transform the body politic. This promotes participatory social accountability.

#### 3.2 System and Software Design

Component reusability and distributed computing are closely linked in an enterprise application. To leverage on this relationship, we designed the n-tier enterprise architecture in Figure 4 for the proposed e-Democracy solution incorporating mechanisms that respond to user requirements.





Figure 4: e-Democracy System Multi-tier Architecture

The n-tier architecture comprises of presentation layer, logic layer and database layer. While government officials and the citizens operate at the presentation layer as end-users using devices like personal computers and phones to voice concerns on initiatives (policies, programmes, projects) and social services, the logic layer made of clustered application servers process the information which is stored in the database layer. The fact that these layers are networked makes dialogue to be online real-time.

The interfaces between the respective ρ-Democracy components are captured in the component diagram in Figure 5. The role of component model (COM+) in this architecture is critical as it provides standards and support services to components though they are not represented physically in the software architecture in line with best practice (Gorton, 2011). The interdependence between e-Government, e-Governance and e-Administration as graphically illustrated underscores the reality that projects conversations policies, on and programmes initiated by one party (say political office holders) can be supported or rebuffed by

another (say the people) which may translate into social services for the well-being of society with high sense of fiscal discipline. Hence, any lamebrain scheme by political office holders to siphon public funds can be checkmated. And this way, resources would be freed for overall and inclusive development.



Figure 5: e-Democracy component diagram

The e-Democracy reusable components (e-Government, e-Governance, and e-Administration) were subsequently built from Microsoft SharePoint using standard components such as Document Library, Custom List, and Tasks and are explained in Table 4.

Table 4:	e-Democracy	components
----------	-------------	------------

S N	e- Democracy Component s	Descripti on	Function Points
1.	e- Governmen t	Useful for accessin g, editing, storing, and deleting informat ion on governm ent	addGovernmentl nputInformation( ), accessGovernme ntInputInformati on(), editGovernmentl nputInformation( ), deleteGovernme ntInputInformati on()



## 12<sup>th</sup>

## **International Conference**

		proposal s on policies, program mes, and projects that will socially and economi cally transfor m the lives of the people.	add(itizonclaput)
2.	e- Governance	Gives the user the right to access, edit, store, and delete informat ion on citizen's inputs or proposal s on policies, program mes, and projects that will socially and economi cally transfor m the lives of the people.	addCitizensInputI nformation(), accessCitizensInp utInformation(), editCitizensInput Information(), deleteCitizensInp utInformation()
3.	e- Administrati on	Empowe rs political stakehol ders to	addServicesInfor mation(), accessServicesInf ormation(), editServicesInfor

access,	mation(),
edit,	deleteServicesInf
store,	ormation()
and	
delete	
informat	
ion on	
social	
services	
designed	
to	
transfor	
m the	
lives of	
the	
people.	

The class diagram for the e-Democracy system is given in Figure 6 below.





Other design tools we used include collaboration diagram, sequence diagrams, class diagram, analysis class, design component and elaborated design class, class elaboration, algorithm, composite (appropriate) interfaces, and elaborated deployment diagram (Pressman, 2009; Martin, 1998).



The researchers used deployment diagram to represent the location of key packages or components of the e-Democracy system (Pressman, 2009) as illustrated in Figure 7 below.



Figure 7: e-Democracy deployment diagram

The study used class elaboration and algorithm to present abstraction details of the components and functions of the proposed e-Democracy system as shown in Figures 8 to 13 below.



Figure 8: e-Government component design

e-Government		
user name user address		
addGovernmentInputsInfo()		
accessGovernmentInputsInfc()		
editGovernmentInputsInfc()		
deleteGovernmentInputsInfo()		

Figure 9: e-Government class analysis

The e-Government algorithm design is as follows:

Procedure addGovernmentInputsInfo() governmentInputsInfo "" while (not endOfGovernmentInputsInfo()) governmentInputsInfo addInput() return(governmentInputInfo)

Procedure accessGovernmentInputsInfo() while (not endOfGovernmentInputsInfo()) getInfo(governmentInputsInfo) return

Procedure editGovernmentInputsInfo() while (not endOfGovernmentInputsInfo()) getInfo(governmentInputsInfo) editGovernmentInputInfo() return(governmentInputsInfo)

Procedure deleteGovernmentInputsInfo() while (not endOfGovernmentInputsInfo()) getInfo(governmentInputsInfo) deleteGovernmentInputInfo() return(governmentInputsInfo)





getInfo(citizensInputsInfo) deleteCitizensInputsInfo () return(citizensInputsInfo)



Figure 12: e-Administration component design

e-Administration	
user name user address	
addServicesInfo()	
accessServicesInto()	
editServicesInfo()	
deleteServicesInto()	

Figure 13: e-Administration class analysis

The e-Administration algorithm design is as follows:

Procedure addServicesInfo() servicesInfo "" while (not endOfServicesInfo()) servicesInfo addInput() return(servicesInfo)

```
Procedure accessServicesInfo()
while (not endOfServicesInfo ())
getInfo(ServicesInfo)
return
```

Procedure editServicesInfo()



Figure 11: e-Governance class analysis

The e-Governance algorithm is as follows:

```
Procedure addCitizensInputsInfo()
citizensInputsInfo ""
while (not endOfCitizensInputsInfo())
citizensInputsInfo addInput()
return(citizensInputsInfo)
```

```
Procedure accessCitizensInputsInfo ()
while (not endOfCitizensInputsInfo ())
getInfo(citizensInputsInfo)
return(citizensInputsInfo)
```

```
Procedure editCitizensInputsInfo ()
while (not endOfCitizensInputsInfo ())
getInfo(citizensInputsInfo)
editCitizensInputsInfo ()
return(citizensInputsInfo)
```

```
Procedure deleteCitizensInputsInfo()
while (not endOfCitizensInputsInfo ())
```



while (not endOfServicesInfo ())
getInfo(ServicesInfo)
editServicesInfo()
return(servicesInfo)

Procedure deleteServicesInfo() while (not endOfServicesInfo()) getInfo(servicesInfo) deleteServicesInfo() return(ServicesInfo)

#### 3.3 Implementation and Unit Testing

This study used Microsoft SharePoint as the development platform for the tailor-made e-Democracy system. SharePoint is a web-based enterprise development tool that makes components available for reuse, the components are called services. It is a component platform that uses Microsoft COM+ as the component model. It provides an integrated development environment (IDE) and its core components are Document Library, Custom List and Tasks, which are not only independent but are distributed (Gorton, 2011). The development of the e-Democracy system was achieved using incremental approach. The minimal e-Democracy system to start with was the e-Government module. Other modules were added on incremental basis. As testing of componentbased system is different from normal software testing, the authors used black box testing as it is more suitable in component based systems (Sirobi and Parashar 2013).

#### 3.4 System Integration

With e-Government as minimal e-Democracy system, regression test was conducted as more modules were interfaced to ascertain that there were no interface errors. Else, if they existed, debugging took place before adding another module. In the final analysis, e-Government was the most tested component in the e-Democracy system. It is the most referenced component in the proposed system. Test cases were developed and used to test the various components (Table 5) prior to integrating them. Then we used system test cases at the point of integration for regression tests. As typical of component-based systems, black-box testing was performed for all components (Beydeda and Gruhn, 2003). Table 5 shows components and their function points.

Table 5. Component testing (e-Democracy function	I
points)	

S N	Component	Function Points
1.	e- Government	addGovernmentInputsInfo(), accessGovernmentInputsInf o(), editGovernmentInputsInfo(), deleteGovernmentInputsInf o()
2.	e- Governance	addCitizensInputsInfo(), accessCitizensInputsInfo(), editCitizenInputsInfo(), deleteCitizensInputsInfo()
3.	e- Administratio n	addServicesInfo(), accessServicesInfo(), editServicesInfo(), deleteServicesInfo()

#### 3.6 System Verification and Validation

We verified and validated the processcorrectness and requirements-compliance of the e-Democracy architecture by examining the various software representations - requirements documents, design documents and program code. Our concern was to ascertain that user requirements had been well catered for in each software representation in the build-up process just as we ensured that the software product met both operational needs of users and emergent properties.

#### 3.7 Operation Support and Maintenance

A number of both technical personnel and endusers were trained to test-run the application. While the end-users operated the software, the technical staff provided sustained support.



Use as many sections as are necessary but don't use more than three levels. Number them according to the format indicated in this guideline.

#### 4. RESULTS AND DISCUSSIONS

In a bid to extract information and measure outcomes of this study objectively, we used software experiment. We also evaluated possible threats to our results.

#### 4.1 Results of Software Experiment

The e-Democracy site was created as a community site using Microsoft SharePoint enterprise development platform. As the name suggests, it is a site where political community members discuss topics of common interest. Behind this software engineering is the political message that in the absence of participatory social and accountability, transparency inclusive development would be a mirage and corruption will thrive. We set up an experimental design in University of Lagos, Nigeria precisely at the Centre for Information Technology and Systems and testrun the system from near (Lagos environs) and remote locations, including Lafia in Nassarawa State and Abuja, all in Nigeria as shown by broken arrows in Figure 14 below. By this act, the researchers used an n-tier web-based e-Democracy system to mimic the sensation of sustained dialogue between the led and leaders in a political setting.



Figure 14: e-Democracy experimental corridor -Lagos, Abuja and Lafia (Nassarawa)

The simulation experiment confirmed that ICTs could bridge the gap between the governed and their governors and more importantly entrench transparency and accountability in conducting government business. The subiects who participated in the experiment concurred that the outcome of the experimental survey was a seamless and robust online real-time communication among political stakeholders on topical public policies, programmes, projects and services that are people- and result-oriented. The bottom line is that the e-Democracy dialogue framework engendered a sense of transparency procedure capable of enhancing participatory social accountability and inclusive development. Though experienced platform-dependent we and hardware-dependent challenges particularly testing from remote locations (Lafia and Abuja), this only suggested that more robust infrastructure was needed for wide-scale implementation.

#### 4.2 Evaluation Threats

There is the possibility that an expanded evaluation of the different components of the e-Democracy system could unearth new perspective of things. In any case, the subjects (who are Nigerians) that participated in the experiment survey have the required experiential knowledge of the Nigerian political domain and its governance challenges - corruption, weak institutions, poor infrastructure, low-skilled human capital and unfavourable investment climate. They equally had sufficient practical engagements with the e-Democracy system. This offered them good basis to make objective comparison between the old way of information exchange between political actors and the new e-Democracy model of interactions. Therefore, there is sufficient reason to take their views seriously.

Equally important is the fact that only two classes of users were involved in the evaluation one representing government while the other represented the masses, which could in a sense limit the statistical significance of the outcome. However, the result of the experiment clearly indicates that both parties who incidentally are the stakeholders in any political space were adequately represented and were functional in the dialogue on



acceptable service levels . This is considered to be a good result because at this juncture in the project, the core objective is to gain a first impression of the degree of transparency and accountability injected into governance by the e-Democracy system. Therefore, despite the constraint of using a limited number of evaluators, there is sufficient grounds to conclude that there is a positive and preferential disposition to the e-Democracy system as a tool for enhancing participatory social accountability and promoting inclusive development. It means optimal utilization of state resources as a consequence of transparency, and translates into wellbeing of the citizenry. We can thus generalize that the CBSE developed e-Democracy system is effective for enhancing participatory social accountability for purposes of inclusive development.

#### **5 CONCLUSION**

Entrenching participatory social accountability in Africa's polity could well be the antidote to poor leadership and massive corruption. Corruption tends to be endemic and most detrimental in African states that are transitioning from one form of governance to another or fragile from violent conflict. New ideas and approaches on how to tackle corruption are being developed on sustained basis. These approaches span from helping to establish good governance to incorporating the power of civic involvement. As part of our contribution, this study used the ICTs approach to entrench transparency and accountability in governance through online real-time interaction between government and citizens on public policies, programmes, projects and services. The component-based approach to e-Democracy system adopted in this study offers benefits to political stakeholders - government, citizens, among others. First, the output of the study is an enterprise software that has practical and commercial value to the governance context in Africa. For the governance process, it will help to inject transparency and accountability. Also, the free-flow of information between government MDAs and citizens on public policies, programmes, projects and services means better utilisation of state resources for the wellbeing of citizens. Yet another benefit is the fact that the incidence of

corruption will be greatly reduced as studies have shown that corruption thrives in the absence of transparency and accountability. It will also promote inclusive development as the masses are empowered to determine the bearing of development through their voices. Hence, the combination of e-Government, e-Governance and e-Administration provides а platform for participatory social accountability and inclusive development. This will in turn promote developmental politics, socio-economic enhancements and attract foreign direct investments as Africa looks forward to becoming the next preferred global investment destination. Finally, the e-Government research community has been presented with a new case study report of the application of e-Government, which adds to the existing body of knowledge in this field.

#### 6. REFERENCES

- Arnstein, S.R., 1969. A ladder of citizen participation. Journal of the American Institute of Planners, 35(4), 216-224, 1969.
- Bhavni, A., Chiu, R. W., Janakiram, S. P., Silarszky, & Bhatia, D., 2008. The Role of Mobile Phones in Sustainable Rural Poverty Reduction, June 15, 2008.
- Bamidele, O., 2013. Corruption, Conflict And Sustainable Development In African States, The African Symposium. An online journal of the African Educational Research Network, 42 Volume 13, No. 1, June 2013, ISSN# 2326-8077.
- Beydeda, S. and Gruhn V., 2001. An Integrated Testing Technique for Component-Based Software, Computer Systems and Applications. ACS/IEEE International Conference.
- Crnkovic, I. and Larsson M., 2003. Building Reliable Component-Based Software Systems, Artech House Publishers, ISBN 1-58053-3272.
- Debayan, B., 2011. Component Based Development -Application In Software Engineering, Indian Statistical Institute.
- Development Informatics ICT4D 2016 (2014). New Priorities for ICT4D Policy, Practice and WSIS in a Port-2015 World, (Paper No. 59). Retrieved 2 November.



- Gasorek, D., 1998. Inclusion at Dun & Bradstreet: Building a High-Performing Company. The Diversity Factor 8/4 (Summer) 2529.
- Gorton, I., 2011. Essential Software Architecture, Second Edition, Springer.
- Heeks, R. and Molla, A., 2012. Impact Assessment of ICT-for-Development Projects: A Compendium of Approaches, Development Informatics Group, Institute for Development Policy and Management University of Manchester, Arthur Lewis Building, Manchester, M13 9PL, UK. Retrieved 7 August.
- Hyter, M. C. and Turnock, J.L., 2006. The Power of Inclusion: Unlock the Potential And Productivity of Your Workforce, John Wiley & Sons.
- Ianchovichina, E. and Lundstrom, S., 2009. Inclusive growth analytics: Framework and application, Policy Research Working Paper Series 4851, The World Bank.
- ICT4D in Africa, 2014. Harnessing the power of ICTs. http://www.infodev.org/. Retrieved November 25.
- Kalika, M., 2000. Internet et les sciences de gestion, PUF, Paris.
- Languepin, O., 2010. How mobile phones can help reduce poverty.
- Masiero, S., 2013. Innovation and Best Practice in Mobile Technologies for Development, Economic and private sector professional evidence and applied knowledge services.
- Matous, P., Yasuyuki Todo, Y., Ishikawa, T., 2014. Emergence of multiplex mobile phone communication networks across rural areas: An Ethiopian experiment, Network Science.
- Michel, H., 2005. e-Administration, e-Government, e-Governance and the Learning City: A typology of Citizenship Management using ICTs. The Electronic Journal of e-Government Vol. 3 Issue 4, pp 213 - 218.
- Miller, F.A. and Katz, J.H., 2002. The Inclusion Breakthrough: Unleashing the Real Power of

Diversity, San Francisco: Berrett-Koehler Publishers.

- Okewu, E., 2013. Design and Implementation of Social Security Scheme: The Role of ICT. Proceedings of the 11th International Conference of Nigeria Computer Society. Iloko-Ijesha, Nigeria.
- Okewu, E. and Daramola O., 2014. Componentbased Software Engineering Approach to Development of a University e-Administration System. IEEE 6th International Conference on Adaptive Science and Technology (ICAST). Ota, Nigeria.
- Pressman, R.S., 2009. Software Engineering: A Practitioner's Approach, 7th ed.
- Raymond, G., 2008. Causes of Civil War in Africa, Introduction to African Politics, UB/12002602.
- Roberson, Q.M., 2006. Disentangling the Meanings of Diversity and Inclusion in Organizations. Group & Organization Management 31/2:212-236.
- Sirobi N. and Parashar A., 2013. Component Based System and Testing Techniques. International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 6, June.
- Sommerville, I., 2011. Software Engineering, Ninth edition.
- Tom, E.J. and Attai, A.J., 2014. The Legislature And National Development: The Nigerian
- Experience. Global Journal of Arts Humanities and Social Sciences, Vol.2,No.9, pp. 63-78, November 2014 Published by European Centre for Research Training and Development UK (www.eajournals.org).
- The Transformational Use of Information and Communication Technologies in Africa, 2014. www.eTransformAfrica.org. eTransform Africa. Retrieved 28 November.
- Transparency International, 2014. Poverty and Corruption in Africa (PCA) Programme.



### INTRODUCING A MODELTO IMPROVE RECENT SUB-SAHARAN AFRICA NETWORKED READINESS INDEX

#### \*P. K. Oriogun<sup>1</sup>, K. K. Agbele<sup>2</sup>, K. D. Aruleba<sup>3</sup> and A. O. Agho<sup>4</sup>

<sup>1,2,3,4</sup>Elizade University, Department of Mathematics and Computer Science, Ilara-Mokin, Ondo State, Nigeria

<sup>1</sup> peter.oriogun@elizadeuniversity.edu.ng, <sup>2</sup> kehinde.agbele@elizadeuniversity.edu.ng, <sup>3</sup> kehinde.aruleba@elizadeuniversity.edu.ng, <sup>4</sup> adrian.agho@elizadeuniversity.edu.ng

#### ABSTRACT

In this paper we specifically address a number of recent recommendations suggested by The Global Information Technology Reports 2013 - 2015 with respect to Sub-Saharan Africa countries slow progress in developing its ICT infrastructure within the current world of Big Data and ICT for inclusive growth; warning that if care is not taking in terms of improving the framework condition for innovation and entrepreneurship there could be a digital divide between the developed and developing economies. In addressing this possible digital divide, we proffered a solution through a model we developed recently at Elizade University for auditing ICT infrastructure projects in a developing economy. We isolated and fully analyzed the sub-Saharan Africa data from the two reports and present our model in this paper. We argue that the model will address the specific shortcomings of the region in terms of strengthening its ICT infrastructure, and will improve the framework for innovation and entrepreneurship.

KEYWORDS: BIG DATA, ENTREPRENEURSHIP, ICTS, INFRASTRUCTURE, NRI

#### 1. INTRODUCTION

The impetus for this paper is to address a number of the recommendations of The Global Information Technology Reports 2014 & 2015 (Bilbao-Osorio et al., 2014; Di Battista et al., 2015), published by World Economic Forum, regarding sub-Saharan Africa slow progress in developing its ICT infrastructure, especially by expanding the share of the population covered by, and having access to, mobile phones and expanding the number of internet users. It further stresses that strong ICT infrastructure can be achieved through improving the framework conditions for innovation and entrepreneurship in order to avoid what it termed as a new digital divide. In addressing this possible digital divide in terms of ICT infrastructure in sub-Saharan Africa, Oriogun, et al., (2015) proposed a model for auditing ICT infrastructure projects in developing economies. We believe that this model will assist government and ICT regulatory bodies in alleviating


poverty and to improve the lives of indigenous communities by building the capacity of target populations to harness the opportunities that ICTs offer.

The Global Information Technology Report 2014 & 2015 (Bilbao-Osorio et al., 2014; Di Battista et al., 2015) offers benchmark for ICT uptake in a world of Big Data and ICT for inclusive growth in terms of what has been termed as the Networked Readiness Index - NRI. Commonly known as the 3Vs, The Gartner IT Glossary (2015) defines Big Data as "highvolume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making". In the context of this paper, we are using the term Networked to mean adaptation and/or implementation of Information Communication Technologies (ICTs). This paper will first analyze through descriptive statistics, the 35 Sub-Saharan Africa countries from the 148 countries investigated by The Global Information Technology Report worldwide, this will be followed by explanation of our proposed model to address the (Bilbao-Osorio et al., 2014; Di Battista et al., 2015) Networked Readiness Index in detail, making references to aspect of our framework that will particularly deal with the inadequacies of ICT infrastructure in the region, followed by some tentative remarks through a brief discussion and finally some concluding remarks.

### 2. THE GLOBAL IT REPORT AND NETWORKED READINESS INDEX FOR SUB-SAHARAN AFRICA (2014 – 2015)

The Global IT Report (Bilbao-Osorio et al., 2014) and the Networked Readiness Index (NRI) were created over 13 years ago based on a new way of organizing and managing economic activity in terms of the new opportunities and the impact that the Internet provided for businesses (Information Technologies Group, 2000; Dutta et al., 2012)

Table 1: NRI Ranking / Scoring for 35 Sub-Saharan Africa Countries -Adapted from (Bilbao-Osorio et al., 2014)

Sub- Saharan Position in 2014 (from 35 countries )	Sub- Saharan Africa Country in 2014 Report	NRI Ranking / Score 2015 (from 143 countries)	NRI Ranking / Score 2014 (from 148 countries)	Sub- Saharan Position in 2014 (from 35 countries )	Sub- Saharan Africa Country in 2014 Report	NRI Ranking / Score 2015 (from 143 countries )	NRI Ranking / Score 2014 (from 148 countries )
1	Mauritiu s	45 (4.5)	48 (4.31)	19	Swaziland	125 (3.0)	126 (3.00)
2	Seychell es	74 (4.0)	66 (4.02)	20	Mali	127 (3.0)	127 (3.00)
3	South Africa	75 (4.0)	70 (3.98)	21	Gabon	122 (3.0)	128 (2.98)
4	Rwanda	83 (3.9)	85 (3.78)	22	Ethiopia	130 (2.9)	130 (2.95)
5	Cape Verde	87 (3.8)	89 (3.73)	23	Cameroon	126 (3.1)	131 (2.94)
6	Kenya	86 (3.8)	92 (3.71)	24	Malawi	133 (2.8)	132 (2.90)
7	Ghana	101 (3.5)	96 (3.65)	25	Lesotho	124 (3.0)	133 (2.88)
8	Botswan a	104 (3.4)	103 (3.43)	26	Sierra Leone	Excluded in 2015	134 (2.85)
9	Namibia	102 (3.5)	105 (3.41)	27	Benin	Excluded in 2015	135 (2.82)
10	Gambia, The	108 (3.3)	107 (3.38)	28	Burkina Faso	132 (2.8)	136 (2.78)





11	Zambia	114 (3.2)	110 (3.34)	29	Mozambiq	129 (2.9)	137 (2.77)
					ue		
12	Nigeria	119 (3.2)	112 (3.31)	30	Madagasca	135 (2.7)	139 (2.74)
					r		
13	Senegal	106 (3.3)	114 (3.30)	31	Mauritania	138 (2.5)	142 (2.61)
14	Uganda	116 (3.2)	115 (3.25)	32	Angola	140 (2.5)	144 (2.52)
15	Zimbab	121 (3.1)	117 (3.24)	33	Guinea	142 (2.4)	145 (2.48)
	we						
16	Liberia	Excluded	121 (3.19)	34	Burundi	141 (2.4)	147 (2.31)
		in 2015					
17	Cote	115 (3.2)	122 (3.14)	35	Chad	143 (2.3)	148 (2.22)
	d'ivoire						
18	Tanzania	123 (3.0)	125 (3.04)				

The concept of *Big Data* relate to every conceivable data that we make use of on a daily basis as humans can be captured and studied as part of what has been termed as *Big Data*. From a total of 49 countries in Sub-Saharan Africa, The Global Information Technology Report 2014 (Bilbao-Osorio et al., 2014) captured 35 countries as shown in Table 1, therefore, only 71.42% of the countries in Sub-Saharan Africa countries were represented. The average NRI 2014 value is 4.12 (average of the lowest score 2.22 Chad and the highest score of

6.06 for Finland). However the average of the highest 12 (35%) of the 35 Sub-Saharan Africa countries captured in the study is 3.67 (Mauritius 4.31; Seychelles 4.02; South Africa 3.98; Rwanda 3.78; Cape Verde 3.73; Kenya 3.71; Ghana 3.65; Botswana 3.43; Namibia 3.41; The Gambia 3.38; Zambia 3.34; Nigeria 3.31). We observe in Figure 1, a graphical representation of sub-Saharan Africa Networked Readiness Index scores for 2014 (Adapted from GITR 2014 using 54 NRI indicators)





Figure 1: Graphical Representation of Sub-Saharan Africa Networked Readiness Index 2014 Scores (Adapted from Bilbao-Osorio et al. (2014) using 54 NRI indicators)

### 3. ANALYSIS OF THE NRI SCORES FOR SUB-SAHARAN AFRICA 2014 – 2015 INCLUSIVE

The analysis of the Global Information Technology Report 2014 for Sub-Saharan Africa included 35 countries as listed in Table 1. The remaining 14 countries (South Sudan, Democratic Republic of Congo, São Tomé and Príncipe, Central African Republic, Republic of the Congo, Equatorial Guinea, Djibouti, Eritrea, Somalia, Comoros, Sudan, Guinea-Bissau, Niger and Togo) were not covered in the report - Bilbao-Osorio et al. (2014), and Liberia, Sierra Leone, Benin were omitted from the report in 2015. The term Environment in this paper refers to Political and regulatory environment as well as the Business and innovation environment in the context of the Networked Readiness Index (NRI) framework. The Affordability of ICT Infrastructure has to match the appropriate knowledge and Skills acquisition before the environment is deemed to be at the state of Readiness. Furthermore, the Environment has to interact with Business, Government and Individual in order to have meaningful state of Usage of available information communication technologies (ICTs) resources. There must also be a method of measuring both the economic and social impacts of the ICT infrastructure environment before we can measure what has been termed as the Networked Readiness Figure 1 shows the overall scores for the 35 sub-Saharan Africa included in the Networked Readiness Index study of 2014. This is closely followed by the sub-indices for Environment, Readiness, Usage and Impacts accordingly. On a scale of 1 to 7 scoring regime, sub-Saharan Africa average is 3.27 (maximum score of 4.31 for Mauritius and minimum score of 2.23 for Chad). The country scoring the maximum NRI for 2014 worldwide is Finland with a score of 6.04. The average NRI 2014 score worldwide is 4.14 (minimum score Chad 2.22, and maximum score Finland 6.04). In terms of statistics, from the 148 countries that took part in the study, we can safely infer that

-NRI (Dutta et al., 2012). From Figure 2, it is evident that there was no data for Liberia, Sierra Leone and Benin as there were no 2015 NRI scores for these countries. It is possible that the reason for this may be due to the Ebola crisis when the data was being gathered, especially in Liberia and Sierra Leone.

The Di Battista et al. (2015) report did not explain the reason for excluding Benin in its 2015 NRI scores. According to Di Battista et al. (2015) the scores of the NRI showed that the ICT revolution has not extended worldwide. They claim that the capacity of a country to benefit from ICTs is strongly influenced by its stage of development'. They concluded that:

The performance of sub-Saharan Africa is particularly disappointing: 30 of the 31 countries included in the sample appear in the bottom half of the NRI rankings. The only exception is Mauritius, at 45th. This country has progressed three places since last year and eight since 2012. Among the large economies of the region, Nigeria drops seven places to 119th. South Africa drops five to 75th—it is now third in the region behind Mauritius and Seychelles (74th). In contrast, Kenya (86th, up six) has been slowly improving since 2012... p.14 Di Battista et al. (2015) noted further that: the developing world still lacks universal, reliable, and affordable Internet. The lack of proper ICT infrastructure and cost of fixed broadband access, mobile broadband is becoming the technology of choice, but it remains prohibitive in too many countries. p.26

59.14% of the maximum score was achieved worldwide (4.14 / 7.0). If we separate the scores of developing countries in sub-Saharan Africa (a total of 35 countries for this study), we discover that only 46.71% of the maximum score (3.27 / 7.0) was achieved from this sample population. This is one of the reasons that Bilbao-Osorio et al. (2014) suggest that sub-Saharan Africa need to develop a more solid ICT infrastructure, and should also improve the 'framework conditions for innovation and entrepreneurship in order to avoid the emergence of a new digital divide' in what Bilbao-Osorio et al. (2014) referred to as possible age of 'digital revolution'





Figure 2: NRI Scores for 35 sub-Saharan Africa Countries (adapted from Di Battista et al. 2015; Bilbao-Osorio et al. 2014)

In Figure 2, the Sub-Sahara Africa countries considered in the (Bilbao-Osorio et al. 2014; Di Battista et al. 2015) reports are shown with the range of their NRI scores.

### 4. A MODEL FOR AUDITING ICT INFRASTRUCTURE PROJECTS IN DEVELOPING ECONOMIES

We offer our model for auditing ICT infrastructure projects in a developing economy (see Figure 4) as

We offer our model for auditing ICT infrastructure projects in a developing economy (see Figure 4) as a way of addressing some of the shortcomings identified in the Global Information Technology Reports (2014 and 2015). At the heart of the model are government and ICT regulatory bodies making policies that will have profound influence on ICT a way of addressing some of the shortcomings identified in the Global Information Technology Reports (Bilbao-Osorio et al. 2014; Di Battista et al. 2015). At the heart of the model are government and ICT regulatory bodies making policies that will have profound influence on ICT investments and use. The model is also robust enough to be able to alleviate poverty and improve the lives of indigenous communities by building the capacity of target populations to harness the opportunities that ICTs offer.

investments and use. The model is also robust enough to be able to alleviate poverty and improve the lives of indigenous communities by building the capacity of target populations to harness the opportunities that ICTs offer. Figure 4 shows our published model for auditing ICT infrastructure projects in a developing economy





Figure 4: A Collaborative Model for Auditing ICT Infrastructure Projects in Developing Economies (Oriogun, et. al, 2015)

The model defines four key stakeholders (this ultimately will include any other possible users of the ICT infrastructure projects) in ICT project delivery in developing countries. These stakeholders are; the government/ICT regulatory bodies, ICT professional bodies, ICT consultants and in-house ICT departments. The inter-relationship amongst these groups creates synergies captured by the proposed model as well. These synergies are encapsulated in ICT professional courses, certification of ICT personnel, together with the actual ICT projects. This knowledge will come from both higher education and from professional certification.

### 4.1 Stakeholders in the Model for Auditing ICT Infrastructure Projects in Developing Economies

Government, through ICT regulatory bodies plays a pivotal role in auditing ICT infrastructural projects. This is done through setting national policies, standards, specifications and requirements to govern the execution of projects. Within the context of developing countries, this role cannot be ignored as "best practices" are yet to be developed and adopted in many parts of the industry. The model recognizes the pivotal role of government in the process by bringing together and regulating the activities of all the other stakeholders in the delivery of ICT infrastructure projects. At the heart of the model are government and ICT regulatory bodies making policies that will have profound influence on ICT investments and use. The intersections of the three major sectors of the model will provide incentives for ICT education and training at all levels, make provisions for strengthening ICT education and training, provide incentives for private sector research and development, together with a transparent ICT infrastructure audit.

The second stakeholders recognized by the model are ICT professional bodies. These are formal associations of ICT practitioners which have as their focus the development and advancement of the various ICT disciplines and technologies. In collaboration with government regulatory bodies, the professional bodies have the responsibility to



develop and maintain a professional "body of knowledge" in ICT. This knowledge is then codified and disseminated through ICT professional courses to create a pool of "certified" ICT professionals. Hence, the model provides for these professional bodies to regulate the quality and quantity of ICT competencies possessed by the practitioners of the discipline.

The third recognized stakeholders by the model are ICT Consultants, they are assumed by the model to mean companies set up in order to provide ICT goods and services. Naturally, the proposed model requires that these consultants/vendors provide these goods and services in collaboration with the in-house ICT departments of the organizations they consult for. This provision of goods and services is done through what this paper recognizes as "ICT infrastructure projects." These projects are codified modules of work to be done in order to meet a specific need of the client. In cooperation with the in-house ICT department of the client, and in conformity to policies of government, these projects are defined and executed.

The model requires all companies wishing to undertake ICT infrastructure projects to have our fourth and final stakeholders, in-house ICT department. This is crucial so that proper technical specifications for the project as well as assessment of work done can be handled by an in-house team of competent ICT practitioners. The size and mix of professionals in these in-house ICT departments will vary based on the size and nature of the organization in question. However, these departments should be staffed by personnel who possess the requisite skill in their various ICT subdisciplines to effectively carry out the work of the organization. This "requisite skill" is determined through certification. In the context of this model, it is expected that the practitioners in this in-house ICT department possess both theoretical as well as hands-on knowledge of the ICT tools and techniques to be deployed in the infrastructure projects.

### 5. DISCUSSION

We are particularly interested in the Networked Readiness Index (NRI) for sub-Saharan Africa, as the authors of this paper are from Nigeria, and are working and operating from Nigeria. We observe that from a total of 49 Sub-Saharan Africa countries, only 35 were included in the analysis, there was no particular reason supplied in (Bilbao-Osorio et al. 2014; Di Battista et al. 2015) to explain the rational for not including the 14 countries that were excluded. We are however of the opinions that since a number of these excluded countries are undergoing some radical social and economic conflicts and unrest, this might have been one of the reasons for their exclusion. As the share of the population in this region having access to mobile telephone, and the use of the internet is expanding, the ICT infrastructure is still in its infancy compared to majority of the developed countries.

It was mentioned in the (Bilbao-Osorio et al. 2014) report that a number of countries in sub-Saharan Africa are slow in their uptake of ICT, with no sign of improvement in the near future. The authors cautioned that the 'gap may hamper their capacity to support further economic and social development as the positive impacts of ICTs become more and more apparent' p17. We believe that our proposed model for auditing ICT infrastructure projects in developing economies is an idea tool to address a number of concerns raised in the reports (Bilbao-Osorio et al. 2014; Di Battista et al. 2015) of the last two consecutive years.

### 6. CONCLUSION AND FUTURE WORK

In conclusion, in order for business and innovation ecosystems to flourish, we offer our model for auditing ICT infrastructure projects to the IT community that, when implemented, will address the specific shortcomings of the region in terms of strengthening its ICT infrastructure, and will improve the framework for innovation and entrepreneurship. The model requires all companies wishing to undertake ICT infrastructure projects to have an in-house ICT department with competent, reliable, dependable, trustworthy



professional ICT practitioners. The four major stakeholders in our proposed model for ICT project delivery in developing countries are: the Government/ICT regulatory bodies, ICT professional bodies, ICT consultants and in-house ICT departments. The inter-relationship amongst these groups creates synergies captured in the model. These synergies are encapsulated in ICT professional courses, certification of ICT personnel, together with the actual ICT projects. This knowledge will come from both higher education and from professional certification. Our future work on the model is to begin to gather empirical data from governments, private and public sector organizations from Sub-Saharan Africa countries in order to recommend a detailed framework for promoting and developing ICT infrastructure projects, such that future Networked Readiness Index for the region will be much enhanced and comparable to that of the developed economies.

### 7. REFERENCES

- Bilbao-Osorio B, et al, 2014. The Networked Readiness Index 2014: Benchmarking ICT Update in a world of Big Data, The Global Information Technology Report 2014,
- World

Economic Forum 2014, ISBN 13:978-92-5044-63-0; ISBN 10:92-95044-63-0, pp3 – 31. Available at www.weforum.org/gitr (Accessed 21st May 2015)

- Di Battista A, et al. 2015. The Networked Readiness Index 2015: Taking the Pulse of the ICT Revolution, Global Information Technology Report 2015, World Economic Forum 2015, ISBN: 978-92-95044-48-7, pp3- 30. Available at www.weforum.org/gitr (Accessed 21st May 2015).
- Dutta, S, et al, 2012. "The Networked Readiness Index 2012: Benchmarking ICT Progress
- and

Impacts for the Next Decade." In The Global

Information Technology Report 2012. S. Dutta and B. Bilbao-Osorio, editors.

Geneva:

World Economic Forum. 3–34

Gartner IT Glossary Big Data (2015). http://www.gartner.com/it-glossary/big-

data

(Accessed 21<sup>st</sup> May 2015)

- Information Technologies Group. (2000). Readiness for the Networked World: A Guide for Developing Countries. Cambridge, Mass.: Center for international Development at Harvard University, 2000
- Oriogun P, et al. 2015. Towards a Model for Auditing ICT Infrastructure Projects in a Developing Economy, British Computer Society Publication, To Appear in SQM 2015 Conference, University of Loughborough UK, 30<sup>th</sup> March 2015, ISBN 978-0-9926958-6-6



### A NEW VECTOR SPACE MODEL FOR GRAMMATICALITY GRADING IN HANDCRAFTED GRAMMAR

\*M. K. Aregbesola<sup>1</sup>, R. A. Ganiyu<sup>2</sup>, S. O. Olabiyisi<sup>3</sup>, E. O. Omidiora<sup>4</sup>, and O. O. Alo<sup>5</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, Elizade University Ilara Mokin, Ondo State, Nigeria

<sup>2, 3, 4, 5</sup>Department of Computer Science and Engineering, Ladoke Akintola University of Technology Ogbomoso, Oyo State, Nigeria

<sup>1</sup>kehinde.aregbesola@elizadeuniversity.edu.ng, <sup>2</sup>raganiyu@lautech.edu.ng, <sup>3</sup>soolabiyisi@lautech.edu.ng, <sup>4</sup> eoomidiora@lautech.edu.ng, <sup>5</sup>ooalo@lautech.edu.ng

#### ABSTRACT

Grammaticality is a concept within the fields of Natural Language Processing and Computational Linguistics that deals with the measure of grammatical correctness or incorrectness of a natural language expression. Among the methods employed for grammaticality evaluation, handcrafted grammar is the most precise. The major drawback of handcrafted grammar however is that its evaluation process scarcely computes gradience of grammaticality. It therefore evaluates the grammaticality of an expression in a binary (right or wrong) fashion. This approach however, for instance, does not provide second language learners a scale for measuring progress made in the learning process. To solve this problem, a number of mechanisms were observed from which a vector space approach was considered to be appropriate, having been successfully applied over a wide range of disciplines, including computational linguistics. A vector space approach was therefore adopted as a suitable means of addressing the gradience problem in handcrafted grammar. In this work therefore, a new vector space model for grammaticality grading was formulated using key identifiable grammaticality error features (missing word, extra word, real word spelling, verb form, puctuation and agreement errors) as elements of the vector space. The results showed that the grammaticality gradience  $\mu$ G is such that  $0 \le \mu G \le 1$ . Therefore, by using the new vector space model with handcrafted grammar, grammaticality can be evaluated on a graded scale between 0 and 1. Where the grammaticality gradience value of o is for a perfectly grammatical sentence while 1 is for a totally ungrammatical expression.

Keywords: Grammar, Grammaticality Grading, Handcrafted Grammar, Model, Vector Space.

### 1. INTRODUCTION

This work is part of an ongoing research aimed at developing an automated system for graded grammaticality evaluation, error detection and error correction in English language expressions. British English is the standard adopted in this study.

Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) that deals with the ability of machines to act intelligently within a language



domain (MacKinlay, 2012; Chu, 2003). Grammar is a linguistic concept often considered within the field of NLP in Computational Sciences. It has attracted a lot of research interest from a wide range of disciplines. It is the branch of linguistics that deals syntax, morphology, and sometimes, with semantics. It determines the set of acceptable constructs of a language (Baldridge et al., 2007). As a derivative of grammar, grammaticality describes whether a sentence is grammatical (well-formed) or ungrammatical (ill-formed). Grammaticality has been the focal point of interest in a number of NLP research (Wagner, 2012). Grammaticality judgment has been applied to significantly improve the performance of a number of NLP systems involving language generation, machine translation, text summarization, second language learning, and automated essay scoring (Nerbornne, 2002; Pauls and Klein, 2012). Several approaches have been adopted in evaluating the grammaticality of sentences based on two prevailing schools of thought. The first argues that the natural way languages are learnt is independent of grammatical rules. This school of thought cites that a child learns to communicate in a language without necessarily having to first learn the grammatical rules of the language. This school of thought often uses induced grammar approaches for evaluating grammaticality. The second school of thought lays emphasis on following a strict set of grammar rules (handcrafted grammar) in determining the grammaticality of a sentence. These grammar rules are designed by linguistic engineers and employed as natural language parsers for parsing (evaluating) expressions of the target language. Irrespective of the school of thought being considered, the degree of grammatical correctness or incorrectness is of significant interest in NLP (Parfitt et al., 1999).

Handcrafted grammar which is the focus of this study, makes it possible to describe the question of grammaticality with precision (Blache *et al.*, 2005). It however scarcely expresses the gradience of grammaticality. Thus making it difficult to quantify the degree of correctness or incorrectness of an expression. This quantification feature is however very desirable, especially for second language (L2) learners, and particularly for providing L2 learners with performance feedback. This study is therefore aimed at formulating a gradience mechanism to facilitate grammaticality quantification in handcrafted grammar. Previous works on grammaticality gradience have been almost exclusively within the domain of probabilisticaly induced grammars (Crocker and Keller, 2006).

After considering a number of alternatives, Vector Space approaches having been successfully employed in a variety of disciplines including computational linguistics was considered a suitable mechanism for computing grammaticality gradience in handcrafted grammar. This paper therefore presents a new vector space model for grammaticality grading in handcrafted grammar.

The formulations in the model showed that grammaticality gradience  $\mu$ G is such that  $0 \le \mu$ G  $\le 1$ . Where the grammaticality gradience values 0 and 1 are for perfectly grammatical sentences and completely ungrammatical expressions respectively. Subsequent studies will further explore the effectiveness of the new vector space model.

### 2. LITERATURE REVIEW

### 2.1 Grammar and Grammaticality

Grammar is a term used to define the body of rules that describe the structure of expressions of a language. This is inclusive of the structure of words, phrases, clauses, and sentences. A grammar is said to under-generate if it excludes grammatical and over-generate if it allows sentence. ungrammatical sentences (MacKinlay, 2012). A grammar should avoid both of these situations. There is thus this tension between avoiding overgeneration and maximizing coverage. Grammar developers therefore seek to make grammars as constrained as possible, while they avoid rejecting grammatical sentences. This therefore calls for grammars to exhibit a high degree of precision. The parsability of a sentence however does not necessarily imply its acceptability (Lin, 2007). Therefore, the real world context of word usage must be put into consideration.

Fillmore (1998) and Kay and Fillmore (1999) looked at Construction Grammar in which all objects are considered as constructions. Blache (2006) looked into the robust and efficient property grammar which employs constraints relaxation to allow the system select the best assignments, and not just



the assignment satisfying the constraint system. Thus, permitting the system to deal with more or less grammatical sentences. Constraints based approach makes it possible to process any kind of inputs. Sampson (2007) on the other hand asked the question of whether it was possible to talk about grammar about without talking grammaticality. The role of grammaticality judgment in the learning (acquisition) of a second language cannot be over emphasised (Gass, 1983). The roles of metalinguistic awareness in the learning of a second language (L2) make grammaticality judgment for L2 learners an observation that can be considered either as an implicit or an explicit concept (Bialystok, 1979; 1981). The study of Clark et al. (2013) examined the relationship between grammaticality and acceptability. It viewed grammaticality as a theoretical notion while acceptability was viewed as an empirically testable property. Acceptability was also further explained to be partly dependent on grammaticality.

Automating grammaticality evaluation has taken the front line in a number of linguistic and NLP research. A system equipped with the capability to determine how users will perceive the grammaticality of a sentence could be useful in fields involving language generation, machine translation, and text summarization. It could also be used for automatic essay scoring and for the learning of a second language. Lau et al. (2014) described an experiment explaining that even the most popular translation tools are in many occasions not congruent in performing a backward translation (from L1 to L2 and then to L1 again) of given sentences. It therefore becomes necessary to ensure the grammatical correctness of written expressions to minimize distortion when translated to other languages.

### 2.2 Handcrafted Grammar

Handcrafted grammars are grammars that are manually specified and evaluated by grammar engineers (Witbrock and Mittal, 1999). Handcrafted grammars differ from induced grammars in that in contrast to requiring large scale treebanks, handcrafted grammars provide a single point of control for the handling of syntactic phenomena (MacKinlay, 2012). Owing to this single point of control, grammar rules can be designed to handle even very rare and complex phenomenon, or to even ignore particularly difficult ones. Such precise control is not possible with induced grammars.

It is quite a labour-intensive task to create a grammar of a language which has reasonable coverage and can create meaningful parse trees and semantics. There has nonetheless been a range of concerted efforts targeted at creating broad-coverage grammars for several languages, some of which are closely related in formalism.

### 2.2.1 Transformational Grammars

Chomsky proposed Transformational (1957) Grammar TG in response to his argument that Context-free Phrase Structure Grammar (CFPSG) capture cannot linguistically significant generalization of natural language syntactic properties. The TGs augment the CFPSGs with rules that map their syntactic structures unto other syntactic forms. Thus for instance, "to die" could be transformed into "to cease to live". However, following founded criticisms from Peters and Ritchie (1973) as to the overly unrestricted nature of the TG, Chomsky developed the theory of and binding, Principles Government and Parameters P&P, and Minimalism, which are more constrained versions of TGs (Chomsky, 1981). TGs were literarily abandoned for over a decade after this period (Malchow, 1992).

### 2.2.2 Feature Based Grammars

Following the crisis with TGs came the Feature based grammars, also known as information based grammars, attribute-value grammars, constraint based grammars or unification grammars (Johnson, 1988; Kay, 1979). Unification is one of the crucial techniques employed by all feature based approaches and it is the emphasis of the unification grammar (Malchow, 1992). Among the feature based grammars are the following formalisms: Functional Unification Grammar (Kay, 1979; 1982); Generalized Phrase Structure Grammar (Gazdar, 1981; Gazdar et al., 1985); Lexical Functional Grammar (Bresnan, 1982); Categorial Unification Grammar (Karttunen, 1986; Uszkoreit, 1986); and Head-driven Phrase Structure Grammar (Pollard, 1984; Pollard and Sag, 1987).

Representing the sets of features of a grammar as one complex symbol using feature based grammars has attracted a lot of research interests. Malchow



(1992) explained that there are three basic ways of representing complex categories. These representations are: Tree diagrams, also known as directed acyclic graphs (dags); Terms; and Matrices. The following sentences demonstrate an example

of constraint on English sentences as discussed by Müller (2013):

- a. Kim loves Sandy
- b. \* I loves Sandy

Where '\*' connotes an error in sentence (b) owing to the violation of the constraint that the subject has to agree with the verb both in person and number.

### 2.2.3 Head-driven Phrase Structure Grammar

Considering the different linguistic categories such as noun phrase and verb phrase, it is observed that the noun plays a very important role in the noun phrase. The noun in the noun phrase determines things like the plurality or otherwise. The same is true of verbs in verb phrases and prepositions in prepositional phrases. This concept is referred to as 'headedness' in linguistics. Headedness is used to describe the fact that phrases usually have an item of high importance referred to as the 'head', which determines the main characteristics of the entire phrase. The head of a noun phrase is the noun, while the head of a verb phrase is the verb. This equally applies to other forms of phrases. Headdriven Phrase Structure Grammar (HPSG) clearly expresses the relationship between the heads and their phrases. Context free grammars (CFG) in the natural form lack this ability to express the relationship between the heads and their respective phrases, thus making CFGs unable to give linguistically acceptable analyses of phrases (MacKinlay, 2012).

Harlow (2010) explained that HPSG is indebted mainly to non-derivational approaches such as Categorial Grammar, General Phrase Structure Grammar, Lexical Functional Grammar, and Situation Semantics. It is equally indebted to NLP studies in Computer Science. The similarities between the Transformational Grammars and the Feature based approach (including HPSG) include: Having the same goal of characterizing human linguistic competence; having the same empirical base of acceptability judgment of native speakers; determination of grammaticality based on the interaction between highly articulated lexical entries and universal principles of grammar; they both make use of binding theory but are nonconfigurational; and though applied differently, they both make use of multiple levels of representation. HPSG however has the following features in contrast from the Chomsky approach. In HPSG, syntax does not take preeminence; it is mathematically rigorous; it employs bottom-up approach; it employs structure sharing but not transformations; it is fractal; it is non-derivational; and it employs only local constraints.

### 2.2.4 Parse Trees

The syntactic structure of natural language sentences are usually expressed using context-free grammars which are often represented as ordered, rooted trees known as parse trees or phrase structure trees. As an example, let us consider the parse tree for the English sentence "The fat pony sleeps in the barn" shown in Figure 1. Terms abbreviated in this example include: sentence as 'S', noun phrase as 'NP', verb phrase as 'VP', verb as 'V', determiner as 'D', noun as 'N', adjective as 'A', preposition as 'P' and the nominal category as NOM. Each of the nodes in this tree is either a root node, a branch node, or a leaf node (Carnie, 2013). There can only be one root node, in a tree. A root node (S in this case) is one with no parents (no branches on top of it). A branch node on the other hand is a non-root node (VP and NP in this case) that connects to two or more child nodes. Finally, a leaf node is a terminal node with no child nodes attached to it. The leaves are the actual lexical tokens of the sentence.





Figure 1: A parse (Phrase Structure) tree for "The fat pony sleeps in the barn" (MacKinlay, 2012)

### 2.2.5 HPSG Feature Structure

Features are the basic tools of linguistic description. They are used to describe the information associated with units of a language. There is a feature structure associated with each unit of written or spoken languages, such as word, phrase or sentence. Each lexical entry equally has an associated feature structure. The feature structure is an abstract representation associated with the lexical item which can be realised as a word in a sentence (MacKinlay, 2012). Feature makes it possible for linguistics to talk about information such as the category of a word, and words that must go with it, also known as its theta grid (Carnie, 2013). Features are paired with a value in an Attribute Value Matrix (AVM), and are thus frequently used for representing feature structures. Features in an AVM can also take feature structures as values. For example, the lexical entry for the word 'cat' is shown in a very simplistic AVM implementation in Figure 2a. The entry shows that the Part of Speech (POS) of 'cat' is noun, the number (NUM) is singular (sg), and the person (PERS) is third-person (3). Considering the complex relationships that often exist between NUM and PERS and the need for them to agree, they could be grouped into a single feature known as "agreement" (AGR). Hence the feature structure can further be represented as shown in figure 3. There are different types of feature structures including those that indicate the "word" versus "phrase" status of the constituent elements. Carnie (2013) explained further that the features for a node are next divided into three main classes namely Syntax (SYN), Argument Structure (ARG-ST), and Semantics (SEM).

SYN has it values as structures which are relevant to syntax. It focuses on the formal grammatical properties of the node including the syntactic category, any inflectional properties, and other elements that must be combined with the node. SYN has three main features namely HEAD, Specifier (SPR), and Complement (COMPS). The HEAD feature is the feature that determines the category of the node and its inflectional properties. SPR is the feature that restricts what kind of nodes appear in the specifier position, while the COMPS feature restricts what kind of nodes appear in the complement position. Figure 3 is an example AVM for the word 'letter'.

The angular bracket (< >) indicate ordered lists. The boxed number ([1]) at both the HEAD and SPR features called conference-tags are used to indicate identical structures. In HPSG, the concept of structural identity expressed by the conference-tag facilitates non-transformational analyses. In this example, the structure of the AGR feature of the SPR determiner (D) is identical to the structure of the AGR feature of the HEAD. It was equally noted that in English language, when single count nouns are involved, specifiers are not optional. To illustrate this fact, consider the following sentences:

- (i) I have a letter.
- (ii) I have the letter.
- (iii) \*I have letter.

The asterisk (\*) in the third sentence is used to indicates that something is wrong with the sentence.

The <(PP)> value associated with the COMPS feature in Figure 3 indicates an optional prepositional complement.

For example:

- (i) a letter from the teacher
- (ii) the letter

Carnie (2013) further explained the argumentstructure (ARG-ST) feature, which is the next major feature whose values are structures (ordered-lists) that represent the theta grid. Values held by this feature are required for binding reasons, independent of the SPR and COMPS features. For example, the ARG-ST feature for the verb like:

- (i) < like, [ARG-ST<NP,NP>]>
- (ii) < likes, [ARG-ST < [NP [AGR 3s]], NP >] >

Various categories of selective restrictions can be imposed on the ARG-ST feature. As shown in the second example, the verb "likes" requires a third



person singular subject, which is indicated by inserting an AVM with the specification into the first NP slot in the ARG-ST list.

The last of the major features is the SEM (semantic) feature, whose values are structures (ordered-lists) that represent the semantic properties of a construct. The SEM gives us information about how words and sentences are to be interpreted.



### Figure 2: AVM for the word "cat"

(b) AVM for the word cat using feature structure as value (MacKinlay, 2012)



Figure 3: AVM for the word letter (Carnie, 2013)

### 2.3 Grammaticality Gradience

Grammaticality gradience or quantification is the degree to which a sentence belongs to the grammar of a language (Lau *et al.*, 2014; Chomsky, 1975). Grammaticality is inherently graded by nature and not absolute (Parfitt *et al.*, 1999; Taylor, 1995). The acceptability of sentences can equally be predicted by quantifying its grammaticality. Crocker

and Keller (2006) explained that probabilistically induced grammars are inherently models of gradience in language processing due to their mode of operation. Handcrafted grammar on the other hand scarcely expresses gradience of grammaticality.

Blache *et al.* (2005) however characterized the result of a parse operation into the set of satisfied properties noted as P+, and set of violated ones noted as P-. Thus the sentence is acceptable if its grammaticality falls within a particular range, and unacceptable if its grammaticality falls outside of that range.

Linear Optimality Theory is very effective when applied in weighting the different syntactic constraints associated with a given grammar. Such evaluated constraint weights is then useful for quantifying sentence grammaticality (Keller, 2000; Sorace and Keller, 2005). In a similar sense, Schroder (2002) and Menzel and Schroder (1998) looked at the use of Constraint Dependency Grammar framework for weighting syntactic constraints.

#### 2.4 Vectors and Vector Spaces

Vectors and points are common concepts considered in many areas of Mathematics and Computer Science. They are applied extensively in data compression, image processing, computer vision, computer graphics, and numerical analysis. Two-dimensional vectors can be defined as directed arrows in the plane. The position of the arrow is not important. The length (magnitude) and direction of the arrow are the important features of the vector, and they determine the vector. They can be added, scaled and rotated (Lindeman, 2008). Vectors having the same length and direction are said to be equivalent (Kambites, 2014). Two vectors in the same direction are said to be parallel. The zero vector has a magnitude of zero and is denoted as <u>o</u>. Figure 4a shows a vector  $\underline{a}$  between two points (x<sub>1</sub>,  $y_1$ ) and  $(x_2, y_2)$ . In Figure 4b, the vectors <u>AB</u> and <u>DC</u> are equivalent, because two-dimensional vectors are distinguished only by length and direction. They are thus treated as equal i.e.  $\underline{AB} = \underline{DC}$ .

#### 2.4.1 Vector Space

The concept of vector space is the focus of a number of literatures (Zybura, 2004; Nowak, 2010; Luenberger, 1969; van Hassel, 2009; and



Kondratieva, 2008). A vector space V can summarily be defined as a set of vectors over the field F (such as real, complex, or natural numbers) which may be added together and multiplied (scaled) by numbers referred to as scalars, such that, given  $\underline{x}, \underline{y}, \underline{z} \in V$ , the following eight axioms are satisfied:

i. Associativity of addition

$$\underline{x} + (\underline{y} + \underline{z}) = (\underline{x} + \underline{y}) + \underline{z};$$

- ii. ii. Commutativity of addition
  - $\underline{\mathbf{x}} + \underline{\mathbf{y}} = \underline{\mathbf{y}} + \underline{\mathbf{x}};$
  - iii. Identity element of addition. There exists an element  $\underline{o} \in V$ , called the zero vector, such that  $\underline{y} + \underline{o} = \underline{y}$  for all  $\underline{y} \in V$ ;
  - iv. Inverse element of addition. For every  $\underline{y} \in V$ , there exists an element  $-\underline{y} \in V$ , called the additive inverse of  $\underline{y}$ , such that

 $\underline{y} + (-\underline{y}) = \underline{o};$ 

- v. Compatibility of scalar multiplication with field multiplication. a(by) = (ab)y;
- vi. Identity element of scalar multiplication

 $1\underline{y} = \underline{y}$ , where 1 denotes the multiplicative identity in F;

- vii. Distributivity of scalar multiplication with respect to vector addition. a(x + y) = ax + ay;
- viii. Distributivity of scalar multiplication with respect to field addition. (a + b)y = ay + by.

(Zybura, 2004; Nowak, 2010; Luenberger, 1969; van Hassel, 2009; and Kondratieva, 2008).

### 2.4.2 NLP Applications of Vectors and Vector Spaces

The concept of vector space is considered in most linguistic and NLP literatures from the perspective of lexical and semantic distribution. Semantic vector space models of language make use of realvalued vectors to denote each word that are typically associated with a particular word. Words that typically occur together are assigned values that often depict their probability of occurring together in a sentence. Vector space is employed in a wide range of NLP operations including grammaticality evaluation and error detection. Detailed semantic and syntactic regularity have been successfully captured using



Figure 4 (a): Vector in the plane (b): Two equivalent and parallel vectors (Lindeman, 2008; Kambites, 2014)

The study of Pennington et al. (2014) came up with global vectors for word representation (GloVe). The study focused on highlighting the properties that made the emergence of such captured regularities possible in word vectors. Schmid (2004) however focused on efficient parsing of highly ambiguous context-free grammars using bit vectors. The study of Stolcke (1991) represented a formalism dubbed Vector Space Grammars (VSG) for deriving phrase structure categories that made use of structured samples of a context-free language. Using the connectionist approach, the entire training process made use of adaptation, competition and error back-propagation, all occurring in a continuous vector space. It advocates the use of vectors instead of symbols for the purpose of linguistic category labeling.

Vector Space Semantic Parsing (VSSP) presented in the work of Krishnamurthy and Mitchell (2013) is a framework for learning compositional models of vector space semantics. It applies Combinatory Categorial Grammar (CCG) to define the



relationship between syntactic categories and semantic representations. The representations were handled as vectors and functions on vectors. Using a CCG based semantic parser, texts were parsed into lambda calculus formulae that compute to equivalent vector space representations.

In general, vector space models make use of vectors and operations on vectors to represent the semantics of natural language expressions (Turney and Pantel, 2010). A number of other studies including Coecke *et al.* (2013), Socher *et al.* (2011), Socher *et al.* (2012), Turney (2006), and Rapp (2003) focused on similar concepts. The studies achieved significant performances that corresponded well with human judgment.

Grammaticality was considered a vector concept within these literatures, having both magnitude and direction. The direction of grammaticality is either towards grammatical correctness, or away from grammatical correctness. Grammatical correctness is a state described as Grammatical Equilibrium (GE) within this literature, and is ascribed a gradience value of zero (o).

On a general note, grammars are designed or induced in such a way as to express the state of grammatical equilibrium. For constraint based grammars (Johnson, 1988; Kay, 1979), each appropriate constraint within the grammar enforces the grammar towards being able to produce or determine sentences that are grammatically correct.

Furthermore, although the word 'grammaticality' is generally used to express the state of grammatical correctness or incorrectness of a sentence, it is sometimes used strictly as a measure of grammatical correctness, especially when used in contrast to 'ungrammaticality'. From this perspective, grammaticality is used as a measure of grammatical correctness while ungrammaticality is used as a measure of grammatical incorrectness.

### 3. PROPOSED VECTOR SPACE MODEL FOR GRAMMATICALITY GRADING IN HANDCRAFTED GRAMMAR

### 3.1 The Vector Space (V) of Grammaticality

Let  $\underline{x}$ ,  $\underline{y}$ ,  $\underline{z}$  be weighted entities associated with grammaticality such as possible error categories

including missing-word, extra-word, real-word spelling, verb-form, punctuation and agreement errors, which are uniquely identifiable within a sentence. Also let  $\{\underline{x}, \underline{y}, \underline{z}, ...\} \in V$ . Like any other standard vector space, V is a set of vectors over the field F (which in this case is the set of real numbers R) which may be added together and multiplied (scaled) by numbers referred to as scalars, such that the eight axioms listed in section 2.4.1 are satisfied.

These grammaticality vectors dubbed Mosesean vectors are one-dimensional over the written plane. Therefore, the addition operation on the elements of V is performed by simple arithmetic summation ( $\Sigma$ ). Furthermore, ungrammatical elements within an input sentence are assigned negative values; while counter measures to correct such ungrammaticality are assigned positive values. Thus, the magnitude of grammaticality the Mosesean vectors introduce into the system at any point in time is totally dependent on the magnitude of existing ungrammaticality. Figures 5, 6 and 7 illustrate these concepts.



Figure 5: Mosesean vectors on a real number line showing –ve Grammaticality Ungrammaticality



Figure 6: Resultants of Mosesean vectors not yet at equilibrium



1

### International Conference



Figure 7: Resultant of Mosesean vectors at equilibrium

#### 3.2 Resultants of Mosesean Vectors

The default value of zero (o) is assigned as gradience value to any sentence introduced into the proposed system of Mosesean vectors. At this default value, the sentence is at equilibrium, and is completely grammatical. This equilibrium is toppled when ungrammatical elements are identified within the sentence. When a sentence is ungrammatical by a certain magnitude, the proposed system attempts to find complementary grammaticality measures to pull the sentence back into equilibrium as depicted in Figure 5. Grammaticality (+g) is generated in response to Ungrammaticality (-g). +g can only be as large as to cancel out -g, thus bringing the system into equilibrium once again.

#### 3.3 Derivation of Grammaticality Gradience **Equations Using Mosesean Vectors**

The default value of zero (0) is assigned as gradience value to any sentence introduced into the proposed system of Mosesean vectors. If the sentence is grammatically correct, the gradience remains unchanged at zero, and requires no further computation. However, if ungrammaticality (-g) is found within the sentence, the cumulative ungrammaticality (-G<sub>sum</sub>) is the arithmetic sum of the individual ungrammaticality values.

Thus:

if the values assigned to ungrammaticality items in a sentence are  $-g_1, -g_2, -g_3, \dots -g_n$ 

then

 $-G_{sum} = (-g_1) + (-g_2) + (-g_3) + \dots + (-g_n)$ 

therefore

$$G_{sum} = \begin{bmatrix} n \\ |g_i| \\ i = 1 \end{bmatrix}$$

Hence, the cumulative grammaticality (G<sub>sum</sub>) required to bring the sentence into equilibrium is:

$$G_{sum} = \frac{|g_i|}{i=1}$$
 2

Furthermore, the grammaticality gradience of an ungrammatical sentence is computed by dividing the cumulative grammaticality (G<sub>sum</sub>) by the total number of leaf nodes in the sentence parse tree. The number of leaf nodes in a sentence parse tree is equal to the number of words in the sentence. It follows that the grammaticality therefore gradience ( $\mu$ G) for a sentence with word-length m is:

$$\mu G = \underbrace{ \begin{vmatrix} g_i \\ i = 1 \end{vmatrix}}_{m} (3)$$

In addition, this model assigns a value of minus one (-1) to each ungrammatical item in a sentence. This follows that the cumulative grammaticality  $(G_{sum})$ can at most be as large as the number (m) of words in the sentence.

Thus in equation (3): when n 0:

∞:

m = m and µG

m

when n

Hence, the grammaticality gradience  $\mu G$  is such that:

0

As grammaticality evaluation extends beyond the evaluation of a single sentence to the evaluation of



multiple sentences, the gradience for each sentence is computed in the same manner, applying equation (3) to each of them. The arithmetic mean of the gradience(s) of the different sentences is then computed to give the gradience of all the sentences put together.

Thus: if there are q sentences with gradience values  $\mu G_1,\,\mu G_2,\,\mu G_3,\,...\,\,\mu G_q$ 

then

$$\mu G_{sum} = \mu G_{1} + \mu G_{2} + \mu G_{3} + \dots + \mu G_{q}$$

$$\mu G_{sum} = \prod_{i=1}^{q} \mu G_{i} \qquad (4)$$

and by extension, the grammaticality gradience  $\mu G$  for multiple sentences is:

$$\mu G = \frac{\underset{i=1}{\overset{q}{\mu}G_{i}}}{q}$$
(5)

### 3.4 Illustrations

To illustrate how the new model for grammaticality grading works, three illustrations are considered.

### Illustration 1: Grammatical sentence

Consider the sentence "The fat pony sleeps in the evaluate this sentence barn". То for grammaticality, HPSG (section 2.2.3) is employed in a bottom-up fashion as shown in the parse tree in section 2.2.4. As each lexical token in the sentence is parsed upward in the parse tree, it is replaced by its corresponding feature structure similar to those shown in section 2.2.5. For reasons of space, Figure 1 only shows the POS components of the respective feature structures for each lexical entry. Hence the grammaticality of the sentence is evaluated as discussed in sections 2.2.2 through 2.2.5. Since the sentence was successfully parsed all the way to the topmost root node (S in Figure 1), and no error (n=o) was identified during the parse process. The sentence is therefore considered grammatical. To compute the grammaticality gradience for this particular sentence:

Using equation (3)

where n = o and m = 7 $\mu G = o/7$ 

### therefore

µG = 0

which is the value expected for a grammatical sentence.

Illustration 2: Ungrammatical Sentence

Consider the following sentence "I loves Sandy". The simplified feature structure for each lexical entry is shown in Figure 8. Figure 9 shows the agreements expected of the lexical entries. The numbers in square bracket (e.g. [1]) show what attributes should agree. Comparing the agreement between "I" and "loves", a subject-verb error was observed (n = 1). "I" is first (1) person singular (*sg*) and therefore expects a first person singular *verb*, but rather it gets a third (3) person singular verb "likes" and vice versa. The ARG-ST feature for the *verb* "loves"

< loves, [ARG-ST < [NP [AGR 3s]], NP >] >

shows that the third lexical entry "Sandy" is in agreement with the object expected by the *verb*.

Thus, to compute the grammaticality gradience for this particular sentence:

Using equation (2)

where 
$$n = 1$$
 and  $g = 1$   
 $G_{sum} = 1$   
Using equation (3)  
where  $n = 1$  and  $m = 3$   
 $\mu G = 1/3$ 

therefore

µG = 0.3333

which is within the range of values expected.

It should be noted that the feature structures shown in Figures 8 and 9 are highly simplified.

### Illustration 3: Multiple Sentences

Finally, consider a text that consists of both sentences in illustration 1 and 2. That is "The fat pony sleeps in the barn. I loves Sandy". To compute the overall grammaticality gradience for the text, the gradience for the individual sentences is computed as has already been done in illustrations 1 and 2. Equation (5) is then used to compute the overall grammaticality of the text:

Using equation (5)



- where q = 2and  $\mu G_i = \{0, 0.3333\}$ (0 + 0.3333)/2 $\mu G =$ therefore µG = 0.1667 POS pronoun NUM sg AGR PERS 1 (a) POS verb NUM sg AGR PERS 3 (b)
  - $\begin{bmatrix} POS & noun \\ AGR & \begin{bmatrix} NUM & sg \\ PERS & 3 \end{bmatrix}$





Figure 9: Parse tree for the sentence "I loves Sandy" highlighting the ungrammaticality weight.

### 3.4 Result

The study showed that using the newly formulated model, grammaticality gradience µG in handcrafted grammar can be computed such that  $0 \le \mu G \le 1$ . This implies that the new model evaluates grammaticality gradience on a graded scale with values ranging between o and 1. Where the grammaticality gradience value of o is assigned to a grammatical sentence perfectly and the grammaticality gradience value of 1 is assigned to a completely ungrammatical expression. Therefore, the new model is designed such that, the higher the grammaticality gradience of an expression, the more ungrammatical the expression is.

#### 4. CONCLUSION

This paper proposes the adoption of the formulated model for the computation of grammaticality gradience in natural language text. Subsequent phases of this research will focus on applying the formulated model in an automated grammaticality evaluation and error detection system. Such automated implementation will facilitate testing across relatively large corpus so as to further validate the effectiveness of this approach.

### ACKNOWLEDGEMENT

We hereby acknowledge and appreciate the Nigeria Computer Society (NCS) for the Postgraduate Scholarship awarded Mr. Aregbesola M. K. in pursuance of this research work.

### REFERENCES

Baldridge, J., Chatterjee, S., Palmer, A., & Wing, B., 2007. DotCCG and VisCCG: Wiki and Programming paradigms for Improved Grammar Engineering with OpenCCG. Proceedings of the Grammar Engineering Across Frameworks (GEAF07) Workshop, pp. 5-25.

Bialystok, E., 1979. Explicit and implicit judgments of L2 grammaticality. *Language Learning*, 29(1): pp. 69-83.



Bialystok, E., 1981. The role of linguistic knowledge in second language use. Studies in Second Language Acquisition, 4(1): pp. 31-45.

Blache, P., 2005. Property Grammars: A Fully Constraint-Based Theory. in H. Christiansen & al. (eds), Constraint Solving and NLP, Lecture Notes in Computer Science, Springer.

Blache, P., 2006. A Robust and Efficient Parser for Non-Canonical Inputs. In proceedings of Robust Methods in Analysis of Natural Language Data, EACL workshop.

Blache, P., Hemforth, B., & Rauzy, S., 2005. Acceptability Prediction by Means of Grammaticality Quantification. Dec., COLING-ACL, <hal-00135422>, pp.57-64.

Bresnan, J., 1982. Control and complementation. In: J. Bresnan (ed.): The mental representation of grammatical relations. Cambridge, pp. 282-390.

Carnie, A., 2013. Syntax: A Generative Introduction. Third Edition. John Wiley & Sons, Inc.

Chomsky, N., 1957. Syntactic structures. Janua linguarum: Series minor 4, The Hague, Netherlands: Mouton.

Chomsky, N., 1975. The Logical Structure of Linguistic Theory. Plenum Press.

Chomsky, N., 1981. Lectures on Government and Binding. Studies in Generative Grammar 9.

Chu, H., 2003. Information representation and retrieval in the digital age. Medford, N. J.: Information Today, Inc. Published for the American Society for Information Science and Technology.

Clark, A., Giorgolo, G., & Lappin, S., 2013. Statistical Representation of Grammaticality Judgements: the Limits of N-Gram Models. Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics, pp. 28-36.

Coecke, B., Grefenstette, E., & Sadrzadeh, M. 2013. Lambek vs. Lambek: Functorial Vector Space Semantics and String Diagrams for Lambek Calculus. Department of Computer Science, University of Oxford.

Crocker, M. W., & Keller, F., 2006. Probabilistic grammars as models of gradience in language processing. In Gisbert Fanselow, Caroline F'ery, R. V. and Schlesewsky, M., editors, Gradience in Grammar: Generative Perspectives. Oxford University Press Inc., New York, USA., pp. 227–245. Fillmore, C., 1998. Inversion and Contructional Inheritance. In Lexical and Constructional Aspects of Linguistic Explanation, Stanford University.

GASS, S., 1983. The Development of L2 Intuitions. The University of Michigan. TESOL QUARTERLY, Vol, 17, No.2.

Gazdar, G., 1981. Unbounded Dependency and Coordinate Structure. Linguistic Inquiry, 12, 2 - pp. 155-184.

Gazdar, G., Klein, E., Pullum, G., & Sag, I., 1985. Generalized Phrase Structure Grammar. Cambridge, Mass., Harvard University Press.

Harlow, S., 2010. Background to Head-driven Phrase Structure Grammar. Retrieved 26-12-2014 from www-

users.york.ac.uk/~sjh1/hpsgintroslides.pdf

Johnson, M., 1988. Attribute-Value Logic and the Theory of Grammar. Stanford CLSI, CSLI Lecture Notes 16.

Kambites, M., 2014. Vectors and Matrices. Lecture notes, School of Mathematics, University of Manchester. Retrieved 05/01/2015 from http://www.maths.manchester.ac.uk/~mkambites/t eaching/cs/lectures.pdf.

Karttunen, L., 1986. *Radical Lexicalism*. Artificial Intelligence Center at SRI and Center for the Study of Language and Information, Stanford.

Kay, M., 1979. Functional Grammar. In Proceedings of the fifth Annual Meeting of the Berkeley Linguistics Society. Berkeley, California.

Kay, M., 1982. Syntactic Models: Parsing in functional unification grammar. Academia.edu. Retrieved 07/01/2015 from https://www.academia.edu

/2803506/Parsing\_in\_functional\_unification\_gramm ar

Kay, P., & Fillmore, C., 1999. Grammatical Constructions and Linguistic Generalizations: the what's x doing y construction. Language.

Keller, F., 2000. Gradience in Grammar. Experimental and Computational Aspects of Degrees of Grammaticality. Phd Thesis, University of Edinburgh.

Kondratieva, M., 2008. Linear vector spaces and subspaces. Memorial University. Retrieved 06/01/2015 from



http://www.math.mun.ca/~mkondra/ linalg2/la2lec1.pdf.

Krishnamurthy, J., & Mitchell, T. M., 2013. Vector Space Semantic Parsing: A Framework for Compositional Vector Space Models. In Proceedings of the Workshop on Continuous Vector Space Models and their Compositionality, Sofia, Bulgaria, pp. 1-10.

Lau, J. H., Clark, A., & Lappin, S., 2014. Measuring Gradience in Speakers' Grammaticality Judgements. In Proceedings of the 36th Annual Conference of the Cognitive Science Society, Quebec City, Canada, pp. 821-826.

Lin, C. J., 2007. Grammaticality and Parsability in Sentence Processing. International Workshop on Grammar and Evidence, Chiayi, pp. 13-15.

Lindeman, R. W., 2008. Computer Graphics: Points, Scalars and Vectors. Lecture Notes, Department of Computer Science, Worcester Polytechnic Institute.

Luenberger, D. G., 1969. *Optimization by Vector Space Methods*. Stanford University, Stanford, California. Retrieved on o6/01/2015 from http://math.oregonstate.edu/~show/old/142\_Luenb erger.pdf, John Wiley & Sons, Inc.

MacKinlay, A., 2012. Pushing the Boundaries of Deep Parsing. PhD Thesis submitted to the Department of Computing and Information Systems, The University of Melbourne, Melbourne, Australia.

Malchow, A., 1992. Syntax and Semantics of Construction. MA Thesis, Fachgruppe Sprachwissenschaft, Universität Konstanz.

Menzel, W., & Schroder, I., 1998. Decision procedures for dependency parsing using graded constraints. In S. Kahane & A. Polgu`ere (eds), Proc. ColingACL Workshop on Processing of Dependencybased Grammars.

Müller, S., 2013. German Grammar and General Linguistics. Symposium: The Semantics of Action, Göttingen, pp 1-25.

Nerbonne, J., 2002. Computer-assisted language learning and natural language processing. In Mitkov, R., editor, Handbook of Computational Linguistics, pp. 670–698, Oxford University Press.

Nowak, R., 2010. Review of Linear Algebra. Lecture Note: ECE 830 Fall 2010 Statistical Signal Processing. Lecture 3. Parfitt, S., Tiilo, P., & Dorffner, G., 1999. Graded grammaticality in Prediction Fractal Machines. *In Proceedings of NIPS'*99, pp. 52-58.

Pauls, A., & Klein, D., 2012. Large-scale syntactic language modeling with treelets. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics, Jeju, Korea pp. 959–968.

Pennington, J., Socher, R., & Manning, C. D., 2014. *GloVe: Global Vectors for Word Representation*. Computer Science Department, Stanford University, Stanford, CA 94305.

Peters, P. S., & Ritchie, R. W., 1973. On the Generative Power of Transformational Grammars. Information Science, 6: pp.49-83.

Pollard, C., 1984. Generalized Phrase Structure Grammars, Head Grammars, and Natural Languages. Dissertation, Stanford University.

Pollard, C., & Sag, I. A., 1987. Information-Based Syntax and Semantics. Fundamentals Vol. 1. Lecture Notes 13. Stanford.

Rapp, R., 2003. Word sense discovery based on sense descriptor dissimilarity. In Proceedings of the Ninth Machine Translation Summit.

Sampson, G. R., 2007. *Grammar without grammaticality*. Corpus Linguistics and Linguistic Theory, 3(1): pp. 1–32.

Schmid, H., 2004. Efficient parsing of highly ambiguous context-free grammars with bit vectors. In Proceedings of the 18th International Conference on Computational Linguistics. Geneva, Switzerland. COLING, pp. 162–168.

Schroder, I., 2002. Natural Language Parsing with Graded Constraints. PhD Thesis, University of Hamburg.

Socher, R., Huval, B., Manning, C. D., & Ng, A. Y., 2012. Semantic Compositionality Through Recursive Matrix-Vector Spaces. In Proceedings of the 2012 Conference on Empirical Methods in Natural Language Processing (EMNLP).

Socher, R., Pennington, J., Huang, E. H., Andrew Y. Ng, & Manning, C. D., 2011. Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP).

Sorace, A., & Keller, F., 2005. Gradience in Linguistic Data. Lingua, pp. 115.



Stolcke, A., 1991. Syntactic Category Formation with Vector Space Grammars. In Proceedings of 13th annual conference of Cognitive Science Society, Chicago.

Taylor, J. R., 1995. Linguistic Categorization: Prototypes in linguistic theory. Claren don, Oxford.

Turney, P. D., 2006. Similarity of semantic relations. Computational Linguistics, 32(3).

Turney, P. D., & Pantel, P., 2010. From frequency to meaning: vector space models of semantics. *Journal of Artificial Intelligence Research*, 37(1).

Uszkoreit, H., 1986. Categorial Unification Grammars. In Proceedings of the International Conference on Computational Linguistics (COLIN), pp. 187-194.

van Hassel, R., 2009. Lecture Notes on Functional Analysis. Retrieved 06/01/2015 from http://www.win.tue.nl/~rvhassel/Onderwijs/2DE08-0809/OwnLectureNotesFA.pdf.

Wagner, J., 2012. Detecting Grammaticality Errors with Treebank-Induced, Probabilistic Parsers. PhD Thesis, School of Computing, Dublin City University.

Witbrock, M. J., & Mittal, V. O., 1999. Ultrasummarization: A statistical approach to generating highly condensed non-extractive summaries. In Proceedings of the 22nd International Conference on Research and Development in Information Retrieval.

Zybura, K., 2004. *Linear Vector Space*. Retrieved o6/01/2015 from http://pirate.shu.edu/~wachsmut/ Teaching/MATH3912/Projects/papers/zybura\_vector \_spaces.pdf.



### PRIVACY-ENHANCING BIG DATA COMPUTING IN FINANCIAL INSTITUTIONS

### <sup>1</sup>C.O. Oloruntoba and <sup>2</sup>A.S. Sodiya

<sup>1,2</sup>Department of Computer Science, Federal University of Agriculture, Abeokuta

scientistore@yahoo.com<sup>1</sup> and sodiyaas@funaab.edu.ng<sup>2</sup>

### ABSTRACT

Big Data, which are high-volume, high-velocity, high-variety and highly complex information assets, has a lot of potential benefits in financial institutions. To compete in a consumer-empowered economy, it is pertinent that financial institutions use their information assets to acquire understanding of customers, markets, products, competitors, suppliers, channels etc. The extraordinary benefits of Big Data are however hampered by concerns of privacy. The issue of privacy has raised serious concerns in Big Data, thus financial institutions need to constantly review the analysis of large datasets of potentially sensitive private information about individuals. This paper introduces the Big Data concept, characteristics, analytics, sources, benefits and its challenges in financial institutions. A framework for enhancing privacy of financial Big Data in the cloud is presented in this work. The analysis of the framework shows a promising system for enhancing privacy of financial Big Data.

Keywords: Big Data, Cloud Computing, Differential Privacy, Financial Institutions, Laplace Noise

### 1. INTRODUCTION

Data has become a driving force behind almost every interaction between individuals, businesses, and governments. Across manysectors, including banking, financial services, energy, healthcare, media,education among others, the exponential automation of business processes is enlarging the scope of data usage and analytics.Data once stored in different offline repositories in a variety of formats, are now available in digital format, ready to be correlated and analyzed in huge chunks of terabytes, petabytes, and zettabytes in real time, leading to the rise of Big Data. (Laney, 2001) describedBig Dataas data sets with three aspects: Volume, Velocity and Variety, whileGartner, (2014), describedBig Dataas highvolume, high-velocity, high-variety and highly complex information assets that demand costeffective, innovative forms of information processing for enhanced insight and decision making. Also, Big Data can be defined as data that exceeds the processing capability of conventional database systems. The data istoo big, moves too fast, or does not fit the structures ofconventional database architectures (Dumbill, 2013).

BigData possesses characteristics which can be described by '6Vs'. They are: Volume, Velocity,



Variety, Value, Variability and Veracity (Russom, 2011; Eaton et al., 2012; (O'Reilly, 2012; Zikopoulos et al, 2012; Bellini et al, 2013; Demchenko et al, 2013; Megahed and Jones-Farmer, 2013; Minelli et al, 2013; Rajpathak and Narsingpurkar, 2013).

a.Volume: The quantity of data that is generated is very important in Big Data. Big Data refers to data size such as Terabytes (TB: approximately 10<sup>12</sup> bytes), Petabytes (PB: approximately 10<sup>15</sup> bytes)and Zettabytes (ZB: approximately 10<sup>21</sup> bytes), etc.

b.Velocity: This refers to the speed of generation of data.

c. Variety: This represents all types of data, such as structured data from relational tables and unstructured data from email messages, texts, streamed video and audio, etc.

d.Value: It is defined by the added-value that the collected data can bring.

e. Variability:This refers to the inconsistence that can be exhibited by the data. Increasing variety andvariability of data increases thepotentiality in providing unexpected, hiddenand valuable information.

f. Veracity: This refers to the varying quality of the data being captured.

Big Data creates enormous value for the global economy, driving innovations, productivity, efficiency and growth. Big Data represents a trend in technology that is leading the way to a new approach in understanding the world and making business decisions (ISACA.org, 2013). It is currently building great relevance in areas such as scientific and medical research, credit risk analysis, security, marketing, management, etc.

Big Data collected requires adequate analysis as the real value of Big Data is the insights it produces when analyzed. Big Data Analytics is the process of examining Big Data to reveal hidden patterns, unknown correlations and other useful information that can be used to make better decisions. With Big Data analytics, data scientists can analyze huge volumes of data that conventional analytics cannot do. High-performance analytics is necessary to process Big Data and as such, using high-performance data mining, predictive analytics, text mining, forecasting and optimization on Big Data guarantees maximum outcomes.

Many organizations seeking to collect, process and analyze Big Data make use of technologies that includes Apache Hadoop and related tools such as MapReduce, YARN, Hive and Pig. Hadoop is an open source framework for storing and processing large and diverse datasets across distributed systems. Hadoop has two primary components, namely, Hadoop Distributed File System and MapReduce programming framework.Hadoop Distributed File System (HDFS) (Borthakur, 2008) is a distributed file system that is used to storedata across cluster of commodity machines while providing high availability andfault tolerance.

The programming model used in Hadoop is MapReduce (Dean and Ghemawat, 2008) which was proposed byDean and Ghemawat at Google. MapReduce is the basic data processing scheme usedin Hadoop which includes two components, the Mappers and Reducers. Mappers read the data from HDFS, process it andgenerate some intermediate results to the reducers. Reducers are used to aggregate intermediate results to generate the final output which is subsequently written to HDFS.A typical Hadoop job involves running several mappers and reducers across differentnodes in the cluster.

Apache Pig and Hive are set of wrappers developed for MapReduce. MapReduce wrappers provide a better control over the MapReduce code and aid in the source code development. These wrappers provide a better environment and make the code development simpler as the programmers do not have to deal with the complexities of MapReduce coding. YARN (Vavilapalli et al, 2013) is a resource management layer and schedules the jobs across the distributed systems.

Cloud computing has become a powerful architecture to perform large-scale and complex computing. The advantages of cloud computing include virtualized resources, parallel processing, data service integration with scalable data storage, etc. Organizations are looking up to cloud computing to



support their Big Data projects. Big Data environments require clusters of servers to support the tools that process and analyze the big data. Cloud computing offers a cost-effective way to support big data technologies and the advanced analytics applications that can drive business value.Some of the first adopters of Big Data in cloud computing are users that deployed Hadoop clusters in highly scalable and elastic computing environments provided by vendors, such as IBM, Microsoft Azure, and Amazon AWS (Chang et al, 2013).

Big Data and Cloud computing are interrelated.As Big Data provides the ability to process multiple datasets, cloud computing provides the underlying engine through the use of Hadoop. Big Data utilizes distributed storage technology based on cloud computing rather than local storage attached to a computer or electronic device. Cloud computing infrastructure serves as an effective platform for data storage required to perform Big Data analysis. MapReduce is used for Big Data processing in a cloud environment.

In this paper, Big Data challenges are highlighted with focus on Privacy.Privacy poses a major threat even as the capabilities and potential applications of Big Data computing are evolving rapidly in financial institutions. The goal of this paper therefore, is to present a framework useful for enhancing privacy of financial big data in the cloud. The sources of Big Data and its benefits in financial institutions were also discussed.

The rest of this paper is organized as follows: Section 2.0 presents an overview of Big Data in financial institutions. The methodology and architecture for enhancing privacy of financial big data are presented in Section 3.0.The analysis of the proposed framework is presented in Section 4.0 while the paper concludes in Section 5.0.

### 2. BIG DATA IN FINANCIAL INSTITUTIONS

Big Data is especially promising for financial institutions as data is one of its important assets. In banking and financial management, hundreds of millions of transactions are conducted daily, each adding another row to the sectors'huge and growing ocean of data. Log data which are machine-generated data, produced to record the details of every operational transaction and automated function performed within the bank's business or information systems, exists and requires adequate analysis. Such internal data has outgrown the ability to be stored and analyzed by traditional systems. In many cases, these data have been collected for years without been analyzed. Many of these institutions are curious onhow to use these data to gain a competitive advantage.

Research shows that 71% of these banking and financial markets report that the use of big data and its analytics is creating a competitive advantage for their organizations, compared with 63% of cross-industry respondents. Compared to 36% of banking and financial markets that reported an advantage in IBM's 2010 New Intelligent Enterprise Global Executive Study and Research Collaboration, this is a 97% increase in two years as shown in figure 1 (LaValle et al, 2010).



**Figure 1:** competitive advantage of financial institutions over their cross-industry counterparts using Big Data and Analytics



Major sources of BigData for financial institutions include internal sources that produce structured data and social mediathat produces unstructureddata. Structured data can refer to data that currently resides in a product and customer-servicing systems. Unstructured data can refer to data not easily captured or stored, such as social media data, voice call logs, emails, website click streams, video files, etc. Figure 2 shows the classifications of Big Data sources, comparing sources for banking and financial markets with their cross-industry counterparts.



**Figure 2:** Sources of Big Data for banking and financial markets.



### 2.1Benefits of Big Data in Financial Institutions

The introduction of Big Data has enabled the financial industry to access more data than before. Previously, unstructured data had no value, but with Big Data, it can be collected and analyzed for the benefit of the industry.Banks and insurance agencies have access to more data than ever before and they must determine how to use this data to deepen the customer relationship (World Retail Banking Report, 2013). Benefits of Big Data in financial institutions include:

Customer-centric Outcomes: a. In financial institutions, customers are the focus around which data insights, technology, operations and systems revolve. Customer analytics will enable better service to customers due to the ability to understand customer needs and anticipate future behaviours. By improving the ability to understand changing markets conditions and customer preferences through the analysis of unstructured data from social media, sentiments and complains of customers can be uncovered which will induce the creation of new customer-centric products and services. This improves customer service leading to higher customer retention.

b.Business-centric Outcomes: Big Data can simultaneously reduce costs and increase revenues, a duality that will give the business a boost.Financial Institutions can use customer insights derived from big data to design marketing strategies, execute campaigns and capture sales leads across all channels, product lines and customer segments. By leveraging innovative business models enabled by Big Data, investment managers, sales teams, and operation teams have the ability to predict market changes and forecast sales.

c.Enhanced Risk and Financial Management: Predictive analytics enhanced by Big Data allows the analysis of data to combat fraud, track anti-money laundering, cash management and alleviate operational risk.

d. Improved Decision Making: For important business decisions, Big Data provides the availability of a wealth of data that can influence timely and quality decision making with confidence. Relationship managers, investment bankers, financial advisors, loan officers and others can have access to data for better and more informed decisions.

e.Improved Accurate Data: Big Data offers the financial industry the opportunity to discover data correlations and patterns that would have remained hidden, which enables theindustry to have access to more accurate information.

#### 2.2Financial Big Data in the Cloud

By developing a comprehensive Cloud-based Big Data strategy, optimized value can be derived from Big Data. In financial institutions, as customers' data, which can be of individuals, organizations or enterprises are analyzed, cloud computing models can help accelerate the potential for scalable analytics solutions. Analytics as a Service (AaaS) can be deployed in the cloud based on various cloud service models. Infrastructure as a Service (IaaS), Platform as a service (PaaS), and Software as a Service (SaaS) are basic cloud service models for Analytics as a Service.

laaS enables the allocation of shared server resources, which are often virtualized to handle the computing and storage needs for Big Data analytics. Cloud operating systems manage high-performance servers, network, and storage resources. PaaS provides tools and libraries to build, test, deploy, and run applications on cloud infrastructure. PaaS serves as a development platform for advanced analytics applications. SaaS can provide specific applications for cloud-based big data analytics. SaaS can be referred to as a standalone application or part of a greater cloud provider solution.

### 2.3 Challenges of Big Data in Financial Institutions

In a financial institution, the usage of big data on a daily basis and in real time encounters challenges, such as:



- a. Data Privacy
- b. Security of Data
- c. Storage of Data
- d, Analysis of large data sets
- e. Identity and Access management
- f. Lack of IT infrastructure, skills and expertise

### 2.4Big Data Privacy Challenge

The ability of an individual or group to seclude themselves, or information about themselves, and thereby express themselves selectively can be referred to as Privacy. Concerns over privacy are causing drawbacks in the benefits of Big Data computing. The ease and efficiency of cloud computing comes with privacy and security risks (Kaufman, 2009). Privacy of data is the main hurdle in implementation of cloud services. Creating an appropriate balance between privacy challenge and big data rewards may very well be the biggest public policy issue in financial institutions (Ira, 2013). The growth of Big Data has influenced different repository storage of personally identifiable credit card data details and transaction data in financial institutions. The storage and analysis of such data have increased the demands of privacy. A breach in privacy will be detrimental to a customer and the institution's overall brand, including a loss in revenue from decreased business or from regulatory fines. Since privacy is critical, government agencies heavily regulate the use of personal information for financial services and insurance companies.

Three broad categories of Big Data privacy threats can be conceived, namely,Surveillance, Disclosure, and Discrimination. Surveillance regards the feeling of being watched resulting from the collection, aggregation, and/or use of one's information (Ryan, 2011). Another threatarises due to the disclosure of data beyond the entity that initially collected it. Other big data privacy problems include discrimination, which is treating people differently on the basis of information collected about them.

Preserving privacy of Big Data is challenging due to the following:

a) Massive increase of transaction data volumes

b) Rise of new forms of interaction data, such as social media

c) Increasingly complex IT environment

d) Use of insecure java-based frameworks, such as Hadoop and its programming paradigm MapReduce.

### 3. BIG DATA PRIVACY FRAMEWORK

Driven by the need to solve privacy issues of BigData in financial institutions, a privacy-enhancing framework was proposed.

### 3.1 Noise Addition

Noise addition perturbation methods can transform data by adding noise to preserve privacy. The noise addition mechanism works by adding stochastic or randomized attributes to confidential data. The stochastic value is selected from a normal distribution with zero mean and a diminutive standard deviation. With stochastic noise, random data is added to confidential attributes to hide the distinguishing values. The added noise is targeted at the following: a)Anonimizing: Hiding the customer.

- b) Hashing: Disguising a customers' identity.
- c) Cloaking: Making a customer invisible.
- d) Blurring: Decreasing accuracy of a query
- e) Lying: Intentionally giving false information.

### 3.2Differential Privacy

With differential privacy, analysts can extract useful information from datasets containing personal information while offering strong individual privacy protections (Salido, 2012)(Heffetz and Ligett, 2013). Differential privacy is a promising method for preserving privacy in Big Data.It aims to minimize the chances of individual identification while analyzing data.

Formally, Differential Privacy is defined as follows:A randomized function A gives  $\varepsilon$ -differential privacy if for all datasets  $D_1$  and  $D_2$ that differs on a single element (i.e., data of one person), and all S  $\subseteq$ Range(A),

 $\Pr[\mathcal{A}(D_1) \in S] \le e^* \times \Pr[\mathcal{A}(D_2) \in S]$ 



Since differential privacy is a probabilistic concept, any differentially private mechanism is necessarily random. We consider the Laplace mechanism that relies on adding controlled noise. The Laplace mechanism adds Laplace noise (i.e. noise from the Laplace distribution), which can be expressed by probability density function.

noise(y)  $\propto \exp(-|y|/\lambda)$ \_ 2which has a mean of zero and standard deviation  $\lambda$ .

The output function of A is defined as a real valued function (called the transcript output by A)as:

where  $Y \propto Lap(\lambda)$  and f is the original real valued function to be executed on the datasets.

 $\mathcal{T}_{\mathcal{A}}(\mathcal{A})$  can be onsidered to be a continuous random variable, where

$\mathrm{pdf}(\mathcal{T}_{\mathcal{A},D_1}(x)=t)$	$\operatorname{noise}(t - f(D_1))$	
$\operatorname{pdf}(\mathcal{T}_{\mathcal{A}, \mathcal{D}_2}(x) = t) =$	$noise(t - f(D_2))$	4

which is at most,

$$e^{|f(D_1) - f(D_2)|} \le e^{\frac{\Delta(f)}{\lambda}} - \frac{5}{5}$$

### 3.3 Proposed Privacy-EnhancingFramework

We can therefore consider to be the privacy factor<sub>ε</sub>.

 $\Delta(f)$ 



The proposed framework combines three concepts, Big data, Cloud computing and Differential privacy. The architecture is shown in figure 3.



Figure3: Privacy-Enhancing Big Data System Architecture

The Cloud provides facilities for the computation and processing of Big Data. Financial Big Data from several sources are analyzed with Hadoop using its programming model, MapReduce and HDFS. The outcomes of the analytics are stored in specific data centers in the cloud. Laplace noise distribution is calculated and added to the analytics to provide enough perturbation such that queries sensitive enough to reveal a customer's data will give responses that hides the individual.

Considering some zettabytes of data,  $\Delta f$  is calculated. A small  $\varepsilon$ value is chosen. Laplace noise is set to  $\Delta f/\varepsilon$ . Thus we arriveat Laplace (o,  $\Delta f/\varepsilon$ ) noise distribution.

Therefore,

Big data analytics plus Noise from Laplace (0,  $\Delta f/\epsilon$ ) is equal to Perturbed datasets.

Research concludes that smaller  $\epsilon$  epsilon value creates a more enhanced privacy.

### 4. ANALYSIS OF THE FRAMEWORK

Consider a large dataset that is a collection of rows. For instance, a row might be a customer's data and transactions. On query f, the customer's data di of the data in row i of the large dataset, computes f (di), i = 1, ..., n. The output bits are represented as 0/1 values in Af (D), for a large dataset D. We denote this set  $\{0, 1\}_{Af(D)}$ , where Af (D) is the perturbed datasets.

For a set of bits  $d_{i,1}$ ,  $d_{i,2}$ ..... $d_{i,n}$ , Af(D) = oif  $\forall i \in \{1, ..., n\}$ ,  $\exists j_1', j_2'$  such that: A(Di,di $\forall i$ ) = Di,j'di  $\forall i$ .



Thus, Di,j′di = Di,di +noise ∀ i.

The framework can be analysed based on two

- measures, namely,
  - a. Privacy Invasion
  - b. Effect of noise

Privacy Invasion: By acquiring a perturbed version of datasets, Af(D) = 0.

Effect of Noise: The privacy factor  $\varepsilon$  in differential privacy usually takes a small value to give an appropriate measure of noise as more noise renders the data useless.

### **5.CONCLUSION AND FUTURE WORK**

As financial institutions create and store more transactional data and logs, big data analysis can be conducted on such data to derive value that gives a competitive advantage. Big Data analytics can perform predictive modelling to detect fraud, track anti-money laundering and as such enhance risk management. The potential benefits of Big Data are however hampered by challenges that include data privacy, data security, processing and storage of data, lack of infrastructure, expertise and skills, etc. Sensitivities around privacy are a big hurdle that financial institutions need to overcome if the economic benefits of Big Data are to be realized. We focused on noise addition as a perturbation methodology to transform large datasets for enhancing privacy. Future work should be targeted at the evaluative results of this framework.

### 6. **REFERENCES**

- Bellini, P. et al, (2013). Tassonomy and Review of Big Data Solutions Navigation. *In Big Data Computing*, Akerkar, R. (Ed.), Chapman and Hall/CRC, pp 57-101.
- Borthakur, D., (2008). HDFS architecture guide. Hadoop Apache Project. http://hadoop.apache.org/docs/r1.2.1/ hdfs\_design.pdf.
- Chang, L. et al,(2013): Public Auditing for Big Data Storage in Cloud Computing – a Survey, In 2013 IEEE 16th International Conference, pp 1128–1135.

Hence, if Af(D) = 0, the privacy of the big dataanalyticshasbeenenhanced.

- Dean, J. and Ghemawat, S., 2008.MapReduce: Simplified data processing on large clusters. *Commun ACM*, Vol. 51, No.1, pp 107–113.
- Demchenko, Y. et al, 2013. Addressing big data issues in scientific data infrastructure. Proceedings of the International Conference on Collaboration Technologies and Systems, IEEE Xplore Press, USA, pp 48-55. DOI: 10.1109/CTS.2013.6567203.

Dumbill, E., 2013. Making sense of Big Data. Big Data.

- Eaton, C. et al,(2012):Understanding Big Data. McGraw-Hill Companies.
- Gartner IT glossary Big data, (2014) June 25. Accessed from <u>http://www.gartner.com/it-glossary/big-</u> <u>data</u>
- Heffetz, O. and Ligett, K., (2013): Privacy and databased research. NBER Working Paper.
- Ira, R., (2013): Big Data: The End of Privacy or a New Beginning?.3 INT'L DATA PRIVACY L.74, pp 77-78.
- ISACA.org. Big Data: Impacts and Benefits, 2013.Retrieved from<u>www.isaca.org/Knowledge-Center/Research/ResearchDeliverables/Pages/Bi</u> g-Data-Impacts-and-Benefits.aspx
- Kaufman, L.,(2009). Data Security in the World of Cloud Computing. *IEEE Security and Privacy*, Vol. 7, No. 4, pp 61-64.
- Laney, D., (2001):3D Data Management: Controlling Data Volume, Velocity and Variety. Retrieved from<u>www.blogs.gartner.com/doug-</u> laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocityand-Variety.pdf
- LaValle, et al,(2010):Analytics: The new path to value: How the smartest organizations are



embedding analytics to transform insights into Value in collaboration with MIT Sloan Management Review. Retrieved from http://www-935.ibm.com/ services/us/gbs/thoughtleadership/ibvembeddinganalytics.html.Massachusetts Institute for Technology.

- Megahed, F. and Jones-Farmer, L.,(2013):A Statistical Process Monitoring Perspective on Big Data. Frontiers in Statistical Quality Control. 11<sup>th</sup> Ed., Springer, New York.
- Minelli, M. et al, (2013). Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Business. 1st Ed., John Wiley & Sons, Inc. Hoboken, New Jersey, USA, ISBN-10: 111814760X, pp 187.
- O'Reilly,(2012). Planning for big data. O'Reilly Radar Team.
- Rajpathak, T. and Narsingpurkar, A., (2013): Managing knowledge from big data analytics in product development. White Paper, Tata Consultancy Services.
- Russom, P., (2011): Big data analytics. TDWI Best Practices Report, Fourth Quarter.
- Ryan Calo, M., (2011): The Boundaries of Privacy Harm, 86 IND. L.J. 1131, pp 1144–1147.
- Salido, J., (2012): "Differential privacy for everyone". White Paper, Microsoft Coporation.
- Vavilapalli, V. et al, (2013): Apache Hadoop Yarn: Yet another resource negotiator. InProceedings of the 4th annual Symposium on CloudComputing,pp5.

action. IBM Institute for Busir
---------------------------------

Zikopoulos, P. et al, (2012):Understanding big data: Analytics for enterprise class hadoop and streaming data. McGraw-Hill, New York.



## Facilitating SMEs for Inclusive Development: The Role of Service-oriented Software Engineering (SOSE)

### **Emmanuel Okewu**

Centre for Information Technology and Systems, University of Lagos, Lagos, Nigeria

okewue@yahoo.com

### ABSTRACT

Small and Medium Enterprises (SMEs) are mechanisms for inclusive development in Africa, breaking barriers of age, gender, social class, religion, and tribe. They are aimed at productively engaging minds for socio-economic growth and development through job creation and wealth generation. However, a major road block to SME sector is the investment climate in Africa. A less than conducive business environment means high cost of doing business, low access to finance, poor access to markets, and low capacity building. Undoubtedly, measures have been proposed to address these challenges facing the sector in view of its huge contribution to the gross domestic product (GDP) of African countries. Against the backdrop that information availability and accessibility is key to tackling these challenges, this work explored the use of Information and Communication Technology (ICT) for repositioning and retooling the sector. Design and engineering skills were applied to achieve the mandate - the study designed and implemented an n-tier enterprise system tagged Automated System for Informal Sector (ASIS) using Service-oriented Software Engineering (SOSE) approach. Requirements were gathered, proposed system modelled using Universal Modelling Language, and development done using Microsoft SharePoint. This paper reports on the promising results obtained so far. The outcome ostensibly indicates that a well articulated and coordinated online real-time information system is one of the measures that could serve as a game-changer in providing condusive business landscape. And ultimately, the envisaged SMEs growth and development would be guaranteed, strengthening inclusive development.

**Keywords:** Inclusive Development, SMEs, Service-Oriented Software Engineering, Automated System for Informal Sector, Socio-Economic Transformation

#### 1. INTRODUCTION

Many challenges are confronting SMEs growth and development in Africa: poor infrastructure, access to markets, access to finance, and unattractive investment climate occasioned by regulatory, institutional and legal barriers. The obstacles vary from country to country. The financing of small and medium enterprises (SMEs) has been a topic of keen interest in recent years because of the key role that SMEs play in economic development and their potentially important contribution to economic diversification and employment in various economies.



Nonetheless, a system of measures is being put in place to address the issues so that SMEs would continue to engage all particularly the underprivileged for socio-economic inclusion. One of such measures is a technology-based SME sector in which ICT is given the pride of place. Information availability, accessibility and reliability will boost investors' and financial institutions' confidence in the sector and the multiplier effects will result in all other indicators looking upward. A case in point is the ongoing Bank Verification Number (BVN) exercise in Nigeria. The initiative involves capturing biometric data of customers, among other data, with potential for providing banks with adequate information about customers under the aegis of KYC (Know Your Customer). No doubt, this will stimulate banks' confidence and propensity to lend.

This research study focused on the provision of ICT infrastructure that will facilitate access to markets, assist in capacity building for SMEs operators to have requisite operational and managerial skills, enhance access to finance, and provide online realtime data to galvanize regulatory, institutional and legal frameworks of the investment climate. Specifically, the study developed an enterprise application tagged Automated System for Informal Sector (ASIS) using Service-oriented Software Engineering (SOSE) approach with the expectation that the online real-time information it provides will fast-track socio-economic transformation.

The application of SOSE for developing an SME system is promising because of the reasons stated in (Pressman, 2009; Debayan, 2011; Sommerville, 2011; Crnkovic and Larsson, 2003; Breivold and Larsson, 2007).

The remainder of this paper comprises the following: Section 2 gives the background of study and related work; Section 3 presents the methodology and the selected case study; section 4 focuses on results and discussions; and finally, the paper is concluded in section 5.

### 2. BACKGROUND AND RELATED WORK

#### 2.1 Challenges of SMEs in Africa

The main challenges facing SMEs growth and development in Africa include poor infrastructure,

access to finance, access to markets, electricity deficit, and weak regulatory, institutional and legal frameworks. Earlier studies revealed that enabling environment lags far behind in Nigeria, Tanzania, and even in Kenya. However, studies have shown that the number of small established businesses is high (Global Entrepreneurship, 2012). Figure 1 that while established illustrates business ownership is relatively high for some African countries it is relatively average in others. It equally emphasizes that the percentage of established owned businesses in some African countries is higher than that of more developed economies such as China, Brazil, France, the United Kingdom (UK) and the United States of America (USA).



Figure 1: Established business ownership rate (%) (Source: Global Entrepreneurship Monitor Survey Data', May 2012, http://www.globalentrepreneurshipmonitor). Downloaded March, 2015.

However, the challenges confronting SMEs in Africa are impacting on the growth of start-ups. Hence, start-up expectations for growth remain low despite the large number of SMEs. Research has shown that start-ups in France, the UK and the USA show greater expectations for their business' growth than those in Africa as shown in figure 2.



Figure 2: Expectations for growth in next 5years (%) (Source: Global Entrepreneurship Monitor Survey



Data',

May

2012, http://www.globalentrepreneurshipmonitor). Downloaded March, 2015.

The macroeconomic environment in which SMEs operate in Africa is less than appropriate. The World Bank Doing Business Report named Chad the most difficult Sub-Saharan African country to conduct business in 2012 (Doing Business, 2012). High tax rate, difficult insolvency regulations, and long processing times make it difficult for SMEs to operate and make profit. Particularly, high insolvency costs make it extremely difficult for entrepreneurs to try again when their businesses fail. Other factors include absence of building permits issued to the private sector and the nightmare of accessing credit (Carstens, 2012; Bangudu, 2013). Furthermore, poor infrastructure, corruption, and unreliable electricity are contributory factors. For instance, Nigeria ranks poorly in the world for ease of doing business, despite Government policy that attempts to encourage entrepreneurship. Specifically in 1999, the Federal Government implemented a Small and Medium Enterprises Equity Investment Scheme (SMEEIS) which aimed to encourage a more productive SME sector. However, the policy objectives have not been met due to poor implementation (Mamman, et al., 2007). This has constrained SMEs from tapping into the market potential resulting from the country's young, large and growing middle class and population, thus slowing down economic growth and employment creation. However, while some countries are grappling with high levels of regulatory and institutional barriers, others have made significant improvements in easing the cost of doing business in their countries (Carstens, 2012). Indeed, the World Bank Doing Business Report 2012 shows that over the last few years, an impressive 78% of African countries undertook meaningful governmental regulatory reform as a means of improving the business climate and encouraging investment (Doing Business, 2012, Carstens, 2012).

Though some governments do more to facilitate and encourage local business development in African countries, access to credit for entrepreneurs remains the key problem for SME

development. The African Development Bank (AfDB) recently reported that only 20% of African SMEs had access to credit and that only 9% of the investments SMEs make are funded by a bank (Calice et al., 2012). This contrasts South America and the Caribbean where 44% of SMEs reported access to credit, and to Europe were 23% of SMEs' investments are financed through bank loans. As it were, it is traditional for number of start-ups and small businesses to fail, and a climate that permits failure positions an entrepreneur to learn from that failure and start afresh. It is in such an environment that innovation and success can most thrive (Les Miserables, 2012). Study also revealed that most entrepreneurs were 'serial entrepreneurs' in that they launched one or two other companies before achieving success (Ernst & Young, 2011).

Notwithstanding this difficultly, there are a number of indicators that suggest that the situation is improving (Calice et al., 2012).

### 2.2 Information Technology and SMEs Challenges

The fact that the ICT strategy is one of many strategies that is auspicious of taking SMEs to the next level underpins this research study. Against the background that the integration of ICTs (mobile systems, biometric systems, etc) into SME initiatives will promote socio-economic inclusion for inclusive development, this study embarked on the design and implementation of an automated system for the informal sector (ASIS). A critical component of an ICT infrastructure is the software. And to fast-track growth and development, information exchange between users and providers needs to be dynamic. In this light, we chose a software engineering approach that dynamically meets the information needs of SME stakeholders. This approach is the Service-oriented Software Engineering (SOSE) model of developing software.

### 2.3 Service-Oriented Software Engineering and Inclusive Development

Inclusion is an organizational practice and goal aimed at bringing together different groups or individuals having different backgrounds like origin, age, race and ethnicity, religion, gender, sexual orientation and gender identity in a culturally and



socially accepted fashion for equal treatment towards the attainment of set common objectives (Miller and Katz, 2002; Gasorek, 1998; Hyter and Turnock, 2006; Roberson, 2006). Ostensibly, SMEs in the African body politic are meant to drive set objectives of growth and development. SME initiatives in Africa mean more than just businesses and entrepreneurship; they represent conscious and concerted efforts to get all strata of the economic ladder to contribute to socio-economic growth and development. This implies that SMEs are tools for social-economic inclusion and provide platform for those shut out of mainstream society like the less privileged to be actively engaged in socio-economic drive. Entrepreneurship ignites utilization of skills, experience and knowledge even among citizens at the bottom of the economic ladder (Arnstein, 1969). The need to coordinate effort and share information dynamically among all players in the economy regardless of gender, ethnicity, age, sexual orientation, disability or poverty is key to the survival of SMEs. Inclusive growth encourages equitable opportunities for economic participants during economic growth with benefits incurred by every section of society. Sustainable economic growth requires inclusive growth. Emphasis on inclusiveness, especially on equality of opportunity in terms of access to markets, resources, and an unbiased regulatory environment, is an essential ingredient of successful growth. The inclusive growth approach focuses on productive employment as a means of increasing the incomes of the poor and excluded groups and raising their standards of living (lanchovichina and Lundstrom, 2009).

Languepin (2010) reports that efforts to promote digital inclusion has made the use of mobile phones predominant among vulnerable people and proven to be a success as the rapid distribution of mobile telephony has made it possible for poor people to have easy access to useful and interactive information, thereby increasing the momentum of SMEs activities.

The United Nations Development Programme (UNDP) opines that many people are excluded from development because of their gender, ethnicity, age, sexual orientation, disability or

poverty. Development can be inclusive - and reduce poverty - only if all groups of people contribute to creating opportunities, share the benefits of development and participate in decision-making. In pursuing inclusive development, African countries have chosen SMEs as a critical element for creating productive and gainful employment. This is often paired with effective and efficient social safety nets to protect those who cannot work or who earn too little (Okewu, 2013).

The desire to provide business data that are easily accessible by all SME stakeholders just-in-time means a dynamic-based model of software development is required.

Service-oriented Software Engineering (SOSE) is a software engineering methodology focused on the development of software systems by composition of reusable services (service-orientation) often provided by other service providers. These services may be provided by others as web services, but the essential element is the dynamic nature of the connection between the service users and the service providers (Cervantes and Hall, 2004; Breivold and Larsson, 2007).

There are three types of actors in a service-oriented interaction: service providers, service users and service registries. Service providers are software services that publish their capabilities and availability with service registries. Service users are software systems (which may be services themselves) that accomplish some task through the use of services provided by service providers. Service users use service registries to discover and locate the service providers they can use. This discovery and location occurs dynamically when the service user requests them from a service registry Stojanović, (2005); Okewu and Daramola, (2014).

Figure 3 shows the service-oriented architecture on a web service platform.



## **12<sup>th</sup>**

### International Conference



Figure 3: Service-oriented architecture (Source: en.wikipedia.org/wiki/Web service). Downloaded February, 2015.

As illustrated figure 3, the service provider sends a WSDL (web service description language) file to a directory named UDDI (Universal Description, Discovery and Integration). The service requester contacts UDDI to know the provider for the data it needs, and then it contacts the service provider using SOAP (Simple Object Access Protocol). The service provider validates the service request and sends structured data in an XML (extensible markup language) file, using the SOAP protocol. This XML file would be validated again by the service requester using an XSD file. Generally, UDDI outlines which software system should be contacted for which type of data. In the event a software system needs a particular report/data, it would go to the UDDI and find out which other system it can contact for receiving that data. Once the system to be contacted is known, SOAP is used to make the contact.

#### 2.4 Related Work

Some of the previous efforts that are related to SMEs in Sub-Saharan Africa in the literature are presented as follows.

Olawale and Garwe (2010) provided information on their investigation of the obstacles to the growth of new SMEs in South Africa using the principal component approach. The most important obstacle was termed Financial which is largely an internal factor. The other obstacles were Economic (external), Markets (external), Management (internal) and Infrastructure (external). Despite highlighting challenges confronting SMEs, the study did not elaborate on using ICT strategy to tackle the challenges. Neither was there mention of the use of any software engineering approach.

International Finance (2015) researched challenges confronting SMEs in Africa and categorized them into four building blocks - conducive business environment (investment climate), limited management and operational capacity, access to markets and access to finance. It is of the view that SMEs make up 90% of all businesses in African markets and are important drivers of growth in economies across Sub Saharan Africa.

Hersman (2012) observed that the dearth of government support for companies has seen a proliferation of Information and Communication Technology (ICT) hubs spread across the entire continent. The mention of the role of ICT in stimulating growth and development is heartwarming even though the study stopped short of mentioning software applications used let alone the software engineering approach used for developing such software. The research also did not elaborate on the place of SMEs in inclusive development.

Leo et al. (2010) was concerned about the role that SMEs play for growth and development in Sub-Saharan Africa. The study pointed to substantial literature showing that there is a strong correlation between business environment and growth opportunities for SMEs. The better the business environment, the more SMEs will be established. It should be pointed out that Okewu (2015) had indicated that corruption is a major factor hindering progress in Africa and suggested the use of information and communications technologies (ICTs) in entrenching transparency and enhancing the African business landscape.

Mousley (2014) highlighted issues, challenges and prospects of developing SMEs in Africa. The work focused on two dynamic indicators - one, transformation of micro and informal firms into formal firms and two, growth of small and medium firms into globally competitive firms. As clinical as


the study was, it fell shy of mentioning that information technology is a viable strategy for addressing the constraints. Additionally, it did not harp on SMEs as a vehicle for inclusive development.

Obeche (2015) reiterated that SMEs hold the ace for growth and development of Africa's economies, but their potential cannot be fully tapped until critical challenges bedeviling the sector are tackled. While lack of financing and infrastructures such as power are critical challenges facing SMEs in Africa, the African Development Bank (AfDB) is partnering with governments, local and international financial institutions to strengthen the SMEs in Africa.

Though the emphasis of this research was more on the challenges posed by finance and power, and less on information technology as an emerging critical infrastructure, it nonetheless brought to the fore the invaluable contribution of infrastructure to providing an enabling environment for SMEs to thrive.

In a nutshell, we observed from the literature that none of the previous studies had focused on the implementation of an SME system in an African context using SOSE approach with a view to strengthening inclusive development, the main motivation for this work.

### 3. METHODOLOGY - AUTOMATED SYSTEM FOR INFORMAL SECTOR (ASIS)

The study used Nigeria as a test bed. This is against the background that it has the largest economy in Sub-Saharan Africa and findings here could, to a large extent, be representative of economic realities on the African continent. The proof of technology was done using Microsoft SharePoint while the underlying theoretical framework used was service-oriented software engineering. Guided by the software architecture process, the study progressed through activities of gathering architectural requirements, designing architecture to validating the service-oriented architecture

### 3.1 Architectural Requirements

The researcher gathered requirements and summarized the cross-cutting functional requirements of the proposed solution in Table 1.

Table	Table 1: Cross-cutting functional requirements			
Req . ID	Requirem ent	Brief Description		
R01	Add SMEs informatio n	The system shall allow authorized users to add SMEs information to the database depending on assigned rights and privileges.		
Ro2	Access SMEs informatio n	The system shall allow authorized users to access SMEs information from the database in accordance with assigned rights and privileges.		
Ro3	Edit SMEs informatio n	The system shall allow authorized users to edit SMEs information on the database in line with assigned rights and privileges.		
Ro4	Delete SMEs informatio n	The system shall allow authorized users to delete SMEs information from the database contingent upon rights and privileges assigned.		

### 3.2 Architectural Design

The deployment diagram for the proposed solution is shown in Fig. 4. Users can view outcome on personal computer (PC) and third party tool such as phone.



# 12<sup>th</sup>

### International Conference



Figure 4: ASIS deployment diagram

For financial deepening with the SME industry, payment systems like automatic teller machine (ATM) and point-of-sale (POS) terminal are incorporated. Also integrated is biometric system for identification and verification of SME stakeholders. The ASIS algorithm design is treated offline in this paper.

The ASIS software architecture pattern is presented as 3-tier architecture as shown in Figure 5.

### 3.3 Architecture Validation

To verify and validate the ASIS architecture, a prototype was developed using Community Site component of Microsoft SharePoint. The webbased solution was then tested using simulated interactive sessions between stakeholders.



Figure 5: Visual overview of the 3-tiered ASIS architecture

As the name suggests, the Community Site component facilitated online real-time conversation between SME stakeholders. Test scenarios were presented where service users (requesters) sought for information from the service providers housed in the service registry. To underscore the dynamic collaboration philosophy of SOSE which encourages real-time exchange between information users and providers, the prototype ensured that information was provided on real-time basis. This confirmed that ASIS is not only web-based, but also serviceoriented. Put in another fashion, service users were able to access service providers warehoused in the service registry of ASIS.

### 4. RESULTS AND DISCUSSION

The study conducted a mock-trial of ASIS using simulated questions and answers posted online real-time between users of the system. The researcher also evaluated possible threats to results obtained. This ASIS Simulator provided new information that online real-time dynamic communications can take place among SME stakeholders.

### 4.1 Results of Software Experiment

Outcomes of the test-run as graphically demonstrated in Figures 6-8 clearly indicate that both information requesters and information providers could exchange information on issues



that bother on the growth and development of SMEs in a dynamic fashion. This ensures that stakeholders are well informed about developments in the sector on minute-by-minute basis, thus boosting their confidence and commitment to the sector. ASIS can also facilitate implementation of regulatory, institutional and legal procedures, giving a facelift to the investment climate.

The following snapshots (Figures 6 - 8) show simulated postings on the ASIS platform. The simulation shows information exchange between the SME stakeholders who are seeking for information and stakeholders who responded on just-in-time.

deres at			with second life second
with the			Che die an
A hite street to	Autor	nated System for the Informal Sector	niffanis (for in an f
	B NR D	<ul> <li>A second state of the second stat</li></ul>	Contract His factor
		<ul> <li>Special descent of an electronic strength of the second strengt of the second strength of the second strength of the second</li></ul>	A REAL PROPERTY AND A
	in.	<ul> <li>Jose and Rey Control Sectors</li> <li>Sector Control Comparison in the last sector in the part (Sector Sector Sector Sector) - protection (Sector Sector Sector</li></ul>	Mechtage est
	and and a second	<ul> <li>Komment of Stells (1, 1), edge (subject on proposed Stells in Super- position on Stells) - an one sign (1)</li> </ul>	2
	Trees.	<ul> <li>In an initial characteristic second and the second s</li></ul>	4

Figure 6: The ASIS sub-site showing simulated items for SME stakeholder's discussion

AND A		1.000	42 million (1997)
	General	5363311 <del>83</del>	the second of
La de	Age op endersjon SM Society great annen Han Personen Sociation An age on and an on Mark Sprace Sections of the social	18	

Figure 7: An ASIS sub-site showing SME stakeholder asking question



Figure 8: An ASIS sub-site showing an SME stakeholder responding to question asked by another stakeholder

#### 4.2 Evaluation Threats

It is quite possible that an elaborate evaluation of the different modules of the proposed system, ASIS, could present new insights and change information obtained. Nonetheless, the subjects that participated in the survey have the required practical knowledge of the challenges confronting Nigeria's SME sector - lack of electricity, access to markets, access to finance, corruption, and investment climate barriers (regulatory, institutional and legal). They equally had sufficient practical engagements with ASIS. This offered them good basis to make objective assessment of the role of online real-time information in promoting growth and development of SMEs. It is therefore safe to say that there are ample reasons to take their views seriously.

In addition, only 2 classes of people were involved in the simulation exercise - one representing information service requester while the other represented the information service provider, which could in a sense limit the statistical significance of the outcome. However, the result of the experiment clearly indicates that both parties, who incidentally are the stakeholders in an SME information space, were adequately represented and were functional in the exchanges on acceptable service levels. This is considered to be a good result because at this point in the project, the core objective is to gain a first impression of the impact of information availability and accessibility on SMEs growth and development. Therefore, despite the limitation of using a limited number of



evaluators, there is sufficient ground to infer that there is a positive and preferential disposition to the use of information technology in confronting the challenges confronting the informal sector and promoting inclusive development through the engagement of the vulnerable groups in job creation and wealth generation. We can thus generalize that the SOSE developed automated information system is effective for facilitating growth and development of SMEs, and by extension overall socio-economic development of Sub-Saharan Africa.

### 5. CONCLUSION

Small and medium-sized enterprises (SMEs) are on a growing scale being acknowledge as productive drivers of economic growth and development for African countries. It is established that SMEs account for 70% of Ghana's gross domestic product (GDP) and 92% of its businesses. They also make up 91% of formalized businesses in South Africa and 70% of the manufacturing sector in Nigeria (Abor and Quartey, 2010; Mamman et al. 2007). SMEs not only contribute substantially to the economy but also is catalyst for economic diversification through their development of new and unsaturated sectors of the economy. Moreover, innovative and technology-based SMEs can provide a solid platform for expanding outside of domestic borders, and entering intra-regional and international markets. Hence, the motivation to build a web-based multi-tier architecture SME system, ASIS that can facilitate the realization of this objective. Taking into cognizance the increasing emergence of SMEs across sub-Saharan Africa and their impact, an automated system for the sector will enhance SMEs development, success and potential across the African continent. This study opined that an information technology architecture that supports SMEs access to markets, access to finance, capacity building and coordinated regulatory/legal frameworks was needed against the backdrop that once this pillars are galvanized, SMEs will prosper. The implementation of ASIS has demonstrated that online real-time information can be shared dynamically among stakeholders for best business decision making. And by extension, growth and development can be guaranteed through inclusive development that is the hall-mark of SMEs.

### 6. REFERENCES

- Abort, J. and Quartey, P., 2010. Issues in SME Development in Ghana and South Africa. In International Research Journal of Finance And Economics, (39), pp. 218-228.
- Arnstein, S.R, 1969. A ladder of citizen participation. In Journal of the American Institute of Planners, 35(4), 216-224.
- Bangudu 2013. World Bank wants Sub-Saharan African banks to serve SMEs more, Premium Time.
- Breivold, H.P. and Larsson, M., 2007. Component-Based and Service-Oriented Software Engineering: Key Concepts and Principles in Software Engineering and Advanced Applications. Proceedings of 33rd EUROMICRO Conference. ISBN 978-0-7695-2977-6.
- Calice, P., Chando, V. and Sekioua, S., 2012. Bank financing to small and medium enterprises in East Africa: Findings of a survey in Kenya, Uganda, Tanzania and Zambia. African Development Bank.
- Carstens, M., (2012). Best and worst sub-Saharan countries for start-ups. July 2012, http://ventureburn.com.
- Cervantes, H. and Hall, R.S., (2004). Technical Concepts of Service Orientation. Chapter 1, pp. 1-26 in Stojanović, Zoran and Dahanayake, Ajanthap. Service-oriented software system engineering: challenges and practices. Idea Group Inc. (IGI), 2004. ISBN 978-1-59140-428-6.
- Crnkovic, I. and Larsson (2003). Building Reliable Component-Based Software Systems. Artech House Publishers, ISBN 1-58053-3272, 2003.
- Debayan, B. (2011). Component Based Development - Application In Software



Engineering. Indian Statistical Institute Press.

- Doing Business (2012). Doing business in a more transparent world. World Bank, 20 October 2011, http://www.doingbusiness.org.
- Ernst & Young, (2011). Nature or nurture? Decoding the DNA of an entrepreneur. Ernst & Young, 2011.
- Gasorek, D., 1998. Inclusion at Dun & Bradstreet: Building a High-Performing Company. The Diversity Factor 8/4 (Summer) 2529.
- Global Entrepreneurship Monitor Survey Data, (2012). http://www.globalentrepreneurshipmonitor.
- Hersman, E., 2012. From Kenya to Madagascar:The African tech-hub boom. BBC News, 19July2012,http://www.bbc.co.uk.
- Hyter, M. C. and Turnock, J.L., 2006. The Power of Inclusion: Unlock the Potential And Productivity of Your Workforce. John Wiley & Sons.
- Ianchovichina E. and Lundstrom S., 2009. Inclusive growth analytics: Framework and application. Policy Research Working Paper Series 4851, The World Bank.
- International Finance Corporation, 2015. Sub-Saharan Africa. World Bank Group.
- Languepin, O., 2010. How mobile phones can help reduce poverty.
- Les Miserables (2012), The Economist, 28 July 2012, http://www.economist.com.
- Leo S.F, Chris A.G. and Green (SQW), 2010. SMEs and growth in Sub-Saharan Africa: Identifying SME roles and obstacles to SME growth. MENON-publication no. 14/2010, MENON Business Economics, Essendrops gate 3, 0303 Oslo, Tlf: 97 17 04 66.
- Mamman A., Elridge, D. and Branine, M., 2007. Skills need of SMEs and the informal sector in

Africa: Problems prospects of employment creation in Nigeria. Scientific Journal of Administrative Development (5), pp. 145-177.

- Miller F.A. and Katz, J.H., 2002. The Inclusion Breakthrough: Unleashing the Real Power of Diversity. San Francisco: Berrett-Koehler Publishers.
- Mousley P., 2014. SME Issues, Challenges and Prospects. FSS 2020 International Conference, Finance and Private Sector Department, Africa Region, World Bank.
- Obeche O., 2015. SMEs key to Africa's economic development. Business Section, BBC News.
- Okewu, E., 2013. Design and Implementation of Social Security Scheme: The Role of ICT. Proceedings of the 11th International Conference of Nigeria Computer Society. Iloko-Ijesha, Nigeria
- Okewu, E. and Daramola, O., 2014. Componentbased Software Engineering Approach to Development of a University e-Administration System. *IEEE 6th International Conference on Adaptive Science and Technology (ICAST 2014)*. Ota, Nigeria.
- Okewu, E., 2015. e-Government, e-Governance, and e-Administration: A Typology of Corruption Management Using ICTs. 15th European Conference on eGovernment (ECEG 2015). University of Portsmouth, Portsmouth, UK.
- Olawale, F. and Garwe, D., 2010. Obstacles to the growth of new SMEs in South Africa: A principal component analysis approach. *African Journal of Business Management*. Vol. 4(5), pp. 729-738, May 2010, ISSN 1993-8233 © 2010 Academic Journals.
- Pressman, R.S., 2009. Software Engineering: A Practitioner's Approach. 7th ed.
- Roberson, Q.M., 2006. Disentangling the Meanings of Diversity and Inclusion in



Organizations. Group & Organization Management 31/2:212-236.

Sommerville, I., 2011. Software Engineering. Ninth edition. Stojanović, Z., 2005. A Method for Component Based and Service-Oriented Software Systems Engineering. Doctoral Dissertation, Delft University of Technology, The Netherlands. ISBN 90-901910-0-3.



### SMART CITY: CONCEPT, APPLICATIONS AND ARCHITECTURAL MODEL

\*Efozia N.F.<sup>1</sup>; Immanuel T.<sup>2</sup>; Odebunmi F.<sup>3</sup>; Giwa O.<sup>4</sup>; Balogun T.<sup>5</sup>

<sup>1,2,3,4,5</sup> Prototype Engineering Development Institute (Pedi) IT Unit, Ilesa, Osun State Nigeria

<sup>1</sup>fenngo31@yahoo.com, <sup>2</sup>toluImmanuel@gmail.com, <sup>3</sup>admin.pedi@naseni.org

#### ABSTRACT

Cities are the economic, social and political hubs of our world. They contribute the most to our world's economy and consume most of its resources. Despite its contributions, cities still face a storm of economic, environmental and demographic challenges. In the process of tackling these challenges, the term "smart city" was conceived and is defined by some authors as the combined use of software systems, server infrastructure, network infrastructure and client devices to better connect some critical city infrastructure components and services, such as transportation, education, healthcare, and so on. Also, with the advancement in new technology platforms like broadband, software applications, mobile technologies, big data analytics and social technologies, there has been an explosion of new products and solutions that reach billions of users, connect billions of devices and create several new applications. This research paper is focused on reviewing the concept of smart city, its applications and possible future architectural design model.

Keywords: Smart City, IoT, Broadband, IoE, Soft Infrastructure, Sensors, Integrated Platform

#### 1.0 INTRODUCTION

Smart City is a powerful paradigm that applies the most advanced communication technologies to urban environments, with the final aim of enhancing the quality of life in cities and provide a wide set of valueadded services to both citizens and administration. A fundamental step towards the practical realization of the Smart City concept isin the development of a communication infrastructure capable of collecting data from a large variety of different devices in a mostly



uniform and seamless manner, according to the Internet of Things (IoT) paradigm. While the scientific and commercial interest in IoT has been constantly growing in the last years, practical experimentation of IoT systems has just begun. (Angelo, et al, 2014)

It is a developed urban area that creates sustainable economic development and high quality of life by excelling in multiple key areas; economy, mobility, environment, people, living, and government. Excelling in these key areas can be done through strong human capital, social capital, and/or ICT infrastructure. The emerging trend towards "smarter" cities strengthens the importance of information and communication technology in urban development. Wireless city initiatives can be seen as a vehicle to fulfil technologyrelated objectives which refer to wireless technology. Smart buildings are one of the major application areas of technologies bound to embedded systems and the IoT. Such systems have to be adaptable and flexible in order to provide better services to its residents.

The IoT represents a variety of things or objects which, through unique addressing schemes, are able to interact with each other and cooperate with their neighbours to reach common goals. These things include different (often embedded) devices including but not limited to sensors of various kinds, actuators, mobile devices, TV sets, vehicle computers; but also non-ICT appliances (dishwashers, microwave ovens. refrigerators), electrical energy sources and building components. Some of the key application areas of IoT are smart cities, smart power grids, smart health, smart transport, as well as smart buildings which includes smart living solutions(Schatten, 2014).

The concept of Smart City is a natural evolution because current global trends in energy supply and consumption are patently unsustainable; environmentally, economically, and socially. The scientific community believes that the future of human prosperity depends on how successfully we tackle the two central energy challenges facing us today: securing the supply of reliable and affordable energy; and effecting a rapid transformation to a low-carbon, efficient and environmentally benign system of energy supply. This can be tagged "an energy revolution". (Bartoli, et al, 2011)

Smart Cities are a future reality for municipalities around the world. These cities will use the power of ubiquitous communication networks, highly distributed wireless sensor technology, and intelligent management systems to solve current and future challenges and create exciting new services. Smart City officials will be essential visionary leaders who drive Smart City progress using public-private partnerships to invest in scalable projects, smart regulation to connect city laws to new digital realities, and innovation clusters to create jobs and vibrant economies. Its technologies integrate and analyze massive amounts of data to anticipate, mitigate, and even prevent many problems. This data is leveraged, for example, to intelligently reroute traffic and reduce accidents, identify crime hot spots and target resources for crime reduction, and connect citizens at work or out on the town. Smart Cities proactively provide services, notifications, and information to citizens such as where to find a parking spot or a new local shop or even to monitor air pollution. It connects citizens to local government and encourages more direct participation, interaction, and collaboration. It also provides solutions that are economically and environmentally sustainable. This is the potential future for every city and town.

When we consider the Internet ofEnvironment(IoE) and the Smart City, the importance of network infrastructure becomes apparent. Smart Cities must be built upon a pervasive broadband network to connect all parts of a city, and data must be sent faster, securely, in real time, and in unprecedented volumes. "Network infrastructure must be a key investment to prepare for our 2020 digital universe." (Ruthbea, 2013)

Hence, the purpose of building smart city is to provide better working and living environment, involving



multi-subjects as governments, businesses, and scientific research units.

#### 2.0 RELATED WORK

The concept of Smart City originated in a U.S. information technology company IBM. The CEO Sam Palmisano put forward this concept of smart earth in the roundtable after Obama took office in January 2009. Obama gave positive affirmation to the idea that smarter planet building would lead a new generation of economic investment and acknowledged such concept as the information superhighway plan. Since then, the construction of smart society has become a focus in countries throughout the world. The concept of "smart cities" has attracted considerable attention in the context of urban development policies. The Internet and broadband network technologies as enablers of eservices become more and more important for urban development while cities are increasingly assuming a critical role as drivers of innovation in areas such as health, environment and business. (Dmitry Namiot and Manfred Schneps-Schneppe, 2012)

Smart city has its root into the world as the Obama administration responded positively to the wisdom of smart earth with \$787 billion economic recovery and reinvestment strategy, in order to support and promote the development in fields of energy, healthcare and the Internet. Since then, development of smart earth has caught the world's attention. The United States, Japan, Korea, Europe and other countries made the IoT technology development into national strategy. U-Korea, The first u-society on the best of uinfrastructure, is the core development strategy in South Korea. Japanese's u-Japan strategic aims at realization of all the Japanese people, things, objects and persons connected (4U =Ubiquitous and Universal, the Useroriented, Unique). The European Commission has put forward the Internet of Things-An action plan for Europe program to promote the right to speak in the initial stage of IOT era. Singapore is building smart country, one of the pioneers constructing smart city. On June 19, 2006,

the Singapore government official launched the 2015 Intelligent Nation Plan. Singapore's Intelligent Transportation System (ITMS) is world's first intelligent transportation systems. The investment in Singapore makes itself the world's most smooth traffic city. Stockholm Sweden also began to use the smart citybuilding concepts to solve the congestion of traffic system. German has started implementation of the T-City project. Ireland set up Galway Bay Smart Bay project. South Korea's Incheon city government announced cooperation with technology companies. Smart City in China's development begins when Premier Wen Jiabao visited WuXi High-techMicro-nano Sensor Network R & D Center in 2009. The concept has been included in the 12thFive-Year Development Plan, focusing on investment in smart grid, intelligent transportation, smart home and smart logistic. (Kang-juanandLiu-ging, 2012)

The IoT is indeed a recent communication paradigm that aims at annexing into the Internet any kind of object, provided it is equipped with a microcontroller, a communication transceiver, and a suitable protocol stack. Hence, the IoT can enable easy access and interaction with a wide variety of devices such as, for instance, home appliances, surveillance cameras, monitoring sensors, actuators, displays, vehicles, and so on. The potentially enormous amount and variety of data generated by such objects will then foster the development of new services to citizens, companies, and public administrations. Therefore, the adoption of the IoT paradigm in a Smart City scenario is very attractive to public administrations, which may become the promoters for the adoption of the IoT paradigm on a wider scale. (Angelo, et al, 2014)

Also, smart city is based on *IoT*, because it improves informational infrastructure and urban management. It is put forward based on the requirement of city developing needs and driven by technology breakthrough. Its infrastructure is the fourth basic infrastructure after water, electricity and gas. Perfection of information technology is an important benchmark to



measure the degree of information, international competitiveness and influence for a city or country. As at 2010, 1500 cities had started smart city construction throughout Asia, Europe, the United States and Africa. The IoT technology is the technology push power of smart city building. Through information technology, the entire city functions as an organism. Any information in a cell (urban component) can be neural networked in real time to the brain (perception layer, and transmission layer). The brain analysis and data processing can then make a scientific decision-making in order to realize successful applications. Known as the third technology wave of industrial revolution, IoT provides technical supports to solve current urban development and social, economic, natural contradictions. Sensor and actuator layer targets at terminal objects, such as assets, environment, vehicles and even physical state of people. This level is made up of hundreds of millions of sensing equipment, to obtain information and data (similar to human nerve fibres feel perception of the physical world) and to perform instructions on terminal objects it receives. The terminal data collection equipment includes sensors, radio frequency identification (RFID) tags, smart meter, camera, GPS, mobile phones and so on. The sensing technology is the fundamental base in realization of a more thorough perception smart city. (Kang-juanandLiu-qing, 2012)

Furthermore, information security is the core and fundamental element of smart city where anyone can obtain any data in any place at any time through wireless network. The easy accessibility of information resources is also a potential leak of confidential information and brings risk of theft. Therefore, the construction of smart city should strengthen network and information security, set safety norms on information collection, and establish security protection and encryption system. The key element for information industry and the *IoT* are technical standards. As related information and technology in smart city is massive and complex, uniform standards must be established. Lack of technical standards, enterprises will be faced with investment risks or repeated switching costs. Research and development might be delayed because of potential replacement costs if there is no standard for technology compatibility. Thus it is required to establish a unified industry standards and transparent criteria. In order to ensure implementation of the decisions made by the *IoT* processing system, a strong executive government and organizational institution are also significant. (Namiot and Schneps-Schneppe, 2012)

There are various definitions of smart city, For example, IBM's Smart Cities digitize and connect infrastructures (IoT) to infuse them with new intelligence (IBM, 2012). As per Forrester, Smart City is the combined use of software systems, server infrastructure, network infrastructure, and client devices to better connect seven critical city infrastructure components and services: city administration, education, healthcare, public safety, real estate, transportation, and utilities. (Namiot and Schneps-Schneppe, 2012)

Consequently, for a city to be considered a Smart City, it must call for the cooperation of a multitude of fields of activities including industry, education, community participation, technical infrastructure, and various 'soft factors': "A Smart City is a city well performing in a forward-looking way in six characteristics (Smart Economy, Smart People, Smart Governance, Smart Mobility, Smart Environment, Smart Living), built on the smart combination of endowments and activities of self-decisive, independent and aware citizens." If analyzed, the above-mentioned definition can be broken down into the following characteristics. A Smart City must have well-developed connectivity obtained through a networked infrastructure. High-tech and creative industries (known as 'soft infrastructure') must emerge from this fertile foundation and over time attract new businesses and investments therefore producing both urban growth and a positive socio-economic performance. (Sauro, et al, 2012)



Furthermore, over the past two decades, communication networks have experienced tremendous growth and expansion all over the world. The explosive growth of many types of mobile devices such as smart phones, variations of tablet computers, and laptops, has fuelled the demand for more bandwidth with varying Quality of Service (QoS), with pervasive connectivity and at affordable costs. This has brought about smart city which is pushed by the application of the paradigm of IoT to an urban context is of particular interest as it responds to the strong push of many national governments to adopt ICT solutions in the management of public affairs, thus realizing the so-called Smart City concept. The final aim of smart city is to make a better use of the public resources, increasing the quality of the services offered to the citizens while reducing the operational costs of the public administrations. This objective can be pursued by the deployment of an urban IoT, i.e., a communication infrastructure that provides unified, simple, and economical access to a plethora of public services, thus unleashing potential synergies and increasing transparency to the citizens. (Andrea, et al, 2014)

### 2.1 ASPECTS OF SMART CITY

Smart Cities have been characterized and defined by a number of factors including sustainability, economic development and a high quality of life. Enhancing these factors can be achieved through infrastructure (physical capital), human capital, and social capital and/or ICT infrastructure.

A Smart City is a place where citizens interact with the city in order to satisfy their necessities (relationship, communication, green energy, economy, connectivity, accessibility, culture) and improve their quality of life using new technologies as shown in Figure 1.



**Figure 1:** Graphical representation of Smart City and its aspects. (Source: Małgorzata Hanzl, et al, 2014)

#### **City environment**

• Cities as inherent parts of the environment.

• Bringing "nature" into the city.

• Lack of proper care for the environment in the past.

• The simplification of urban landscape generates a loss of biodiversity.

• Trials to respect and protect the environment inside the city, creating green areas, parks and using green energy in order to reduce the pollution of the atmosphere and prevent climate changes.

#### Communication and relationship

• Communication - "the process by which information is passed between individuals and/or organizations by means of previously agreed symbols" (Little, 2012).

• Technology enables effective communication but its speed may constrain the ability to build meaningful relationships.

• Effective communication anticipates citizen's requirements. It uses codes: colours, light, sounds, signs, it enhances mass collecting of information, i.e., crowd-sourcing, it assists process organisation or just provides a communication platform (Hanzl, 2007).

**Culture and innovation** 



• Cities compete to provide comfort, which requires satisfaction of human needs: security and safety, education, etc.

• The recognition of the city's own culture and history is key for knowledgeable and respectful development.

• Every area of human activity may be enhanced with the use of technology: education, sports or musical and theatre spectacles.

• The constraints are creativity and organisation, the background is the adjustment to local culture.

#### Connectivity and accessibility

• The transportation system: accessibility of goods and citizens' mobility – effective management.

• Availability of interactive information increases the comfort of commuters, thus significantly improving travel conditions.

• The role of pedestrian movement as a form of transportation.

• Smart-smart urbanism should follow specific planning principles, privileging the complexity of ground-plane design, recognising the cognitive value of pedestrian experience(Senett, 2012;Hanzl, et al, 2014).

#### 2.2 DESCRIPTION OF A SMART CITY

A smart city can be described as a city that:

• Allows real-world urban data to be collected and analyzed by the use of software systems, server substructure, network infrastructure, and client devices.

• Implements solutions, with the support of instrumentation and interconnection of sensors, actuators, and mobile devices.

• Can combine service production and an intelligent environment, exploits accessible information in its activities and decision making and adopts information flows between the municipality and the urban or business community.

The city may be considered as a service organization with citizens as the customers - it provides services to its citizens. There is a demand for smarter, effective, efficient and more sustainable cities, pushing the collective intelligence of cities onward, which can improve the ability to forecast and manage urban flows, and integrate the dimensions of the physical, digital and institutional spaces of a regional agglomeration. Urban development and improvement of the city has been turning towards technology. Smart cities use different information and communication technologies (ICT).

Hence, smart city services and applications focuses on how to shape future Internet based services and applications from a smart city perspective. The deployment, implementation and approval of innovative internet based services and applications have to be made in order to permit facing the challenges of advanced cities. This change and prospect especially involves people who work with knowledge and information. The creation of information stuff is not restricted to a particular location, and the resulting products are typically delivered through the network. Smart city services are also available through wireless mobile devices and are enabled by services oriented enterprise architecture including web services, the extensible mark-up language (XML), and mobilized software applications. (Novotný, et al, 2014)

#### 2.3 CHARACTERISTICS OF SMART CITY

Smart cities refer to the urban centers which are made safe, environmentally green and efficient. The management of all utility services whether power, water, gas or transportation are maintained using advanced integrated sensors, electronics and networks. Finally, the services are interfaced with computerized systems comprised of databases, tracking and decision making algorithms. Technology has changed the traditional ways of city development. The characteristics of a smart city are:

(a) A broadband infrastructure which is widely available and affordable to all, including developed and undeveloped area.



(b) Applications and services in different areas such as safety, health, education and economy.

(c) An interconnection between the communities through integrated service architecture.

(d) A platform for innovation that promotes the development of new applications and services.

With the Internet and worldwide web, teleconferencing, video communication devices, cell phones, people and companies now have freedom to choose how they want to be organized in a smart city. Moreover, wireless technologies, have gained great momentum by the cities around the world as quick and effective technologies for enabling Municipal Area Networks (MAN). The availability of advanced broadband communications, services and effective electronic applications help in bridging the "digital divide" and improved economic competence. In short, Information and Communication Technology (ICT) are increasingly recognized as a powerful instrument for reducing poverty, promoting and facilitating sustainable good governance(Shahnaj and Fouzia, 2012).

### 3.0 APPLICATION AREAS OF A SMART CITY

A smart city is recognized by the availability of some very common applications. These front-end services are mostly electronically managed and maintained. The application areas of a smart city are numerous. Some of them are as follows:

**Education:** The spread of education in the smart city is ensured through distant learning methods. More benefits are derived by establishing interactive learning process. Although not much effective but off-line digital learning is also a modern tool for education. Digital learning includes the provision for digitized lecture notes, electronic books and journal. Apart from the education itself, the system of education is computerized by providing web-based results, course content, students' record and/or teachers' performance. **Public Utility Services:** All of the utility services provided to the citizens: electricity, water, gas etc. use state-ofthe-art techniques using leading-edge ICT to change the way of services. Such techniques are used in order to improve customer service, productivity, effectiveness and efficiency.

**Public Health Care:** A prompt and accurate diagnosis is the fundamental requirement for an effective public health care system. For that a distributed database is developed with patient records. And most importantly accessibility of the patient records to the doctors is established. This data includes the test reports of the patients.

**Public safety:** At the key installation and security sensitive area, network based security camera is installed for crime prevention. Moreover, in order to make public life easy, smooth and reduce traffic congestions, synchronous traffic signalling is computerized. An online traffic update system is also established for the public to follow roads with less traffic.

**Business:** Different ICT projects in public and private sectors are encouraged to involve local ICT educated graduates as this creates more ICT related jobs for the younger generation. Experts are developed gradually for creating opportunities for more business from home and abroad. This, in turn, improves internal communication to help strategic planning and prioritizing resources as well as promote innovative thinking and collaboration. Within a short span of time, people learn how to become a successful entrepreneur.

**Standard of life:** In a smart city, the government delivers new "value-added" services to its citizens using leadingedge technology to improve their quality of life. It is expected that each individual can derive economic benefits by accessing information in cost-effective ways. A smart city provides the tools and infrastructure to let citizens and community organizations take advantage of the information age and to participate and express their views as part of local decision-making process. By providing information open for all, everything becomes transparent. By this, the misappropriations and corruption in the society are reduced and economic



development of the country is ensured. (Shahnaj and Fouzia, 2012)

### 3.1 NEEDS FORA SMART CITY

Countries can operate effectively in the new global economy only if they meet two conditions. They must (i) command adequate information and communication technology infrastructure and (ii) generate the human resource to operate it. The ability to maximize the use of knowledge is now considered to be the single most important factor in deciding the competitiveness of countries as well as their ability to empower their citizens through enhanced access. The smart city is needed for several reasons such as:

- a) Cities and Towns are growing in size and influences making the world becoming more urban and global.
- b) Government is being reformed e.g. decentralization /devolution of governments
- c) Technological developments and e-commerce are having a profound effect on society.
- Increasing pressure to alternate service delivery i.e. find creative methods through which municipalities can mobilize energy capacity and resources outside the municipality for the development of the area.
- e) Increasing pressure to become service and customer oriented.
- f) Increasing pressure to involve the community to the work of municipality
- g) Increasing pressure to provide all relevant information to other levels of government, communities and other stakeholders.
- h) Lack of integration between departments and directorates.
- i) Increasing bureaucracy huge reliance on manual processes and people based processes.
- j) IT systems focused mainly on cost reporting not business enablement.
- k) IT enabled governance, socio-economic development as well as Administration/Service delivery.

(Shahnaj P. and Fouzia F., 2012).

### 4.0 MODEL AND FRAMEWORK DESCRIPTION

In turning talk into action, a Smart City Framework is a simple decision methodology that enables both the public and private sectors to plan and implement Smart City initiatives more effectively. Most cities actually undergo this process in an intuitive way rather than in a clearly structured manner. A structured method not only will enable efficiencies in city infrastructures, but also transparencies into how cities work. Smart City Framework as shown in Figure 2, starts with city objectives as its base, against which all initiatives are then measured(Gordon and Shane, 2012).



**Figure 2:** Smart City Framework Layers (from bottom to top). (Source: Gordon and Shane, 2012)

Figure 3 represents the city services intended for a smart city model and services are divided into three different groups:

- 1. Infrastructure services
- 2. Community services
- 3. Noncommunity services



water, energy, electricity, waste, and data- and The community telecommunication. *services*include services that are provided directly to citizens and visitors of the city. These are services like education and healthcare, but the community service layer also includes things like proximity to commerce, culture, sports and parks. Furthermore, the non-community services consist of other services where ICT could be used and which will have an impact on sustainability of the city. The finance services include for instance creation of an attractive business environment, while a workplace service relates to improve work efficiency, and persuasive information services to influence people's choices.

The city needs to attract both people and business. Many people move to cities aiming for a better life, where "better" stands for different things depending on your situation. Employment and the household budget are Important factors for groups of various ages, while health becomes more and more important with age (Uchida, et al, 2011). For visitors, the availability of culture, architecture, restaurants, commerce, sports, parks and nature are of special interest. The growth state of a city, expressed as mature, transitional, and emerging, is related to what kind of challenges the city needs to tackle. For instance a city which grows fast will have major challenges providing people with basic needs such as access to food, water and safety. For transitional and mature cities the challenges are more related to choice and convenience, and lifestyle and independence. (Loft, 2006)



Figure 4: is another model that has mostly technical view, which concentrates on how (sub) systems interact with each other supported by telecommunications and information technology. The city is divided into the built environment (including homes, offices and shops and the devices within them), infrastructure-based sectors (e.g. energy and waste) and service-based sectors (e.g. healthcare and education). There is possible interaction between elements within any of these subsystems as well as between subsystems. Smart city infrastructure sectors, such as telecommunications, information technology and electronics, enable and support this interaction. A common theme in the example smart city models is the use of sensors to collect data from the city which, through platforms, can be combined, stored, analysed and displayed. This provides decision support for actors in the city who can then act and make changes,

theeffect of which can in turn be measured. (Koen, 2014).

#### 4.1 ARCHITECTURAL DESCRIPTION

The illustration in Figure 5 is an architectural data flow model of a Smart City, and it shows how data is created in social and physical systems, collected, transmitted, stored and possibly shared before the data can be analysed, displayed and finally used to make decisions. At each step there are different actors involved (e.g. the person whose actions lead to new data, the owner of a

#### Figure 4: A Smart city Model

(Source: Koen, 2014)



Figure 5:Data Flow of a Smart City Model (Source: Koen, 2014)

sensor, the maintainer of a web platform, and various owners of the data) in the social system. Also, technical challenges will be addressed such as those given on the right in the figure above, with the first column highlighting a few technical challenges (e.g. related to interfaces and interoperability) while the second shows a number of social issues (e.g. privacy, security, monetization). Such data flows can be observed within as well as between the systems as shown in Figure 4.



According to Nina. and Anna (2013), the following requirements are to be considered before the implementation of the architectural design of any smart city in order to avoid waste of resources.

• Selection of indicators: to have a manageable, yet sufficient amount of indicators both on a city level and for specific ICT solutions used in the city.

• **Data:** to handle case specific as well as general publicly available data, considering both city-related data and evaluation indicators, with the possibility to define baseline, reference year, etc.

• **Transparency in city boundary:** to have transparency in the definition of the city boundaries, both geographically and around what impacts are included in the assessment e.g. impacts occurring within the defined city boundaries, in the surrounding region, on a national level or on a global level, e.g. whether import/export is included.

• **Transparency in results:** avoid merging impacts into too few categories and avoid translating different impacts into one unit (e.g. money).

• Life cycle thinking: use a life cycle perspective when possible and especially for the ICT solutions.

• **Realistic scenario for ICT implementation**: the scale of the ICT solution and its impacts should be based on relevant data for the specific city.

After the above have being considered certain research, questions need to be answered so as to access the effectiveness of implementing smart city and broadband;

- Identify ICT solutions to apply to a city's implementation process and its cost effect.
- Define the system boundary for the city and a functional unit for the assessment.
- Build scenarios and assess ICT solutions at a city level.
- Select indicators to connect solution specific impact results to overall city level sustainability.
- Scale ICT solution scenarios between cities.

### 4.2 IMPLEMENTATION REQUIREMENT Hardware Requirement:

- Strong wireless broadband network (Wi-Fi or WiMAX) which is crucial for communication within the same region, same country and global network
- Technology (i.e. Indicators or sensors) for communication and dissemination of information.

#### Software Requirement

- Integrated Database for each incorporated city service in the smart city model
- Development of applications for sectors, utilities, governance e.t.c
- Measurements: application to take statistics of the impact, intelligence, innovation and efficiency of smart city.

#### **Human Requirement**

- Programmers to Model each of the smart city service, develop the database required for each service and also program the sensors.
- Technical Personnel
- Network Experts

### 5.0 CONCLUSION

The concept, characteristics, application areas as well as reasons and importance of Smart Cities was discussed in this paper. It has gained importance in the last years, as a means of enabling services and applications available to the citizens, companies and authorities that are part of a city's system. It aims at increasing citizens' quality of life, and improving the efficiency and quality of the services provided by governing entities and businesses. This perspective requires an integrated vision of a city and of its infrastructures, in all its components and it has to



incorporate a number of dimensions that are not related to technology, e.g., the social and political ones. As a critical infrastructure element of future society, its architectural design model was highlighted. Like in the aspect of transportation, smart city traffic routing will be of great benefit as its being presently implemented in cities such as Lagos and Ibadan in Nigeria in an attempt to decongest traffic in those cities. Also, Smart City requires the highest levels of security. A comprehensive architecture with security built in from the beginning is necessary. In order to achieve user consent, trust in, and acceptance of Smart Cities, integration of security and privacy-preserving mechanisms must be a key concern of future research. We recommend that our country Nigeria starts taking the building of smart cities serious and include it as a long term budget plan.

### REFERENCE

Andrea Z, Nicola Bui, Angelo Castellani, Lorenzo Vangelista, Michele Zorzi, (2014) "Internet of Things for Smart Cities" DOI 10.1109/JIOT.2014.2306328, IEEE Internet of Things Journal. Retrieved on 23-3-2015 from www.06740844.pdf/ Angelo C., Andrea Z., Lorenzo V., Michele Z, (2014) "Padova Smart City: an Urban Internet of Things Experimentation" The International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM) 16-19 June 2014 Sydney, Retrieved on 23-3-2015 from Australia. http://patavinatech.com/PDSmartCity\_CR\_preprint.pdf/

Bartoli, A. Hern'andez-Serranoy J., Sorianoy, M., Dohler, M., Kountourisz, A. and Barthelz, D. (2011) "Security and Privacy in your Smart City". Retrieved on 23-3-2015 fromhttp://www.SmartCitySecurity.pdf/

Dmitry N. and ManfredS., (2013) "Smart Cities Software from the developer's point of view" retrieved on 23-3-2015 from www.1303.7115.pdf/ Gordon F., and Shane M., (2012) "Smart City Framework: A Systematic Process for Enabling Smart + Connected Communities". Cisco IBSG,

www.cisco.com/go/smartconnectedcommunities. Retrieved on 30-3-2015 from www.smart- cityframework.pdf/

Koen V.M.,(2014) "Mapping Smart City Standards: Based on a data flow model" www.bsigroup.com/smartcitymapping. Retrieved on 30-3-2015 from http://www. BSI-smart- citiesreport-Mapping-Smart- City-Standards-UK-EN.pdf/

Loft H, G. (2006) "Megacity challenges – a stakeholder perspective", Economist Intelligence Unit, Siemens, Munich Germany. Retrieved from http://id.siemens.commegacity-challenges.pdf/

Lv K.andXu L, (2012) "Study on the Smart City and Its Economic and Social Effect Basedon the Internet of Things" International Conference on Information Engineering (ICIE) 2012 Lecture Notes in Information Technology, Vol.25.

Małgorzata H, Antonina M., Alizee T., Marcos M. Iban M., Fanny N., (2014) "Smart City - a Quest for Innovation within the EPS Framework" Retrieved on 23-3-201from www.CORP2014 SLIDES64.pdf/

Nina L, and Anna B.,(2013):Evaluating sustainability of using ICT solutions in smart cities methodology requirements".Retrieved on 1- 4-2015from http://www.loevehagen-bondesson-evaluatingsustainability-of-using-ict-solutions.pdf/

Novotný R, Kuchta R, Kadlec J (2014) "Smart City Concept, Applications and Services".

J.TelecommunicationSystem Management 3: 117. doi:10.4172/2167- 0919.1000117, an open access journal3(2)1000117. Retrieved on 25-3-2015 from www.smart-city-concept-application-and-services-2167-0919-117.pdf/



Ruthbea, Y. C., (2013) "Smart Cities and the Internet of Everything: The Foundation for Delivering Next-Generation Citizen Services" October 2013, IDC Government Insights #GI243955. Retrieved on 25-3-2015 from www.ioe\_citizen\_svcs\_white \_\_paper\_idc\_2013.pdf/

Sauro V, Sara B., Alberto S, (2012) "How to Co-Create Internet of Things-enabled Services for Smarter Cities" SMART 2012: The First International Conference Smart Systems, on Devices and Technologies ISBN: 978-1-61208-225-7.Retrieved on 25-3-2015 from http://www.smart 2012 2 50 40077.pdf/ Schatten, M. (2014): "Smart Residential Buildings as Learning Agent Organizations in the Internet of Things",BusinessSystems Research, 5(1),pp. 34-46. DOI:10.2478/bsrj-2014-0003.Retrieved on 25-3-2015 from www.bsrj-2014-0003.pdf/Shahnaj, P. and Fouzia F., (2012) "TheConceptofSmart City:Digital Bangladesh Perspective" InternationalJournal ofEngineering Research and Developmente-ISSN:2278-067X,p-ISSN:2278-800X,5(5)pp.48-54.

Uchida, Y., Ohtake, F., Komamura, K., Hiroi, Y., Makino, Y., Miyamoto, M., Yamauchi, N., and Yamada, M., (2011) "Measuring national well-being – proposed well-being indicators".Retrieved on 1-4-2015 from http://www5.cao.go.jp-proposed-well-beingindicators.pdf/

86



### A USER-CENTRIC DESIGN TOWARDS PERFORMANCE DEVELOPMENT FOR ADAPTIVE E-LEARNING SYSTEMS

### **Kingsley Okoye**

School of Architecture Computing and Engineering, University of East London, UK

u0926644@uel.ac.uk

#### ABSTRACT

Effective utilization and profitability in today's e-learning systems demands a well-designed and easy to use system that meets the goal of the learners or intended users. To meet the needs of intended users, there is requirement for e-learning systems to embody technologies that support learners in achieving their learning goals and this process don't happen automatically. This calls for the development of e-learning systems that exhibits a high level of usability and achievements on performance; since what most users of these learning platforms do is to read, view and interact with the contents that are embodied in these educational products. This paper considers the conditions and standards that are critical to ensuring quality of usability and performance centered on e-content users. The focus is on identifying some of the common design problems with e-learning systems as means to revealing the implications for designers to stick to usability standards when building their products. Designers of e-learning systems must put into consideration what the learners expect from a particular product, and must be planned and developed to meet the intended users need. To this end, this paper proposes a personalised adaptive learning system to help address some of the design problems with present e-learning systems in order to improve Learners engagement and outcome on performance.

Keywords: Accessibility, Adaptive Learning System, E-learning, Performance, Usability

#### 1. INTRODUCTION

E-learning has been described as a digital learning platform that integrates various multimedia components (text, images, sounds, videos, graphics etc.) which are developed with the intention to be used for educational and training purposes within various organisations including, schools, universities and further education centres (Okoye, et al, (2014). Current innovations is what motivates software engineers over the years to provide personalised adaptive e-learning systems. With such system, learners can proceed at their own pace, get recommendations about what learning content best fits their learning needs, practice as much as they need on their own, and move ahead to greater challenges when motivated by interest, or data that demonstrates they have mastered a skill. Personalised adaptive e-learning system is imperative and necessary to provide continuous intelligent recommendations, guidance and feedback on learner's performance. Studies have shown that challenges in current information-rich world is not only to make information available for



learners at any time or in any form, but should essentially offer the right content to the right user and in the right format [Huang and Shiu, 2012, ; Yu Z. et al, 2011). E-learning systems should dynamically support different adaptive learning paths and contents to fit learner's diverse needs based on the users' profile, background, learning styles and goals. This should take into account the fact that there is an additional task of matching these persons (user profiling) with solutions that best fits their particular learning needs (personalization). Elearning systems should enable users to create, share and collaboratively edit contents to suit their individual learning needs and styles.

This paper, explore some of the common design problems with e-learning systems performance and usability that needs to be solved in order to provide a user-centred intuitive, and effective use of such systems. We reveal the implications for designers to stick to certain design standards when building e-learning applications and then propose an adaptive e-learning system to help address some of the design problems with present e-learning systems in order to improve learners' engagement and outcome on performance as well as provide road map for future improvements. The research looks at the determining factors from three perspectives; Accessibility, Presentation Design, and Standard Compliance; bearing in mind that if elearning systems developers do not take into consideration these factors along with the user experience, that most users tend to turn away from the readily available products looking for easier to use and intuitive e-learning platforms.

The rest of the paper is structured as follows; in the next section, we discuss and analyse appropriate related works in relation to this area of research. Next, we identify some of the key design criteria necessary towards achieving performance and usability in e-learning systems, and then propose design principles that serves as an effective guide towards the design and development of a usercentric adaptive learning system. In addition, we propose an automated learning system that is capable of detecting changing trends in learning behaviours and abilities in order to help address some of the design problems with present elearning systems. Subsequently, we discuss the importance of the design methods described in this paper and its necessity towards the achievement of performance and effectiveness in e-learning systems. The last section concludes the paper and points out directions for future research.

### 2. RELATED WORKS

In recent years, there has been a combination of factors that affect learners experience with elearning products including the level of efficiency in use and effortlessness learning, lack of satisfaction with the system, confusing navigation, slow in loading and download time, frequency of errors and difficulty for first time users in learning the interface quickly.

There is evidence, that in the design of e-learning products that it is essential to define usability goals, and to specify the intended context of users and that the adoption of clear, visible and concise guidelines for assessing usability in e-learning products will help both learners and designers in achieving an intuitive, in-depth and effective use of learning technologies. Okoye, et al., (2014) put forward a novel framework for practice that enables a deeper understanding of accessibility requirements for dynamic web applications centred on a particular user group. Through evaluation of collected facts and analysis; the authors propose and implement solutions to the dynamic web accessibility issues by designing and deploying a software application, and its implementation for best practice that informs developers on how best to significantly and effectively approach the design of e-contents/applications with accessibility issues in mind as well as considering the user groups.

There has been remunerations both in theory and in practice towards developing standardized and strategic means of evaluation of e-learning applications to accord with the changes in demand for users. Okayed, K. et al (2014), authors proposed theoretically means to rich and enhanced usability e-contents and validated the in technological impact by suggesting two types of methodology; one that suggests the removal of unnecessary and non-essential contents, new materials to be added to a content repository and/or re-use of already existing contents (the nofrills methodology) - and another which identifies



the potential users as well as the content type, creates prototypes fitting the expected user's need and finally evaluates the developed product with test/heuristics to analyse its impact on the users (the Agile methodology).

Niu and Kay, (2010) mentions that an important pragmatic concern is that it must be inexpensive to create e-contents and its presentations for different learners. The authors argue that elearning platforms should dynamically update the representation of knowledge to take into account the changing state of the leaners and changes in the information that is relevant to each user over time. They further suggest that such system is applicable and effective especially for users with learning difficulty; to manage their learning progress and also help improve their activity of daily life, thereby bridging the gap between such users and other learners seen as non-disabled.

In another research, Okoye, et al, (2014) used the idea of process mining to discover, monitor and improve the set of recurrent behaviours that can be found within learning processes. The authors utilised the technique in order to address the problem of determining the presence of different learning patterns within a learning knowledgebase. The study constructs a semantic learning process model; a User-Oriented Learning Knowledge-Base system (UOLKBs) that is of great impact and significance in this area to drive learning using process mining techniques to discover new rules through semantic reasoning, and adopting web languages such as Ontology Web Language and Semantic Web Rule Language. The result of the research shows that learning is the flow of activities within a Learning process Knowledge-Base (workflow), and being able to use Description programming Logics and languages to automatically compute the class hierarchy of learning activities is one of the major benefits of building personalized adaptive e-learning systems. Annotation properties are used to add information (Metadata - data about data) to the classes, individuals and object/data properties within the learning knowledge base. The outcome is relevant in bridging the gap between the levels of learning for different users by providing them with the same learning opportunity; through a system that adaptively support the personalisation of contents based on data regarding the users learning behaviour or actions.

Process discovery, which lately has been seen as the most important and most visible intellectual challenge related to semantic mining of processes e.g. learning process, aims to automatically construct useful models like Petri net (Murata, (1989)) or a BPMN model (Van der Aalst (2011)) and describes causal dependencies between learning activities (Fahland and Van der Aalst, (2012)). In principle, one could use process discovery to obtain a model that describes learning in reality. The second type of process mining is conformance checking where, an existing learning process model is compared with an event log of the same process to check if in reality it conforms to the resulting learning model (Adriansyah, et al, (2011); Calders, et al, (2009); Cook, and Wolf (1999); Munoz-Gama, Rozinat, and Van der and Carmona, (2011); Aalst, (2008); Weerdt, at al, (2011)). Conformance check could imply that the model does not describe the executed learning process as observed in reality or is being executed in a different order. It could also mean that activities in the model are skipped in the log or that the log contains events not described by the learning model. Given this drawback, the last type of process mining; model enhancement comes into play. Van der Aalst et al (2011) used the idea of an enhanced existing model to maintain compliance and to quantify deviations using information about the actual process recorded in the event logs from a given process.

According to the authors in (Huang and Shia, (2012)) searching for suitable learning paths and content for achieving a learning goal is time consuming and troublesome especially on dynamic learning platforms. To tackle these problems, the authors proposes a User-Centric Adaptive Learning System (UALS) that uses sequential pattern mining to construct adaptive learning paths based on users' collective intelligence and recorded events, and then employs Item Response Theory (IRT) with collaborative voting approach to estimate learners' abilities for recommending adaptive materials.

This paper differs from these previous works in several aspects. First, an automated learning system that ensures performance through user-



centric design was provided. The focused on personalizing learning based on user's profile or learning behaviour as opposed to most existing systems that provide guidance based on views of designers or experts. Second, the paper also support e-learning process bearing in mind the determining design factors that allows for usability performance from three perspectives; and presentation design, accessibility and standard compliance. Third, this work is not only intended to ensure learner's ability to learn or meet their learning needs but is expected to be useful in providing learning path and guidance based on individual differences. This is achieved by collecting user's initial capabilities and preferences on interaction and then determine which adaptations or further assistive measures are best suited or may be required through time.

### 3. DESIGN APPROACHES AND METHODS.

In this section, we ascertain some of the key design criteria necessary towards achieving performance and usability in e-learning systems, and then propose some design principles that serves as an operational guide towards the design and development of user-centric adaptive e-learning systems.

### 3.1 Accessibility Design

Accessibility is essential for e-learning systems usability, as learners will not be able to interact and complete tasks if they find it difficult to use the system. Studies have shown that many e-learning applications do not provide fully sufficient level of accessibility to the users (Okoye, et al (2014); Parmanto and Zeng, (2005); Saldaño, et al., (2013); Ali, et al, (2008)). These studies suggests that designing and developing a reasonably accessible elearning application should be a priority to the designer as most learners are more satisfied with accessible platforms, attracts more users and are more likely to make return visits.

The increase in dominance of various computer user interface results in conditions whereby learners with disabilities finds it cumbersome to use and access the e-contents. This group of learners may unintentionally be kept out from the feasible benefits of e-learning; if designers do not conform or put into consideration the problems with accessibility of design structure and outcome on performance. Take for instance, in a situation whereby users who are hard of hearing or deaf are being provided with e-learning materials that includes sound and does not contain texts or captions, or on the contrary, e-learning platforms that tends to exclude users with visual impairments by presenting the contents only in text or graphical forms without sound. According to Okoye, et al (2014) provision of multiple ways to operate the technology and retrieve information so that users can choose alternatives based on their physical capabilities; is a great way to help ensure users have an improved accessibility and support for the learning technologies they use. For instance, ensuring keyboard accessibility and navigation, and providing accessible user interface control over the font size or page colours. This means that a good design practice should provide various or diverse means of accessing the learning resources.

The IMS Global Learning Consortium (IMS, 2006) provided guidelines for developing accessible elearning platforms. They stated specifically that in dealing with e-learning applications, that these guidelines addresses accessibility issues in tests, interactive exercise, presentation tools, repositories, schedule organizers, threaded message boards and synchronous collaboration tools such as text chat and video conferencing. They proposed and recommended the use of standard technical formats in providing accessibility for e-learning applications, and that in providing accessibility; e-learning application developers requires a detailed understanding or knowledge about these standards and guidelines.

The International Standards Organization (ISO) in part of ISO 9241 (Travis, (2009) contains general recommendations to improve the accessibility of learning technologies. The standard advocates a conceptual structure for accessibility that comes in four dimensions.

Comprehension and specification of context to be used focusing on diversity of the users, the significance of objective, resources and interface features that impacts accessibility.



- Identify the accessibility requirements for users, knowing who the intended users are.
- Designing the model and algorithm for the product considering accessibility.
- Evaluating the accessibility of the user profile and design algorithm of the elearning product with the intended user group.

From the arrangement, we see without doubt that e-learning application developers should bear in mind the need and requirements of the learners in their context goal when designing new learning products. The process involves putting into consideration the requirements and purpose for various users together; including users with diverse background, educational capabilities and disabilities. For instance, stipulating list of terms and definitions or summary of the necessary knowledge the learner must have already as to be able to accomplish their learning purpose will make the e-learning product more accessible especially for first time users. Consideration of accessibility factors is a great way of increasing performance and usability of e-learning systems.

### 3.2 Design Presentation

Presentation Design is the quality of content and components in e-learning products (Leacock and Nesbit, (2007)). This is relevant to all expository media with regard to text, images, sounds, videos, graphics and animations. Presentations that are very high in quality usability standard is expected to incorporates aesthetical design values and development of learning information in formats that are efficient and consistent with the fundamental essence of research and knowledge of educational multimedia products, by demonstrating standards in the development of elearning applications.

The principles about clear and concise expression for data graphics and writing style recommended by Tufte back in 90's represents the significance of the presentation design subject in the design of elearning products, Tufte, (1997). Further, findings in support of Tufte design presentation results from the provenance of the abstract quality of the human working memory as expressed by Mayer's principle for developing educational multimedia products (Mayer and Moreno, 2003)

Mayer & Moreno (2003) described presentation design as being a basic and intrinsic element in the process of learning due to interactions among the components of the systems designed for learning, and that these components cannot be impaired without affecting the purpose of the system. To this end, an effective presentation design is relevant to e-learning systems development and is expected to efficiently contribute to intellectual and learning development. E-learning systems that have poor presentation design can result to extraneous information representation, which tends to reduce the quality and amount of information available for learners understanding.

The International Standards Organization (ISO) are engaged in developing a new standard called the ISO/AWI 23973 "Software ergonomics for World Wide Web user interfaces". They has been developing ergonomics standards for over 20 years in the field of human-system interaction. Most of the set standards contain general principles from which appropriate interfaces and procedures can be derived. Their purpose and strategy has been focus on the following aspects of user interface design; high-level design decisions and design strategy, content design, navigation and search as well as content presentation. These principles have been used by software developers as a framework guide their design towards improved to performance in e-learning applications as we show in Figure 1.



Fig 1. Human-System Interaction Design



Following these practice, we propose some design principles that can serves as an effective guide to minimizing these extraneous information content through;

- Clear and logical principles that suggest the excluding of materials that are irrelevant and are not needed by the intended users.
- Contiguity rule that suggests the presentation of components that the learners can manage and integrate rationally in time and space.
- Modality rules in addition to standard which suggests explanation of graphics or animations with some form of audio narration other than text only (Okoye, et al (2014)).

These method has been adopted by the software developers' community and has proven to be effective towards the design and development of personalized and adaptive e-learning systems (Okoye, et al (2014); Stephen, (2012)). There is strong evidence that presentations which combine graphics with text, most of the time tends to magnanimously ease and improve learning when compared with text only designs (Stephen, (2012)). Designers have to understand the implication of graphical representation in presenting speech notions so prominent from their more selfexplanatory use in presentation of useful data and information. Table 1 shows the technical criteria and benefits of performance and user-centric design.

Table	1.	Technical	Criteria	and	Benefits	of
Perfor	mar	າce/User-cer	ntric Desig	gn		

Critical Success Factor	Technology Enabler	Projected Benefits
	Flexibility –	Return on
Protect	Writing fewer	expected
usability by	lines of code	outcome
avoiding	results in	Commercial
resource loss	quicker delivery	viability of
	of product	usability
	offerings.	protection
	Enabling	program –
	intuitive learning	Increased agility
	model based on	to meet learning

	Web services	demands
Enhance e- learning operational efficiencies	Interoperability – XML-based data exchange simplifies application integration with back-end systems and legacy environments enabling data mining and rich management reports	Real-time noise detection and direct notification to users at learning stage. High available and scalable solution Data transfer and storage as XML enables integration with back-end systems and improved learning analysis
Reduce development time and support cost	Manageability and Supportability– Multi-language and object- oriented development environment enables faster time to application development cycle. Code reuse enables greater efficiency and productivity with fewer support calls from users.	Product development cycle from 6 to 8 months reduced to 6 weeks ⊕ 75% to 80% Available development time reduced from 130 person- days to 30 person- days to 30 person- days ⊕ 77% IT support costs ⊕ 19% Lines of code ⊕ 44% Development costs ⊕ 66% Writing lines of reusable code as .NET Framework components Rapid learning

As shown in Table 1, designers should put into consideration how the platforms they build make use of text, sounds, conceptual diagrams, videos, animations etc to effectively communicate their ideas or purpose to the intended users. Table 1, also shows that Presentation design that is flexible and consistent in usage, clear, concise, aesthetically pleasing, and which effectively integrates the various components of multimedia (text, images, graphics, sounds and videos) in formats that are



suitable for the learning platforms can profitably provide an exhaustive and efficient application that is convenient for learning.

### 3.3 Standards and Compliance Design

In e-learning application design, standard promotes best practice and attitude. This is important especially for system performance and usability. Standards compliance is an essential prerequisite in the development stages of many software applications including e-learning applications. With the rapid advancement in learning technologies, it is appealing to state that there is a lot more too proper and ethical design than simply making use of standard (Travis, 200923). In spite of all ideas, usability standard is still critical and plays an essential role in the development of an intuitive elearning system because the idea of compliance to standard;

- Makes sure consistency is achieved in automated learning systems. The approach provides measures that assist educational multimedia designers in avoiding noise and inconsistencies in user interfaces.
- Determines and make clear good design practice and routine.
- Helps the designers put into consideration the user profiles/requirements and the issues with design of graphical user interface.

Standards compliance is essential in solving some of the technical problems that affect the usability of e-learning products.

According to Travis, (2009), the ISO revealed comprehensive design principles that e-learning application developers must put into consideration to ensure the design of a usable learning platform. The standard consists of five scopes of concept;

- Design decisions in addition to strategic designs - what are the objectives of the elearning product? How does it meet the need and requirement of its intended users? Who are the intended learners and what are their objectives?
- Design of content- how is the content arranged, how would the system handle other issues like offering user-specific

customization and privacy? What is the conceptual model of the system?

- Search and navigation how would the content of the e-learning application be arranged so that learners can navigate the pages with ease? How can the learners search for information and contents of the e-learning product?
- Presentation of the content- how will the individual pages of the product be designed to enable learners in accessing and in making use of the content or information in the system? How will the links be designed?
- General design view how will the developers design for an international audience? How will the system provide help? What is the download time? And is it acceptable?

To accomplish a high level of standard compliance towards solving these common problems; designers of e-learning applications should stick to all applicable standard and should make resources available for the intended users. The quality of how information is described and how it closely confirm to the learning objective is critical in assisting learners advance from searching to discovering (Duval, and Hodgins, (2006); Okoye, et al, 2014a).

As the prevalence of e-learning applications continue to increase, the significance of a utilitarian, consistent and distributive information system continue to increase as well. The compliance to/and consistent use of standardized information representation means of will significantly increase the interactivity of e-learning depositories. With consistency of data and information, searching tends to be more precise and broad, organization of work becomes uniform and simple and there is also efficiency and accuracy in sharing of information. Compliance to standards is a good way of assessing the quality of e-learning systems, and designers must make sure that the combination of resources associated with the systems complies to the international standards and ascertain whether they have completed them with the applicable details and exactness; to allow others make use of the content in evaluating pertinence of the resulting systems.



### 4. USER-CENTRIC DESIGN METHOD TOWARDS AUTOMATED LEARNING

It is evidenced that one of the key challenges in developing automated system for learning is to build effectively represented user profile, learning styles and goal to help support reasoning about each learner (Nganji, et al, (2011); Huang and Shiu, (2012)). It should also be possible to dynamically update the representation of users' performance to take into account the changing state of leaners and the variations in the information that is relevant to each user over time. There is need to consider the fact that there is an additional task of matching such learners (user profiles) with solutions that best fits their particular learning needs/requirement (Nganji, et al, 2013). The work in this paper propose a personalised adaptive learning system that is expected to collect routines and monitor changes in user's behaviour during the learning process. This is to determine which adaptations technique is suitable or may be required progressively through time. The approach as described in Figure 2, is expected to take into account users profile (prior knowledge of learners' background), learning behaviour and actions when using the system. The propose system is an automated learning system that is capable of detecting changing trends in learning behaviours and abilities. The goal is to discover user interaction patterns, and respond by making decisions based on adaptive rules centred on captured user profiles. The approach also focuses on the personalization of learning contents for the individual users. The tactics makes it possible to efficiently generate learning patterns based on the sequence or control-flow individual of each learning patterns/behaviours.

The proposed approach uses learning process adaptation and discovery technique (Okoye et al., 2014a; Van der Aalst, (2011)) to allow for traces not present in an existing learning process to be discovered, by using adaptive rules to generalise and allow for behaviours unrelated to the ones in the learning knowledge-base to be observed. These observations are further enhanced by revealing the most likely underlying learning activity that is not invalidated by the next set of observations. The important aspect of our approach, is that we aim to analyse the sequence of activities to produce the behaviour of a particular learner (user profile) and can be used to extract, transform and load new/more enhanced system for learning capable of increasing learners performance or goal in an unswerving path as shown below;

These Activities as they happen sequentially take a workflow log  $W \subseteq T$  as input and result in a workflow net being constructed.

Workflow Logs  $W \subseteq \mathbb{R}^{d}$  definitive relationship management algorithm: where

W is a finite set of Events.

T is a finite set of transitions where W /T ¼;

Hence,  $W \subseteq T^*$  is a set of directed pattern, called the flow-relation.

Our goal is to create rules of the form;  $X \rightrightarrows Y$ 

IF X THEN Y,

where **X** = Learning pattern (Antecedent) and **Y** = Learning pattern extension (Consequent)

Driven by these variables, it is possible that the following learning path can be suggested to improve the performance or learning pattern of a particular user;

Rules like "Learners that have similar learning characteristics as Learner(X) are most likely to come across LearningActivities(Y) of Learner(X)" can be established. Thus; (forall (?X)

(=> (Learner? X)
 (exists (? Y)
 (And (someActivity? Y)
 (Learning Activity?X?Y)))))

)





Complete?

Send details to Database

Recommend

A Yes

Measure Aftributes

Pattorns & Learning Styles

Matching Alfribult - 7 Jyze Return Best Match



Figure 2: Architectural Diagram of proposed Personalised Adaptive Learning System

No

In figure 2, it is easy to check whether or not all the traces of activities within the learning process are possible in the model. The initial marking of activity is enabled because of the token at the *Start* of the learning activities. The control-flow for its execution  $(X \Rightarrow Y)$  (Okoye et al., 2014a) results in the marking that learner (X) performs an activity that is casually followed by activity (Y), hence, the stronger the relation between Learner X and its learning activities Y. Now this enables the execution of the remaining learning activities to the final event, *End* in the path.

### 5. DISCUSSION

Change

Criteria

To achieve effectiveness in design and development of e-learning systems, designers must stick to all relevant standards and specifications and should make available to the intended users the adequate and specific information required.

Considering a situation whereby different caption is being provided for the same element eg. "goal" vs "objective" the points being raised is that these factors are critical and should be put into consideration by e-learning system developers during the development and evaluation of the system. When the determining factors (accessibility, presentation design and standard compliance) are narrowed down and put into consideration, it is perceived to assist e-learning product designers to procure a fit for purpose usercentered planning and instructional design. We point out that designers have to understand who their targeted audience are before moving into the design as many learners tends to turn away from some e-learning system if they find it difficult to intuitively use the presented platform. As a final point, it may be seen that without clear standards by which usability in e-learning products is



evaluated or measured, it will be difficult for designers to have knowledge on how to ensure the system will be comprehensively usable. We show that the adoption of clear, concise and visible guidelines will help both learners and designers in achieving an exhaustive, intuitive and effective use of learning technologies. Optimistically, with the wide spread of the rate at which these evaluation standards would be adopted for design, we believe a large scale in proportionate of e-learning systems will exhibit rich usability and prove effective for elearning.

### 6. CONCLUSION

This paper ascertain some of the key design criteria necessary towards achieving usability in e-learning systems and revealed variables and measures that are of relative importance when considering the performance and usability of e-learning systems. The outcome shows that the technological response to satisfying the increasing learning demands is by providing a personalized adaptive learning system for the users. From all indication, adoption of these new technological the developments will spark a great success for many learners. It is therefore, of great significance for elearning system developers to maintain a higher degree of performance and usability design standards that is centered on the intended users, when developing their systems in order to ensure learners satisfaction and constancy. As part of its contribution in this area of research, the work in this paper propose an adaptive e-learning system to help address some of the design problems with present e-learning systems in order to improve learners engagement and outcome on performance. This is required and necessary to provide continuous intelligent recommendation, guidance and feedbacks on learner's performance as well as in achieving their learning goal.

Future research can be conducted by replicating the present study to uncover the perception of performance and usability in various segment of digital platforms; ranging from e-commerce to ebusiness, e-work to e-marketing to e-government etc.

- 1. Adriansyah, A. et al, (2011). Conformance Checking using Cost-Based Fitness Analysis. Proceedings of EDOC 2011 IEEE Computer Society, pp. 55-64
- 2. Ali, L. et al, (2008). E-accessibility of higher education websites. Presented at the ECEL Seventh European Conf. on e-Learning, pp. 6–7
- 3. Calder's, T. et al, (2009). Using Minimum Description Length for Process Mining. Proc. of SAC 2009. ACM Press '09. pp. 1451–1455
- Cook, J.E. and Wolf, A.L., (1999). Software Process Validation: Quantitatively Measuring the Correspondence of a Process to a Model. Journal of ACM Transactions on Soft. Engr & Methodology. Vol. 8(2), pp. 147–176
- Duval, E. and Hodgins, W., (2006). Standardized uniqueness: Oxymoron or vision for the future? Journal of Computer. Vol. 39(3), pp. 96–98.
- Fahland, D. and Van der Aalst, W. M. P., (2012). Repairing process models to reflect reality. *In: Business Process Management 2012*, vol. 7481 of Lecture Notes in Computer Science, Springer. pp. 229–245.
- Huang S.L. and Shiu J.H., (2012) A User-Centric Adaptive Learning System for E-Learning 2.0. International Journal of Educational Technology & Society, Vol. 15 (3), pp. 214–225
- IMS guidelines for developing accessible learning applications, version 1.0. IMS Global Learning Consortium, (2006). [online] (Available at: http://www.imsglobal.org/accessibility/accv1po /imsacc\_guidev1po.html [Accessed 17 Dec 2014].
- 9. Leacock, T. L. and Nesbit, J. C., (2007). A Framework for Evaluating the Quality of Multimedia Learning Resources. *Educational Technology* & *Society*, Vol. 10 (2), pp 44-59.
- 10. Mayer, R. E. and Moreno, R., (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, vol. 38(1), pp. 43–52.

#### 7. **REFERENCES**



- 11. Munoz-Gama, J. and Carmona, J., (2011). Enhancing Precision in Process Conformance: Stability, Confidence and Severity. *Proc. of CIDM. IEEE*, 2011.
- 12. Murata, T., (1989). Petri Nets: Properties, Analysis and Applications. Proc. of the IEEE, vol.77 (4), 541–580
- Nganji, J.T. et al, (2011). Ontology-Based E-Learning Personalization for Disabled Students in Higher Education. *Journal of ITALICS*, Vol. 10(1) pp. 1473-7507
- Nganji, J.T. et al, (2013). Ontology-driven disability-aware e-learning personalization with ONTODAPS. Journal of Campus-Wide Information Systems, Vol. 30(1), pp.17 – 34
- 15. Nia, W.T. and Kay, J., (2010). PERSONAF: personalized Framework for ontological reasoning in pervasive computing. International Journal of User Modelling and User-Adaptation., Vol. 20(1), pp.1–40. [Online] (Updated 2010) Available at: http://link.springer.com/content/pdf/10.1007%2 Fs11257-009-9068-2 [Accessed 5th Nov, 2014].
- Okoye, K. et al (2014). Accessibility of Dynamic Web Applications with Emphasis on Visually Impaired Users. *IET Journal of Engineering, JOE*, pp. 7 [Online] Available at: <u>http://digitallibrary.theiet.org/content/journals/10.1049/joe.</u> 2014.0136
- 17. Okoye, K. et al, (2014a). A Semantic Rule-Based Approach Supported by Process Mining for Personalised Adaptive Learning. In Journal of Procedia Computer Science, Elsevier, Vol. 37(c), pp. 203-210. In proceedings of the 5<sup>th</sup> International Conference on Emerging Ubiquitous Systems and Pervasive Networks, EUSPN, Halifax, Canada, 22-25 Sept. 2014. In UEL Research Open Access Repository ROAR <u>http://roar.uel.ac.uk/3920/</u> [Online] Available at:

http://www.sciencedirect.com/science/article/ pii/S187705091400996X

18. Parmanto, B. and Zeng, Q., (2005). Metric for web accessibility evaluations. *Journal of Society* 

for Information Science and Technology, Vol. 56(13) pp. 1394-404.

- 19. Rozinat, A. and Van der Aalst, W. M. P., (2008) Conformance Checking of Processes based on Monitoring Real Behaviour. Journal of Information Systems, Vol. 33(1), pp. 64–95
- 20. Saldaño, V. et al, (2013). Web accessibility for older users: a southern argentinean view. Proc. of the Eighth Int. Conf. on Software Engineering Advances ICSEA, 2013.
- 21. Stephen P. N., (2012). Reading for Evidence and Interpreting Visualizations in Mathematics and Science Education. General Education, English Humanities, Social Sciences & Law, SensePublishers, 2012.
- 22. The ISO Guidance on World Wide Web user interfaces [Online] Available at: http://www.iso.org/iso/iso\_catalogue/catalogu e\_tc/catalogue\_detail.htm?csnumber=37031 [Accessed 29 Dec 2014.]
- 23. Travis D., (2009). Bluffers' Guide to Usability Standards. ISO 9241. [7th Ed.] Userfocus Itd 2003-2009. Retrieved Dec 22, 2014 from the World Wide Web: http://www.userfocus.co.uk/articles/ISO9241\_u pdate.html
- 24. Tufte, E. R., (1997). Visual explanations: Images and quantities, evidence and narrative. Cheshire, CN: Graphics Press 1997.
- 25. Van der Aalst W.M.P., (2011) Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer, Berlin, 2011.
- 26. Van der Aalst W. M. P. et al, (2011). Time Prediction Based on Process Mining. *Journal of Information Systems*, Vol. 36(2), pp. 450–475
- 27. Weerdt, J.D. at al, (2011). A Robust F-measure for Evaluating Discovered Process Models. *Proceedings of CIDM,* IEEE, pp. 148–155
- 28. Yu Z. et al, (2011). Ontology-Based Semantic Recommendation for Context-Aware E-Learning. Journal of Ubiquitous Intelligence and Computing, vol. 4611, pp 898-907.





### SECURE AUTOMATED TELLER MACHINE (ATM) USING FINGERPRINT AUTHENTICATION AND SHORT-CODE MESSAGE IN A CASHLESS SOCIETY

\*J. B. Awotunde<sup>1</sup>, R. G. Jimoh<sup>2</sup> and O. E. Matiluko<sup>3</sup>

<sup>1</sup>University School, University of Ilorin, Ilorin <sup>2</sup>Department of Computer Science, University of Ilorin, Ilorin <sup>3</sup>Center for Systems and Information Services, Landmark University, Omu-Aran

<sup>1</sup>jabonnetbylinks@gmail.com, <sup>2</sup>jimoh\_rasheed@yahoo.com, <sup>3</sup>matiluko.opeyemi@lmu.edu.ng

#### ABSTRACT

In cashless policy, security is vital to protect customers' funds. One major problem in the world of the cashless economy is Internet related fraud. Nigeria is at the center of electronic related fraud and this can only be expected to increase as we march into the cashless economy. Another major problems people face is the loss of the password and not remembering it again, these have cause a lot of damage to many people and organizations. In order to fight or reduce the aforementioned problem a secure automated teller machine using fingerprint and short-code system was proposed.

The proposed system demonstrated a three-tier architectural structure, the verification system which centered on the enrolment, enhancement, feature extraction and matching of fingerprints after which a short-code will be send to the holders phone number. The backend database system that serves as storeroom of the templates of all ATM account holders' pre-registered fingerprints. The results obtained confirm that the proposed approach could significantly reduce ATM fraud if not totally eradicate it.

Keywords: ATM Machine, Authentication, Account holder, Minutiae, Bank verification number

### 1. INTRODUCTION

Information Technology plays a significant role in bringing about sustainable development in every aspect of life of any nation. Without a determined use of Information Technology, no country can realize a speedy socio-economic growth and development. The evolution of a cashless economy will increase a lot of security issues in our financial establishment.

In everyday fast insecure world, the need to maintain a well secure Automated Teller Machine is

increasingly significant and increasingly complicated. Fraudulent, hackers and identity theft are beginning to spread from states to states, from countries to countries and thus a proper security approach needs to be adopted by the government in other to protect the account holders and their money. It have been proved by my researchers that the perpetrators of this acts make used of stolen card, password and pin of bank customers and abandon it after finish using it for their evil acts.



A cashless society is a culture where no one uses cash, all purchases being made are by credit cards, charge cards, cheques, or direct transfers from one account to another through mobile banking and other electronic systems. The cashless society anticipated refers to the widespread application of computer technology in the financial system (NSACC, 2011). In the wider society, effective implementation of the policy would curb corruption in all forms of transactions. The expert also expressed optimism that Nigeria would truly move on progressively as a cashless society (NSACC, 2011). ATM fraud has been very common in all banks and the problems has created fear in many costumers heart that they prefer going to the bank to collect their money than to use an ATM, also many people are illiterate that they don't know how to use the ATM card. So, the main aim of introducing ATM machine has not been met because you still find long queues in almost all banks in Nigeria.

In line with this, the apex bank, Central Bank of Nigeria (CBN) has announced the introduction of Bank Verification Numbering (BVN) in all our banks and this is to increase the security of ATM machine and other related systems. The BVN gives each bank customer a matchless identity across the banks, and can be used for easy identification and verification of account holder.

The BVN elucidation is to ensure accountability, protect bank customers' account from illegal access, reduce exposure to fraud, check identity theft, enhance credit advancement to bank customers, and also encourage financial attachment. The BVN will promote a safe and sound financial system in Nigeria, especially as it will keep records of suspected falsified individuals in the banking system. It will make life and banking operations easy for bank customers as it will be accepted as a means of identification across all banks in Nigeria. This will improve speed of service and shorten queues in our banks. It involves capturing customers' fingerprint and signature, among others, at the account holder banks, the BVN enables each individual to have one credentials within the banking system and gives each customer maximum protection and security of transactions. Biometric technologies have been used to evaluate human characteristics as an enriched form of authentication for immediate security improvements. Biometric refers to identifying an individual based on physiological or behavioral characteristics such as fingerprint, face recognition, signature, voice among others. The customers unique BVN is accepted as a means of identification across all banks in Nigeria. Each customer can only enroll once in any banks that he/she has account, while his/her BVN will be linked to all his bank accounts across Nigeria banks (NSACC, 2011).

The apex bank had also taken a giant step to gain the confidence of ATM consumers following the circular enforcing migration from Magstripe type of debit card to chip and Pin (EMV compliance) type of debit card. Statistics show that this effort has reduced the fraud incidences by 70 per cent. Many customers are now embracing these electronic (ATM and PoS) channels in their transactions because of near-impossible efforts of would-be fraudsters to clone debit cards to perpetrate fraud as was the case during the pre-migration era. Interswitch also helps customers with the availability of its e-payment solutions such as Paydirect, Autopay, Direct Debit, Verve Card, Quickteller, Webpay and Smartgov". ``The Biometric authentication for POS and ATMs to address safety of customers' fund and avoid losses through compromise of PIN is being considered by the apex bank and it will be implemented by 2015, (Leadership New Paper, 10-07-2013).

Fingerprint based authentication is a prospective contestant replace password-based to authentication. Among all the biometrics, fingerprint based identification is one of the most mature and proven technique. At the time of transaction fingerprint image is acquired at the ATM terminal using high resolution fingerprint scanner. Security measures at banks can play a critical, contributory role in preventing attacks on customers (Kuykendall, et al., 2003). Because fingerprint-based authentication offers several advantages over other authentication methods, there has been a significant surge in the use of biometrics for user authentication in recent years (Akwaja, 2010).

The cashless economy policy will help the security agencies to control the menace of insecurity in the country by reducing the sequence of armedrobbery, because thieves will not attack people's



houses when they are aware that people are no longer keeping money at home. This policy will not only help in making Nigerian economy cashless, but as the axiomatic silver capsules it represents, would assist in resolving other socio-economic problems troubling Nigerian as a nation.

When there is a massive reduction in the movements of huge cash, it will invariably help the apex bank to checkmate the surge of inflation, because whenever there is a reduction in bulk cash transactions; definitely inflation rate is going to subside. Cashless policy will also help financial crime commissions like EFCC and ICPC to fight financial crimes by making money laundering to be expensive because of charges that apply when making lodgments or withdrawals above the stipulated limits of N500,000 and N3 million for individual and corporate accounts, respectively. Therefore, those who are in the ugly habit of carrying huge money around will now have to pay for their excesses, especially Nigerian politicians who convert their cars into bullion vans and bedrooms to strong rooms, mainly during the electioneering campaigns(NSACC, 2011).

The paper proposed a convenient cost effective system that uses fingerprint and a short-code message to authenticate account holders transaction.

The remainder of the paper is organized as follows: Section two provides a literature review on state of the art review on automated teller machine using fingerprint verification and related work on shortcode SMS. Section 3 presents the overview of the method. In section four the paper outlines the implementation process and finally, section five discuss conclusions and the future work.

### 2. RELATED WORK

### 2.1 Automated Teller Machine using Fingerprint Verification

For customers to really embrace the use of ATM for their major transactions the issue of ATM security must be taken with all seriousness. ATM cards must be very secure even when the owner misplaced or lost the card. This will prevent any attacker from using the card on any ATM machine. Since security measures at ATM centers play a significant role in preventing attacks on customers money, several researches have proposed the used of fingerprint in a like manner of this paper, to shift from PIN to biometric based security. Fingerprinting has been the most widely used during the 20th century. The maturity of biometric techniques and generally the dramatic improvement of the captured devices have led to the proposal of fingerprinting in multiple applications but in the last years, minutiae have been the main type of algorithm used. The minutiae are relatively stable and robust to contrast, image resolution and global distortion as compared to other fingerprint representation (Fatai, et al., 2014). (Jeroen, et al., 2011) provided a better understanding of the benefits and limitation of integration of biometrics in a PIN-base payment authentication system. Based on their review they proposed a biometric that can be integrated in a PIN-based authentication infrastructure by binding a fixed binary, renewable string to a noisy biometric sample. The South African Social Security Agency (SASSA) has introduced a new SASSA Payment Card that has a fingerprint authenticated features. The card is a SASSA-branded smart payment MasterCard, which has an embedded chip containing personal details, fingerprint and secret PIN, with the card the customers can easily withdraw and make payment at point-of-sale (POS) center, purchase airtime, pay water and electricity bill from the accounts, or open accounts (SASSA, 2013). (Fengling, et al, 2005) proposed a smartcard based encryption/authentication scheme for ATM banking system. The first layer of the scheme is used to perform authentication based on available information on the smartcard. Fingerprint based authentication via feature and minutiae matching then followed on the second layer. (Das and Jhunu, 2011) focused on vulnerabilities and the increasing wave of criminal activities occurring at ATMs and presented a prototype fingerprint authentication for enhancing security. The systems adopt the same measure as the current work by formulating modules for fingerprint enrolment, enhancement, feature extraction and database and matching. (Santhi and Kumar, 2012) proposed an ATM security enhancing method with secured Personal Identification Image (PII) process. A detailed study on various existing biometric systems is also presented stating the strengths and limitations.



(Bhosale and Sawant, 2012) and (Ibiyemi, et al., 2012) present ground-breaking models for biometric ATMs which replaces card system with biometric technology. The proposed systems hybridize feature-based fingerprint, iris and PIN to provide reliable and fool-proof ATM authentication. (Mali, et al., 2012) provided a network security framework for real time ATM application using a combination of PIN, thumb scanning and face recognition to foster security. The proposed framework is expected to register thumb and face features to be stored at a server side in encrypted format. Authentication is done by decrypting patterns from database, and matching with input pattern before access is granted for ATM operations. The integrated system uses Principal Component Analysis (PCA) and Eigen algorithm for face recognition, LSB algorithm for stegnogaphy and AES algorithm for cryptography. Though the framework looks promising, its practicality is not supported by detailed implementation and evaluation. (Abayomi-Alli, et al., 2012) proposed an enhanced e-banking system where customer can access multiple accounts over different banks institutions with a single ATM card with fingerprint authentication. A match-on-card technique was used that relies on a one-to-one matching where the data from the ATM fingerprint sensor is compared only to the template stored on the user's ATM card. This will help in privacy concern of users; the system will also help the users to have access to multiple accounts with a single ATM card. It is secured and help in reducing ATM fraud. The paper used the characteristic features of fingerprint to overcome the limitations of the PIN based ATM authentication. However, the proposed method presented adequate implementation and evaluation to back-up the performance claim. The proposed system is different from others approaches because it makes use of the UML modeling in designing the system, used a three-tier architectural structure and minutiae for the extraction of the fingerprint.

### 2.2 Related Work on Short-Code SMS

(Vandana and Deepali, 2012) proposed a system GSM Modem Based Data Acquisition that is a process control system. The system help in gathered data from various processes present at distant places. It helps to scrutinize and monitor parameters like temperature, rainfall, humidity etc. With the proposed system the service personnel need not to visit distant sites for data collection .The data is collected automatically formed a data base and stored in a computer. Information gathered from the computer server can be used in industries as well as in home automation.

(Ramamurthy, Bhargavi, and ShashiKumar, 2010) proposed the development of a Low-Cost GSM SMS-Based Humidity Remote Monitoring and Control system for Industrial Applications. They proposed a wireless solution, based on Global System for Mobile Communication networks for the monitoring and control of humidity in industries. This system provides ideal solution for monitoring critical plant on unmanned sites.

(Ghose, et al, 2011) presented the design and development of microcontroller based SMS gateway for GSM Mobile. In their work, a microcontroller based SMS gateway for GSM mobile has been designed and developed. Most of the SMS gateway system was controlled by PC based software where microcontroller only used for controlling and sending status of devices or any appliances connected with the system. Successful completion of the design and testing of the SMS Gateway indicates that the PC as an SMS gateway can easily be replaced by a PIC microcontroller. Beside this, the additional IC, MAX232, used for voltage adjustment between the mobile and PC is no longer needed in the proposed micro-controller based system. It also reduces the complexity and the overall development cost of such a system.

(Oke, et al., 2013) proposed a system that used GSM module for receiving short message service (SMS) from user's mobile phone that automatically enable the controller to take further action like switching ON and OFF electrical appliances such as fan, airconditioner, light etc. The system was integrated with microcontroller and GSM network interface using C language. MPLAB software was utilized to accomplish the integration. The system is activated when user sends the SMS to the controller at home (regarded as Smart Home). Upon receiving the SMS command, the microcontroller unit then automatically controls the electrical appliances by switching ON or OFF the device according to the user's order.



In other word, it reads message from the mobile phone and respond to control the devices according to the received message.

(Pany and Das, 2011) proposed an Embedded Automobile Engine Locking System, Using GSM Technology, the proposed system deal with the design & development of an embedded system, which is being used to prevent /control the theft of a vehicle. The developed instrument is an embedded system based on GSM technology. The instrument is installed in the engine of the vehicle. An interfacing GSM modem is also connected to the microcontroller to send the message to the owner's mobile. The main objective of this instrument is to protect the vehicle from any unauthorized access, through entering a protected password and intimate the status of the same vehicle to the authorize person (owner) using Global System for Mobile (GSM) communication technology.

### 3. METHODOLOGY

Account holders' fingerprints were captured using SecuGen fingerprint optical scanner and Unified Modeling Language (UML) was used for the formulation of the fingerprint biometric model to show the flow of the proposed system. The model was simulated using Java Programming Language. In addition, customers' details as well as the template of account holder fingerprint were stored using MySQL database management system. The Minutiae-based was also adopted due to its acceptance and it is a well-known method for fingerprint verification. It is the most prevalent ones being included in almost all existing fingerprint identification and verification systems. Minutiae-base represents the fingerprint by its local characteristics like, terminations and bifurcations called minutia. Minutiae are small points of interest in the fingerprint image.

### 3.1 Framework of the Proposed System

The proposed system consists of the different modules as shown in Figure 1.

### 3.1.1 Fingerprint Image Acquisition

A fingerprint is a unique pattern of ridges and valleys on the surface of a finger of an individual. A ridge on the fingerprint is defined as a single curved segment, and a valley is the region between two adjacent ridges. Minutiae points are the local ridge discontinuities, which are of two major types: ridge endings and bifurcations. It was revealed that a good quality image has around 40 to 100 minutiae (Roli, Priti and Punam, 2011). This fingerprint is captured using SecuGen fingerprint optical scanner sensors.

### 3.1.2 Fingerprint Features

Fingerprint feature can be categorized into two, which are global ridge pattern and local ridge pattern. Global ridge pattern relies on the ridge structure, global landmarks and ridge pattern characteristics. Fingerprint ridge pattern, which consist of right loop, left loop, whorl, arch and tented arch. Equally, local ridge patterns are the discontinuities of local ridge structure referred to as minutiae. The proposed system was executed using minutiae fingerprint feature extraction method for extracting the fingerprint features of card holder for the purpose of template generation and subsequent identification.






#### 3.1.3 Fingerprint Image Enhancement

Enhancement of the fingerprint image is frequently employed to reduce the noise and improve the definition of ridges against valleys. The commonly use approach is the Gabor filtering which has four main stages namely: normalization, ridge orientation estimation, ridge frequency estimation and filtering. This work used the approach employed by (Raymond, 2003) by including three additional stages with the four stages stated above which are segmentation, binarization and thinning for the fingerprint image enhancement.

### 3.1.4 Fingerprint Feature Extraction

Crossing Number (CN) approach was employed since it the most commonly used method of minutiae feature extraction (Roli et al., 2011; Raymond, 2003). This method extracts the ridge endings and bifurcations from the skeleton image. The minutiae features are extracted by scanning the local neighborhood of each ridge pixel in the image using a 3 x 3 window. The CN value is then computed, which is defined as half the sum of the differences between pairs of adjacent pixels.

### 3.1.5 Fingerprint Template Generation

This phase cuttings the minutiae features of the fingerprint image and generates a template from these features. The procedure involves defining a



window of size 3x3 pixels fixed at the black pixel. The algorithm finds the number of pixels, N, within the window which determines the feature of the minutiae. The template extracted is stored into the database along with other card holder biographical information.

### 3.1.6 Fingerprint Template Matching

For efficient matching process, the extracted data is stored in the matrix format. The data matrix is as follows. Number of rows: Number of minutiae points. Number of columns: 4

Column 1: Row index of each minutia point. ( $^{\times}$  Coordinate)

Column 2: Column index of each minutia point. ( $^{>}$  Coordinate)

Column 3: Orientation angle of each minutia point.

(Minutiae angle <sup>b</sup> of the particular minutiae point to be paired i.e. input image and template image).

Column 4: Type of minutia. (A value of '1' is assigned for termination, and '3' is assigned for bifurcation). In other words  $(x_{j}y_{j}, \theta)$  and the minutiae type of input template would be paired

with the same of  $X_1 y, b$  and the minutiae type of registered template).

During the matching process, each input minutiae point is compared with template minutiae point

considering the above properties i.e.  $(x, y, \theta)$  and the type of minutiae). In each case, template and input minutiae are selected as reference points for their respective data sets. The reference points are used to convert the remaining data points to polar coordinates. Equation (1) is used to convert the template minutiae from row and column indices to polar coordinates.

$$\begin{bmatrix} \eta_{k}^{T} \\ \emptyset_{k}^{T} \\ \theta_{k}^{T} \end{bmatrix}_{=} \begin{bmatrix} \sqrt{\left(row_{k}^{T} - row_{ref}^{T}\right)^{2} + \left(col_{k}^{T} - col_{ref}^{T}\right)^{2}} \\ tan^{-1} \left(\frac{row_{k}^{T} - row_{ref}^{T}}{col_{k}^{T} - col_{ref}^{T}}\right) \\ \theta_{k}^{T} - \theta_{ref}^{T} \end{bmatrix}$$
.....(1)  
$$\mathbf{r}_{k}^{T} = \text{Radial distance of kth minutiae.}$$

 $\theta_{k=}^{T}$  Orientation angle of Kth minutiae RowTref, colref T= row index and column index of reference points currently being considered. Similarly the input matrix data points are converted to polar coordinates using the Equation (2)

$$\begin{bmatrix} r_{m}^{i} \\ g_{m}^{i} \\ g_{m}^{i} \\ g_{m}^{i} \end{bmatrix} \begin{bmatrix} \sqrt{\left(row_{m}^{i} - row_{ref}^{i}\right)^{2} + \left(cvl_{m}^{i} - col_{ref}^{i}\right)^{2}} \\ tan^{-1} \left(\frac{row_{m}^{i} - row_{ref}^{i}}{col_{m}^{i} - col_{ref}^{i}}\right) + rotatevalues(k, m) \\ \theta_{m}^{i} - \theta_{ref}^{i} \end{bmatrix}$$
.....(2)

Rotate values (k, m) represents the difference between the orientation angles of Tk and Im Tkand Im represent the extracted data in all the columns of row k and row m in the template and

### 3.1.7 Template Database Storage

input matrices, respectively.

This part stores data that is, all the templates and information that are been generated from the process of minutiae extraction and rotation and displacement of image. Fingerprint templates' and information that been generated from the extraction and rotation and displacement of image were stored using MySQL database management system.

### **Minutiae Algorithm**

Input: Gray-scale fingerprint image.

Output: Verified fingerprint image with matching score.

- ✓ Fingerprint is binarized
- Thinning on binarized image
- ✓ Minutiae points are extracted. Data matrix is generated to get the position, orientation and type of minutiae.
- ✓ Matching of live test fingerprint with template
- ✓ Matching score of two images is computed, if matching score is greater than the threshold then it stops else it returns to enrollment

### 4.0 IMPLEMENTATION PROCESS

The process of enrolment involve the card holder opening an account and register with the bank of their choice this will enable the bank to have all the enrollee's information and all necessary details that concern the enrollee and take the biometric data



captured of the person that own the account and store in the database, which will be used later for the process of verification and further update of information.

The process of extraction and verification make used of minutiae-base techniques. This is to obtain an efficient and thoughtful result in order to reduce or eradicate the problems which is associated with the use of card-PIN and high rate insecurity people faced in using ATM machine.

### 4.1 Enrollment Process

Before an Account holder being identified or verified by a biometric device, the enrollment process must be completed. The aim of this enrollment process is to create a summary profile of the user (Card Holders'). The process consists of the following:

### 4.1.1 Bio data

This comprises the Following: Surname, First Name and Last Name which take alphabetic characters, Account Type: Current, or Saving this also take alphabetic characters, picture of the enrollee which can take binary characters, Nationality of the enrollee take alphabetic and string characters, date of birth take string characters and the date account was issue take string characters too.

### 4.1.2 Fingerprint Image Capture

The Account Owner fingerprint will be captured with fingerprint scanner for a minimum of two or three biometric readings, by placing a finger in a fingerprint reader. Not all the samples will be stored; the technology analyzes and measures various data points unique to each individual. The number of measured data points varies in accordance to the type of device.

### 4.1.3 Minutiae Feature Extraction from Image

This is where the minutiae extraction is done and of course processes like binarization, thinning and bifurcation would be done have a perfect minutiae feature extraction from the image.

### 4.1.4 Rotation and Displacement of Image

This is where the image is normalized to get an authentic and effective image to be stored in the database, which aids the process of matching.

### 4.1.5 Template Database Storage

This part stores all the templates and information that are been generated from the process of minutiae extraction and rotation and displacement of image.

### 4.1.5 Conversion and Encryption

The Account Owner measurements and data points are converted to a mathematical algorithm and encrypted. These algorithms cannot be reversed to obtain the original image. The algorithm may then be stored as a user's template in the database servers and on the ATM card

### 4.1.6 The Enrollee Storage

This has all the details of all the people that have been enrolled and its stores them with the account number. when there is need to view enrollee's details or make amends this can easily be done with the use of account number to trace individual's details and it makes the process of verification easier and faster as it saves time (Figure 2)





Figure 2: Flowchart for Enrollment process

### 4.2 IDENTIFICATION AND VERIFICATION PROCESS

Once the account holder has been enrolled in a system; he/she can start to use biometric technology to have access to his/her account via the ATM machine or related system to authorize transactions.

### 4.2.1 Identification

This is a one-to-many match. The user provides a biometric sample and the system looks at all user templates in the database. If there is a match, the user is granted access, otherwise, it is declined.

### 4.2.2 Verification

This is a one-to-one match it requires the user to provide his/her identification such as a PIN and valid ATM card in addition to fingerprint. The account holder is to establish who he/she is and the system simply authenticates if this is correct. The biometric sample with the provided identification is compared to the previously stored information in the database. If there is a match, access is provided, otherwise, it is declined Figure 3, while Figure 4 shows the architectural design of the proposed system.

### 4.2.3 Authentication

This is the part of the database that allows access into and out of the database, this part monitors the kind of people that uses the database and controls unwanted users and unnecessary logins and access into the database for more secured and protected database environment. For security purposes servers must also address the problem of authentication. In a networked environment, an unauthorized client may attempt to access sensitive data stored on a server. Authentication of clients is handled by using cryptographic techniques such as public key encryption or special authentication servers such as in the OSF DCE system.

### 4.2.4 Short-Code Message

After the verification and identification process, the system send a short-code inform of SMS to the phone of card holder, which in turn the card holder will enter the short-code. If the shortcode is enter correctly access will be given to the card holder to make all necessary transactions with the card.







Figure 4: Architectural design of the system

### <u>KEY</u>

- DMZ Demilitarized zone
- **Z9PE** D8- This is used for more reliable
- network
- VPN Virtual Private Network

### 4.3.0 Account Details Capture

This is where data collected from account holder are been collected and stored in the database i.e. extraction and matching process.

### 4.3.1 VPN

This is known as virtual private network, it is a dedicated network for the system that protects the data been captured and stored in the database for present and future purposes.

### 4.3.2 Fire Walls

This is a kind of security that helps protect the network from spam ware and other attacks like hacking, virus attacks etc. of the proposed system. A firewall protects networked computers from intentional hostile intrusion that could compromise confidentiality or result in data corruption or denial of service. It may be a hardware device or a software program running on a secure host computer. In either case, it must have at least two network interfaces, one for the network it is intended to protect, and one for the network it is exposed to. A firewall examines all traffic routed between the two networks to see if it meets certain criteria. If it does, it is routed between the networks, otherwise it is stopped. A firewall filters both inbound and outbound traffic. It can also manage public access to private networked resources such as host applications. It can be used to log all attempts to enter the private network and trigger alarms when hostile or unauthorized entry is attempted. Firewalls can filter packets based on their source and destination addresses and port numbers.

### 4.3.3 Collation Server

This is responsible for distribution of information to all clients on the network and even share of



## 12<sup>th</sup>

### International Conference

resources and other necessary things that are needed on the network.

### 4.3.4 Z9PE-D8

This provides additional layer of security to the proposed system's local area network, it provides a good and secured motherboard for more reliable networking environment as it has in built dual Intel, Ethernet that leads to lower CPU utilization and yet it is very affordable

### 4.4 Minutiae Matching

Once the minutia feature extraction is done the next phase is to compare the live template with the stored template, the system fetches the template from the template database storage and compare it with the live template. Once the matching is done, the matching score from the templates is computed (Fatai, et al., 2014).

### 4.5 Short-Code Message

Once the identification and verification of an account holder is established, the machine will send a short-code to the phone of the customer which in turn enter the code. If the code enter is correct the transaction will be allow if not the ATM will rejected the card.

### 5.0 CONCLUSION

ATM has provided evidence that it is useful in reducing queue in our banks, since you don't have to presently be in bank before you can withdraw your money. It can also be used to make others transactions like Point-of-Sale (POS), on-line transaction and a lot of other benefits. It is 24/7 hours operations which enable the account holder to withdraw at anytime and anywhere. But notwithstanding the above advantages the account holders has not embraced the use of ATM due to it security challenge. This paper tried to find a way to eradicate or reduce the insecurity associated with the use of ATM machine by introducing a Smart card-based ATM with fingerprint authentication and shortcode system to ameliorate these challenges. The proposed system is cost-effective and much secured compared with the PIN-base ATM card. The proposed system was developed using programming language, the java programming language was used because of its platform independence, scalability, easy integration, implementation and upgrade. The system can run on any operating systems e.g. Windows, Linux etc. with .NET framework and MySQL. These tools were chosen for their coding versatility, friendliness and compatibility. The proposed system is much secured and will reduce the ATM machine theft if not totally eradicate theft associated with the use of ATM machine. Furthermore, mobile phones can be found in almost every part of the world in both urban and rural areas, therefore the system provided a better authentication and identification with the use of smart-phones to further enhanced the system.

### 5.1 FUTURE WORK

Future researchers can work on how to do away completely with PIN-Card authorization by the introduction of Bimodal biometrics like Palm and Finger vein, fingerprint and face recognition authentication which is very fast, accurate and difficult to contravened.

### 6 REFERENCE

NSACC (2011): Imperatives of advancing a cashless Nigeria. Nigerian South African chamber of commerce. Retrieved on January 23, 2012 from http://nsacc.org.ng /imperatives-of-advancinga-cashless-nigeria/

http://leadership.ng/news/100713/cashless-policycbn-introduce-biometrics-atms-pos-

2015#sthash .lo15wd2G.dpuf Leadership New paper, 10-07-2013.

- Kuykendall, Lavonne and W. A. Lee. 2003. "Intermediary Risk? Card Hack Puts OSOs in the Hot Seat," American Banker, February 21.
- Akwaja Chima (2010). "Nigeria Connects 99 million Subscribers", Fin. Standard., 10: 15-512.
- Fatai, O.W., Awotunde, J.B. and Matluko, O.E (2014) A Novel System of Fingerprint Recognition Approach for Immigration Control, IOSR Journal of Computer Engineering (IOSR-JCE) 2278-8727Volume 16, Issue 3, Ver. III (May-Jun. 2014), PP 39-42.
- Jeroen B., Ileana B., Koen G., and Emile K. (2011) Evaluation of a template protection approach



to integrate fingerprint biometrics in a PINbased payment infrastructure.

- The South African Social Security Agency (SASSA) (2013) A leader in the delivery of social security services. Re-register to get yours: SASSA HOUSE 501 Prodinsa Building Cnr Steve Biko and Pretorius Street Pretoria.
- Wang Y, Hu J, and Phillips D (2007) A fingerprint orientation model based on 2D Fourier expansion (FOMFE) and its application to singular-point detection and fingerprint indexing. IEEE Trans Pattern Anal Mach Intell 29: 573-585.
- Fengling H, Jiankun H, Xinhuo Y, Yong F, and Jie Z (2005) A novel Hybrid Crypto- Biometric Authentication Scheme for ATM Based Banking Applications. Lect Notes Comput Sci 3832: 675-681.
- Das SS, and Jhunu D (2011) Designing a Biometric Strategy (Fingerprint) Measure for Enhancing ATM Security in Indian E-Banking System. International Journal of Information and Communication Technology Research 197-203.
- Santhi B, and Kumar RK (2012) Novel Hybrid Technology in ATM Security Using Biometrics. Journal of Theoretical and Applied Information Technology 37: 217-223.
- Bhosale ST, and Sawant BS (2012) Security in e-Banking via Card-less Biometric ATMS. International Journal of Advanced Technology & Engineering Research 2: 9-12.
- Ibiyemi T. S, Obaje S. E, and Badejo J (2012) Development of Iris and Fingerprint Biometric Authenticated Smart ATM Device & Card. 24th National Conference of the Nigeria Computer Society (NCS).
- Mali P, Salunke S, Mane R, and Khatavkar P (2012) Multilevel ATM Security Based On Two Factor Biometrics. IJERT 1: 8.
- Abayomi-Alli A., Omidiora E.O., Olabiyisi E.O., and Ojo J.A. (2012) Enhanced E-Banking System with Match-On-Card Fingerprint Authentication and Multi-Account ATM Card. The Journal of Computer Science and Its Aplication, An International Journal of the Computer Society of Nigeria (NCS), Vol. 19, No.2 December, 2012.
- Vandana and Deepali (2012). GSM Modem Based Data Acquisition System, International Journal

Of Computational Engineering Research (ijceronline.com) Vol. 2 Issue.5 Pg1662-1667.

- Ramamurthy B., Bhargavi S., and ShashiKumar R., (2010). "Development of a Low-Cost GSM SMS-Based Humidity Remote Monitoring and Control system for Industrial Applications", International Journal of Advanced Computer Science and Applications, (IJACSA), Vol. 1, No. 4: 20 – 26 http://ijacsa.thesai.org.
- Ghose S.,Rahman Md. S., Sharmin D.,Hussain I.,Yousufzai T.K., (2011). "Design and Development of Microcontroller Based SMS Gateway for GSM Mobile", International Journal of Advanced Engineering Sciences and Technologies (IJAEST). Volume No. 2, Issue No. 1: 090 – 098, Bangladesh, Dhaka.
- Oke A. O., Emuoyibofarhe J. O., Adetunji A. B., (2013). "Development of a GSM based Control System for Electrical Appliances", International Journal of Engineering and Technology Volume 3 No. 4: 443 – 448. IJET Publications UK.
- Pany J. K. and Das Choudhury R. N., (2011). "Embedded Automobile Engine Locking System, Using GSM Technology", International Journal of Instrumentation, Control and Automation (IJICA) Volume-1, Issue-2: 49 – 53.



### OBJECT CLASSIFICATION IN VIDEOS USING NEURAL NETWORK: A CASE STUDY OF VEHICLES AND PEDESTRIANS CLASSIFICATION FOR SURVEILLANCE SYSTEMS

Akintola K.G., Akinyokun O.C., Olabode O., Iwasokun G.B.

Computer Science Department, Federal University of Technology Akure Ondo StateNigeria

### ABSTRACT

Moving object classification is a requirement in smart visual surveillance systems as it allows the system to know the kind of object in the scene and be able to recognize the actions the object can perform. This paper presents a neural-network machine learning approach for real time object classification in videos. This is necessary for the higher layer surveillance systemto detect actions being performed by the moving objects. Fast kernel density estimation background subtraction algorithm is used for object segmentation. Distance signal features are then extracted from the silhouettes of the detected objects. The distance signals features are then normalized and fed into a multilayer feed-forward neural network to classify the object as human or vehicle. We implemented the classification algorithm on a set of objects detected from real life video surveillance of human and videos. We trained the network with some of the objects and perform recognition with the rest. The algorithm performs excellently at discriminating between these two classes of objects with the recognition rate of 98.5%.

Keywords: algorithm kernel density, neural network, segmentation, normalized,

#### 1.0 Introduction

Object classification in videos is an important requirement in surveillance systems as it aids understanding of the intentions or actions that the object can perform. For instance humans beings can sit, walk, run, fall while cars can move, run, overspeed, crash etc. Object classification is a challenging task because of various object poses, illumination, occlusion, etc. The desire to carry out this research is to use the result of the object classification in higher layer of surveillance system for action recognition. In such systems a higher degree of recognition accuracy and higher response time is highly required. We proposed to use a parallel- based kernel density estimation (Akintola etal., 2011) for an adaptive background subtraction which has been applied to both outdoor and indoor environments and was found very fast and robust. Unlike the previous systems in the literature that

shape information. It has been recognized by Zhao etal.(2000) that the shape information from silhouettes extracted from the segmented region is invariant to color and texture changes. Moreover, computing the distance signals is very fast and demands small storage space which is good qualities for a surveillancesystem which requires fast operation and storage minimization. After object segmentation, connected component analysis is carried out using morphological opening and closing operations. The resulting blobs extracted are then used to extract the feature vectors. The feature vectors are gotten by calculating the distance from the centroid of the object to the contour outline starting at right hand side and moving in anticlockwise direction. The feature vector is thennormalizing by dividing by the sum of the lengths. These vectors are then used to

uses motion information, our systems is based on



train a neural network using supervised learning approach.

### 2.0 Related works.

Many researches have been carried out in literature on object classification using neural networks (Modi etal., 2011, zhao etal., 2000, Teschioni etal., 2008, Khashman etal., 2008, Karital etal., 2012, Collins etal, 2000). In many of the researches, the background subtraction adopted is not robust to quazzibackgrounds, sudden illumination stationary changes, and efficient coding. The problems addressed in this research are on coding of the silhouettes of the actors that can be processed in real timeand that will require moderate storage space, using silhouette shape features that is robust to sudden illumination changes, and issue of better algorithm for background subtraction that can work in a quazi-stationay background. A robust and fast object tracking algorithm was used. Thus, this research focuses on fast background subtraction using parallel adaptive kernel density technique, fast object tracking technique and fast object classification algorithm.

There are two categories of methods used for object classification (shape-based approach and the motion –

based approach).Shape - based classification is the use of an object shape properties such as the bounding rectangle, area, silhouette of detected object regions. Collins etal., (2000) classifies moving object blobs into general classes such as "humans" and "vehicles "using viewpoint-specific neural networks, trained for each camera. Each neural network is a standard three-layer network. Learning in the network is accomplished using the back propagation algorithm. Input features to the network are a mixture of image-based and scene based object parameters: image blob dispersedness (perimeter2/area (pixels)); image blob area (pixels); apparent aspect ratio of the blob bounding box; and camera zoom. There are three output classes: human; vehicle; and human group. This approach discriminate object with fails to similar dispersedness.Modi et al., (2011) presents Neural Network approach for recognition of human motion using stationary camera. It is noted that task to classify and identify objects in the video is difficult for human operator. Object is detected using background subtraction technique. The detected moving object is divided into 8x8 non-overlapping blocks. The mean of each of the blocks is calculated. All mean value is then accumulated to form a feature

vector. A neural network is trained using the generated feature vectors. Experiment performed shows a good recognition ratebutthe object detection algorithm used cannot work under a quasistationary background. Zhao etal., (2000) presents stereo-and neural network approach for pedestrian detection in videos. The motivation for the research is to develop a surveillance system that can avoid dangerous situations. This is achieved using a stereo-based segmentation and neural network based recognition. This system performs well in pedestrian detection especially stationary pedestrians but fails to incorporate motion cues into the system which should have enhanced the performance.

Motion method uses the object's motion characteristics to distinguish the object. Zhou etal., (2006) used the variance of compactness of the object to classify target object as single person, people group, or vehicle. It is noted in the paper that vehicles are more consistent in their motion because they are rigid objects whereas humans shift some parts forward to maintain balance. So the variance of motion direction is employed to measure motion consistency. These features are got from the optical flow of the motion. The computation of optical flow can be computationally expensive.



Figure 1System description

### 3.1 Object Segmentation

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. This model has been used by (Elgammal et al., 2002)for foreground detection. In this algorithm it is assumedthat a local kernel function is centered upon each reference point, and its scale parameter (the bandwidth). The common choices for kernels include the Gaussian and the Epanechnikov kernel.



(a) ba

(c) No

(e)]

Vehi

### **International Conference**

Let  $\mathbf{x}_{1}, \mathbf{x}_{2}, ..., \mathbf{x}_{n} \in \mathbb{R}^{d}$  be a random sample taken from a continuous, univariate density *f*. The Kernel Density Estimator (KDE) is given by:

 $f(x,h) = \frac{1}{nh} \sum_{i=1}^{n} k\left(\frac{x-x_i}{h}\right)$ 

(1)

K(.) is the function satisfying:

$$\int k(\mathbf{x})d = 1$$

(2)

Thus, K(.) is referred to as the Kernel, h is a positive number, usually called the bandwidth or window width. The Gaussian Kernel is given by:

 $K_{N} = (2\pi)^{-\frac{d}{2}} \exp\left(-\frac{1}{2} ||\mathbf{x}||^{2}\right)$ 

(3)

The Epanechnikov kernel is given by:

$$K_{\rm ff} = \begin{cases} -\frac{1}{2} t_{\rm ff}^{-1} (d+2) (1-||\mathbf{x}||^2 t_{\rm f} ||\mathbf{x}|| < 1 \\ 0 & 0 \ her \end{cases}$$
(4)

Kernel Density Estimation (KDE) for background modeling involves using a number of frames (training frames) to build the probability density of each pixel location. A histogram of each pixel location of the image is then constructed. After the histogram is built, we find the adaptive threshold of each pixel.

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 color. No classification is done for these learning frames. Classification is done for the 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 pixels are ordered based on their corresponding value of the histogram bin based on the adaptive threshold in the previous stage. The pixel intensity values for the subsequent

frames are estimated. The corresponding histogram bin is evaluated and the bin value corresponding to this intensity is determined.

If the value is < threshold Classify as foreground

Else

Classify as background.

The classification process is performed on all the pixel locations in the current frame as shown in figure 2.



Figure 2: Human and Vehicle Objects Detection Results

### 3.2 A spatio-color Histogram Algorithm for Scalable 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 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. 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.

### 3.3 Feature Extraction.

After extracting the contour, the centroid of the contour is calculated using equation (4).From that centre a pre-defined number of axes are projected outwards at specified regular angles to the nearest edges of the contour in an anti- clockwise direction as shown in figure 3.

$$\left(\epsilon_x,\epsilon_y\right)=-\frac{1}{n}\left(\begin{array}{cc}n\\t-1\end{array}, \begin{array}{cc}n\\t-1\end{array}\right)\ldots$$

### (5)

The distance from the contour's centre to its nearest edge along that a predefined angle is then stored. This is done for all the other angles to for a set of values called a vector. The dimension of the vector equals to the number of axes being projected from the centre. The vector is then normalized. This ensures that the vector is scale invariant as the largest value in the vector will at this point be 1.0, which will be the longest of the axes projected.

Let S be the segmented object region within the frame. Let line be a line projected at angle I from the centroid to the object boundary at angle i to the horizontal line passing through the centoroid of the object. The length of each line to the contour boundary of the object is given by

linei=  $\prod_{k=c}^{W} \delta(p(k, l))$ 

(6)

Where k and I are the co-ordinates in the x and y

directions respectively. L is given by ktan(6)  $\delta(.)$  is a binary function that returns 0 or 1  $\delta(p(k, l)) = \begin{cases} 1 & i \cdot p(k, l) \in S \\ 0 & o hen \end{cases}$ 

> (7) is the ni

p(k,l) =is the pixel value of the object The number of lines in each image containing an object as well as the number of neurons in the input layer is j where j = {1, 2, 3, n} and n is given by N=360/ $\theta$  where  $\theta$  is the smallest of the angles.Angle size of 10 degrees interval have been used in this research e.g. (10,20,30,360) will divide the object into 36 regions. Out of these 36 lines we only used 32 of them as feature vectors.



### 3.4 Design of Neural Network for Object Classification

We developed a neural network model to train the network. Figure 4 shows the architecture of the neural network developed. The data collected from the contour of the moving objects were used for the modeling.

The BPN learns during a training epoch, it will probably go through several epochs before the network has sufficiently learnt to handle all the data you've provided it and the end result is satisfactory.



A training epoch is described below: For each input entry in the training data set:

• feed input entry data into the network (feed forward).

Initialized weights

• check output against desired value and feedback error (back-propagate)

- Where back-propagation consists of:
- calculate error gradients

• update weights (in our case the weights between output and hidden layers hidden and hidden layers and input and hidden layers were updated in that order)

The network parameters for our system model consist of an input layer with 32 neurons, one hidden layer with 4 neurons, and an output layer with 1 neuron (MLP 32 : 4: 1). The number of epochs used is 1000, with the momentum of 0.5 and learning rate of 0.3. The initial weights were randomly initialized to small random numbers less than 1 using random number generators.

For the stopping conditions there are various commonly used stopping conditions for neural networks, such include; desired accuracy, desired mean square error and elapsed epochs. In this project we used the 1000 elapsed epochs to train the Network. Back-propagation algorithm is a supervised learning algorithm where an error function is defined

(based on the training set) and the error function is minimized by adjusting the weights using hill climbing algorithm. The Mean Square Error (MSE) is the performance index. The error is calculated as the difference between the target (t) expected and the actual value (a) of the network output. Mean square error of one output neuron over all n examples is given by:

$$m = \frac{1}{n} \frac{n}{k-1} (l(k) - u(k))^2$$
(8)

Multilayer perceptions use the back-propagation algorithm to adjust the weights and biases of the network in order to minimize the mean square error overall output and all examples. It is a generalization of the least mean square algorithm whereby an error function is defined to be minimized by using the gradient descent algorithm. This is achieved by adjusting the weights. The generalized delta rule does this by calculating the error for the current input example and then back-propagating this error from layer to layer. The training algorithm as presented by (Koprinska,2013) is as follow:



Figure 4 The Neural Network Model for the Pedestrian/Car Detection.



Consider a neural network with one hidden layer (Figure 4) indexes i over output neurons, j over hidden neurons, p over input patterns. The MSE over all neurons over all input patterns is given by:

$$E = \frac{1}{2} p \left( d_i^p - a_i^p \right)^2$$

(9)

where  $d_t^p$  is the target output of neuron i for input pattern p,  $u_t^p$  is the actual output of neuron i for input pattern p

If E is expressed in terms of weights and input signals, the input for the hidden neuron j and for a given pattern p is:

$$ne_{j}^{p} = k w_{k} e_{k}^{p} + l$$

(10) Activation of neuron *j* as function of its input:

 $e_1^p = f(n_1^p) =$ 

$$f\left(\begin{array}{c} k w_k \cdot u_k^p + b_j \right)$$
(11)

Input for the output neuron i is given by:  $n \quad {}_{t}^{p} = \int_{j} w_{j} \cdot v_{j}^{p} + b_{j} = \int_{j} w_{j} \cdot f\left( \begin{smallmatrix} & w_{k} & .v_{k}^{p} + l_{j} \end{smallmatrix} \right) + b_{t}$ (12) Output for the output neuron i is given by:  $v_{t}^{p} = f\left( \begin{smallmatrix} & p \\ & t \end{smallmatrix} \right) = f\left( \begin{smallmatrix} & J \\ & J \end{smallmatrix} \right) = f\left( \begin{smallmatrix} & J \\ & J \end{smallmatrix} \right) = b_{t} + b_{t} = b_{t}$ (13) Substituting 13 into 9 gives:

Substituting 13 into 9 gives:  $E = \frac{1}{2} p \left[ d_t^p - f\left( {}_j w_j \cdot f\left( {}_k w_k \cdot c_k^p + t_j \right) + b_t \right) \right]^2$ (14)

Steepest gradient descent adjusts the weights so that the change moves the system down the error surface in the direction of the locally steepest descent, given by the negative of the gradient, that is;

$$w_{j} = -\eta \cdot \frac{\delta}{\delta w_{j}} =$$

$$\eta \cdot p(d_{t}^{p} - u_{t}^{p}) \cdot f(n_{t}^{p}) \cdot u_{j}^{p}$$
(15)
$$\eta \cdot p(d_{t}^{p} - u_{t}^{p}) \cdot f(n_{t}^{p}) \cdot u_{j}^{p} = \eta \cdot p\delta_{t}^{p} \cdot u_{j}^{p}$$
(16)

Thus the change in weight for output neuron is given by:

$$d_{\mathfrak{l}}^{p} = (d_{\mathfrak{l}}^{p} - u_{\mathfrak{l}}^{p}) \cdot f\left(n \quad {}_{\mathfrak{l}}^{p}\right)$$

(17) For hidden neuron, the derivatives is calculated using the chain rule as follow:

$$w_k = -\eta \cdot \frac{\partial}{\partial w_j} = -\eta \cdot \frac{\partial}{\partial w_j^D} \cdot \frac{\partial u_j^D}{\partial w_k}$$

$$= \eta \sum_{p} (\delta_{t}^{p} - o_{t}^{p}) f(n_{t}^{p}) w_{j} f(n_{j}^{p}) o_{k}^{p}$$
$$= \eta \sum_{p} \delta_{t}^{p} w_{j} f(n_{j}^{p}) o_{k}^{p}$$

 $= \eta \cdot p \delta_j^p \cdot a_k^p$ 

(18)

where the change in weight for hidden neurons  $\delta_{1}^{p}$  is given by :

$$\delta_j^p = f\left(n \begin{array}{c} p \\ j \end{array}\right), \quad t \in W_j, \quad \delta_t^p$$
(19)

In general for a connection from x to y:

$$w_x = \eta. \quad i_1 \qquad \delta_x. \, c_y$$

$$w_x n = w_x o + w_x$$
(20)

(21) where o is activation of an input or hidden neuron and  $\delta$  is given either by Equation 17 for the output neuron or Equation 19 for the hidden neuron. From the formulas for  $\delta$ , the derivatives of f can be calculated. For a sigmoid transfer function the derivative is calculated as follow:

$$f_{\cdot} \begin{pmatrix} n & p \\ l \end{pmatrix} = d_{R}^{p} = \frac{1}{1+e^{-n} \frac{p}{l}}$$

$$(22)$$

$$\frac{\partial d_{l}^{p}}{\partial n \frac{p}{l}} = \frac{\partial \frac{1}{1+e^{-n} \frac{p}{l}}}{\partial n \frac{p}{l}}$$

$$(23)$$

$$\frac{e^{-n} \frac{p}{l}}{(1+e^{-n} \frac{p}{l})^{2}}$$

$$(24)$$

 $u_l^p \cdot (1 - u_l^p)$ (25)

The Back propagation rule for sigmoid transfer function is as follows:

=

The change in weight calculation for the output neuron is given by:

$$\delta_{t}^{p} = (d_{t}^{p} - u_{t}^{p}).u_{t}^{p}.(1 - u_{t}^{p})$$

$$w_{j} = \eta. \quad p(d_{t}^{p} - a_{t}^{p}).a_{t}^{p}.(1 - u_{t}^{p}).u_{j}^{p}$$
(26)
(27)

The change in weight calculation for the hidden neuron is given by:

$$\delta_{j}^{p} = u_{j}^{p} \cdot (1 - u_{j}^{p}) \cdot w_{j} \cdot \delta_{t}^{p}$$

$$w_{j} = \eta \cdot k(d_{j}^{p} \cdot u_{k}^{p}) = p u_{j}^{p} (1 - u_{j}^{p}) \cdot t w_{j} \delta_{t}^{p} \cdot u_{k}^{p}$$

$$(29)$$



### 4.0 Experiment and Result

ANN models' performances can be measured by the mean relative percentage error. It measures the accuracy of prediction through representing the degree of scatter. Eq.1 was utilized to calculate the relative error for each case in the testing set. Then, the calculated values were averaged and factored by express in percentages. (actual)-100 to (predicted)/actual \*100%.In this research the network was trained using 80 data items consisting of forty vehicles and 40 pedestrians. The class vehicle was assigned 1 while the class human was assigned o. The network is then trained using backpropagation algorithm. The weights from the training are then saved. For testing, 343 data items were used. Out of this 120 were vehicle data while 223 were human data. We decided from the ROC curve (figure 5) to make the threshold for vehicles to be greater or equal to 0.9 while that of the human is set to less or equal to 0.1. Out of the 120 vehicles, only one was misclassified while for 223 human class data items, only four were misclassified. Error in forecast is 1.46% so the accuracy is 98.54%



Figure 5: ROC curve of the Neural Network.

### 5.0 Implementation

The model has been applied in real life scenario to recognize vehicles and human. Figure 6 shows the performance of the algorithm on the international surveillance data called PETS 2002 dataset provided by third IEE international workshop on performance evaluation of tracking and surveillance. Figure 7 shows the performance of the algorithm in an indoor environment. Figure 8 shows the performance of the algorithm in an outdoor environment at FUTA community. It is observed that the classifier is able to classify the objects irrespective of their sizes and environmental conditions. Thus the algorithm is adaptive to changes.







Figure 6Classifications of Vehicles in an

Outdoor Environment









Figure 7 Classification of Humans in an Indoor Environment



Figure 8: Classification of Vehicles in Outdoor

Environment

### 6.0 Conclusion

The essence of this paper is to present a neural network model and lightweight object shape descriptors for fast object classification in videos. Distance signal features are extracted from the silhouettes of the detected objects. The distance signals features are then normalized and fed into a multilayer feed-forward neural network to classify the object as human or vehicle. We implemented the classification algorithm on a set of objects detected from real life video surveillance of human. We trained the network with some of the objects and perform recognition with the rest. The algorithm performs excellently at discriminating between these two classes of objects with the recognition rate of 98.5%. It can be seen that neural network on the shape descriptor performs excellently at object classification. Future work can be done on this study by considering other classifiers such as radial basis network, recurrent network and support vector machines.

### 7.0 References

Akintola K.G , Tavakollie A., (2011), Robust Foreground detection in Videos using Adaptive Colour Histogram Thresholding and shadow removal 7<sup>th</sup>International Symposium on Visual Computing Las-Vegas, NV, September, 2011.

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.No. 0802mds2008020001.

Elgammal A., Duraiswami R, Harwood D., and Davis L.S (2000)Backgroundand Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance, PROCEEDINGS OF THE IEEE, VOL. 90, NO. 7, JULY 2002.

Kavita P. Patil S.V. (2012) Tracking and counting human in visual surveillance systems, international journal of electronics and communication engineering and technology (IJECET) 3(3) 139-146 www.iame.com/ijecet.asp



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) 25(6) July 2011 Teschioni A., Oberti F., and Regazzoni C. (2008), A neural network approach for moving objects recognition in color image sequences for surveillance applications

Khashman A. (2008) Automatic detection, extraction and recognition of moving objects international journal of systems applications, engineering and development issue 1, vol, 2 2008

Koprinska I. (2013). Lecture Notes on Artificial Neural Networks, School of Information, the University of Sidney. Accessed July 2013.

Zhao L., Thorpe C.E., (2000) Stereo and neural network based pedestrian detection IEE transactions on intelligent transportation systems vol. 1No. 3 September 2000



### EXPERIMENTAL FORMULATION OF A VECTOR SPACE FOR GRAMMATICALITY GRADING IN HANDCRAFTED GRAMMAR

\*M. K. Aregbesola<sup>1</sup>, R. A. Ganiyu<sup>2</sup>, S. O. Olabiyisi<sup>3</sup>, E. O. Omidiora<sup>4</sup>, and O. O. Alo<sup>5</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, Elizade University Ilara Mokin, Ondo State, Nigeria

<sup>2, 3, 4, 5</sup>Department of Computer Science and Engineering, Ladoke Akintola University of Technology Ogbomoso, Oyo State, Nigeria

<sup>1</sup>kehinde.aregbesola@elizadeuniversity.edu.ng, <sup>2</sup>raganiyu@lautech.edu.ng, <sup>3</sup>soolabiyisi@lautech.edu.ng, <sup>4</sup>eoomidiora@lautech.edu.ng, <sup>5</sup>ooalo@lautech.edu.ng

### ABSTRACT

Grammaticality is a concept within the fields of Natural Language Processing and Computational Linguistics that deals with the measure of grammatical correctness or incorrectness of a natural language expression. Among the methods employed for grammaticality evaluation, handcrafted grammar is the most precise. The major drawback of handcrafted grammar however is that its evaluation process scarcely computes gradience of grammaticality. It therefore evaluates the grammaticality of an expression in a binary (right or wrong) fashion. This approach however, for instance, does not provide second language (L2) learners a scale for measuring progress made in the learning process. To solve this problem, a number of mechanisms were considered from which Linear Vector Space (LVS) was observed to be appropriate, having been successfully applied over a wide range of disciplines, including computational linguistics. LVS was therefore considered a suitable mechanism for addressing the gradience problem of handcrafted grammar. In this work therefore, a linear vector space of grammaticality was formulated using key identifiable grammaticality error features (missing word, extra word, real word spelling, verb form, puctuation and agreement errors) as elements of the vector space. Equations for grammaticality grading were further derived based on the formulated vector space. The results showed that the grammaticality gradience  $\mu$ G is such that 0  $\leq \mu$ G  $\leq$  1. Therefore, by using the formulated vector space with handcrafted grammar, grammaticality can be evaluated on a graded scale between 0 and 1. Where 0 is for a totally ungrammatical and unrecognisable expression while 1 is for a perfectly grammatical sentence.

**Keywords:** Experimental Formulation, Grammaticality Grading, Handcrafted Grammar, Vector Space Of Grammaticality





### 1. INTRODUCTION

This work is part of an ongoing research aimed at developing an automated system for graded grammaticality evaluation, error detection and error correction in English language expressions; using British English as a standard.

Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) that deals with the ability of machines to act intelligently within a language domain (MacKinlay, 2012; Chu, 2003). Grammar is a linguistic concept often considered within the field of NLP in Computer Science. It has attracted a lot of research interest from a wide range of disciplines. It is the branch of linguistics that deals with syntax, and sometimes, semantics. morphology, It determines the set of acceptable constructs of a language (Baldridge et al., 2007). As a derivative of grammar, grammaticality describes whether a grammatical (well-formed) sentence is or ungrammatical (ill-formed). Grammaticality has been the focal point of interest in a number of NLP research (Wagner, 2012). Grammaticality judgment has been applied to significantly improve the performance of a number of NLP systems involving language generation, machine translation, text summarization, second language learning, and automated essay scoring (Nerbornne, 2002; Pauls and Klein, 2012). Several approaches have been adopted in evaluating the grammaticality of sentences. There are two prevailing schools of thoughts about the concept of grammaticality. The first argues that the natural way languages are learnt is independent of grammatical rules. This school of thought cites that a child learns to communicate in a language without necessarily having to first learn the grammatical rules of the language. This school of thought often uses induced grammar approaches for evaluating grammaticality. The second school of thought lays emphasis on following a strict set of grammar rules (handcrafted grammar) in determining the grammaticality of a sentence. These grammar rules are designed by linguistic engineers and employed as natural language parsers for parsing (evaluating) expressions of the target language. Irrespective of the school of thought being considered, the degree of grammatical correctness or incorrectness is equally of significant interest in NLP (Parfitt *et al.,* 1999).

Handcrafted grammar which is the focus of this study, make it possible to describe the question of grammaticality with precision (Blache et al., 2005). It however scarcely expresses the gradience of grammaticality. Thus making it difficult to quantify the degree of correctness or incorrectnes of an expression. This quantification feature is however very desirable, especially for second language (L2) learners, and particularly significant for providing L2 learners with performance feedback. This study is therefore aimed at formulating a gradience grammaticality mechanism facilitate to quantification in handcrafted grammar. Previous works on grammaticality gradience have been exclusively within the domain of almost probabilisticaly induced grammars (Crocker and Keller, 2006).

#### 2. LITERATURE REVIEW

### 2.1 Grammar and Grammaticality

Grammar is a term used to define the body of rules that describe the structure of expressions of a language. This is inclusive of the structure of words, phrases, clauses, and sentences. A grammar is said to under-generate if it excludes grammatical sentence, and over-generate if it allows ungrammatical sentences (MacKinlay, 2012). A grammar should avoid both of these two situations. There is thus this tension between avoiding overgeneration and maximizing coverage. Grammar developers therefore seek to make grammars as constrained as possible, while they avoid rejecting grammatical sentences. This therefore calls for grammars to exhibit a high degree of precision. The parsability of a sentence however does not necessarily imply its acceptability (Lin, 2007). Therefore, the real world context of word usage must be put into consideration.

Fillmore (1998) and Kay and Fillmore (1999) looked at Construction Grammar in which all objects are called constructions. On the other hand, Blache (2006) looked into property grammar which can be very efficient and robust. Constraints relaxation allows the system to select the best assignment,



and not just the assignment satisfying the constraint system, thus permitting to deal with more or less grammatical sentence. Constraints based approach makes it possible to process any kind of inputs.

Sampson (2007) asked the question of whether it was possible to talk about grammar without talking about grammaticality. The role of grammaticality judgment in the learning (acquisition) of a second language cannot be over emphasised (Gass, 1983). The roles of metalinguistic awareness in the learning of a second language (L2) make grammaticality judgment for L2 learners an observation that can be considered either as an implicit or an explicit concept (Bialystok, 1979; 1981). The study of Clark et al. (2013) examined the grammaticality between relationship and acceptability. They considered grammaticality a notion while theoretical acceptability was considered an empirically testable property. Acceptability was also further explained to be partly dependent on grammaticality.

Automating grammaticality evaluation has taken the front line in a number of linguistic and NLP research. A system equipped with the capability to perceive determine how users will the grammaticality of a sentence could be useful in fields involving language generation, machine translation, and text summarization. It could also be used for automatic essay scoring and for the learning of a second language. Lau et al. (2014) described an experiment explaining that even the most popular translation tools are in many occasions not congruent in performing a backward translation (from L1 to L2 and then to L1 again) of given sentences. It therefore becomes necessary to ensure the grammatical correctness of written expressions to minimize distortion when translated to other languages.

### 2.2 Handcrafted Grammar

Handcrafted grammars are grammars that are manually specified and evaluated by grammar engineers (Witbrock and Mittal, 1999). Handcrafted grammars are completely different from induced grammars. In contrast to requiring large scale treebanks, handcrafted grammars provide a single point of control for the handling of syntactic phenomena (MacKinlay, 2012). Owing to this single point of control, the grammar can be designed to handle even very rare and complex phenomenon, or to even ignore particularly difficult ones, which is not possible with induced grammars. It is a very labour-intensive task to create a grammar of a language which has reasonable coverage and can create meaningful parse trees and semantics. There has nonetheless been a range of concerted efforts targeted at creating broad-coverage grammars for several languages, some of which are closely related in formalism.

### 2.2.1 Transformational Grammars

Chomsky (1957) proposed Transformational Grammar TG in response to his argument that Context-free Phrase Structure Grammar (CFPSG) linguistically cannot capture significant generalization of natural language syntactic propoerties. The TGs augment the CFPSGs with rules that map their syntactic structures unto other syntactic forms. Thus for instance, to die could be transformed to ceases to live. However, following founded criticisms from Peters and Ritchie (1973) as to the overly unrestricted nature of the TG, Chomsky developed the theory of Government and binding, Principles and Parameters P&P, and Minimalism, which are more constrained versions of TGs (Chomsky, 1981). TGs were literarily abandoned for over a decade after this period (Malchow, 1992).

#### 2.2.2 Feature Based Grammars

Following the crisis with TGs came the Feature based grammars, also known as information based grammars, attribute-value grammars, constraint based grammars or unification grammars (Johnson, 1988; Kay, 1979). One of the crucial techniques, unification, employed by all feature based approaches is the emphasis of the unification grammar (Malchow, 1992). Among the feature based grammars are the following formalisms: Functional Unification Grammar (Kay, 1979; 1982); Generalized Phrase Structure Grammar (Gazdar, 1981; Gazdar et al., 1985); Lexical Functional Grammar (Bresnan, 1982); Categorial Unification Grammar (Karttunen, 1986; Uszkoreit, 1986); and Head-driven Phrase Structure Grammar (Pollard, 1984; Pollard and Sag, 1987).

Representing the sets of features of a grammar as one complex symbol using feature based grammars



has attracted a lot of research interests. Malchow (1992) explained that there are three basic ways of structuring complex categories. These structures are: Tree diagrams, also known as directed acyclic graphs (dags); Terms; and Matrices.

The following sentences demonstrate an example of constraint on English sentences as discussed by Müller (2013):

- a. Kim loves Sandy
- b. \* I loves Sandy

Where '\*' connotes an error in sentence (b) owing to the violation of the constraint that the subject has to agree with the verb both in person and number.

### 2.2.3 Parse Trees

The syntactic structure of natural language sentences are usually expressed using context-free grammars which are often represented as ordered, rooted trees known as parse trees. As an example, let us consider the parse tree for the English sentence "The fat pony sleeps in the barn" shown in Figure 1. Terms abbreviated in this example include: sentence as 'S', noun phrase as 'NP', verb phrase as 'VP', verb as 'V', determiner as 'D', noun as 'N', adjective as 'A', preposition as 'P' and the nominal category as NOM. Each of the nodes in this tree is either a root node, a branch node, or a leaf node (Carnie, 2013). There can only be one root node, in a tree. A root node (S in this case) is one with no parents (no branches on top of it). A branch node on the other hand is a non-root node (VP and NP in this case) that connects to two or more child nodes. Finally, a leaf node is a terminal node with no child nodes attached to it. The leaves are the actual lexical tokens of the sentence.



Figure 1: A phrase structure tree for "The fat pony sleeps in the barn" (MacKinlay, 2012)

### 2.3 Grammaticality Gradience

Grammaticality gradience or quantification is the degree to which a sentence belongs to the grammar of a language (Lau et al., 2014; Chomsky, 1975). Grammaticality is inherently graded by nature and not absolute (Parfitt et al., 1999; Taylor, 1995). The acceptability of sentences can be predicted by quantifying its grammaticality. Crocker and Keller (2006) explained that probabilistically induced grammars are inherently models of gradience in language processing due to their mode of operation. Handcrafted grammar on the other expresses hand scarcely gradience of grammaticality.

Blache *et al.* (2005) however characterized the result of a parse operation into the set of satisfied properties noted as P+, and set of violated ones, noted P-. Thus the sentence is acceptable if its grammaticality falls within a particular range, and unacceptable if its grammaticality falls outside of that range.

Linear Optimality Theory is very effective when applied in weighting the different syntactic constraints associated with a given grammar. Such evaluated constraint weights is then useful for quantifying sentence grammaticality (Keller, 2000; Sorace and Keller, 2005). In a similar sense, Schroder (2002) and Menzel and Schroder (1998) looked at the use of Constraint Dependency Grammar framework for weighting syntactic constraints. Blache (2005) looked at fully constraint-based syntactic formalism called Property Grammars that take into account automatically derived parameters and empirically determined weights in calculating a grammaticality index.

#### 2.4 Vectors and Vector Spaces





Vectors and points are common data structure considered in many areas of Mathematics and Computer Science. They are applied extensively in data compression, image processing, computer vision, computer graphics, and numerical analysis. Two-dimensional vectors can be defined as directed arrows in the plane. The position of the arrow is not important. The length (magnitude) and direction of the arrow are the important features of the vector, and they determine the vector. They can be added, scaled and rotated (Lindeman, 2008). Vectors having the same length and direction are said to be equivalent (Kambites, 2014). Two vector in the same direction are said to be parallel. The zero vector has a magnitude of zero and is denoted as o. Figure 2a shows a vector a between two points  $(x_1, x_2)$ v1) and (x2, v2). In Figure 2b, the vectors AB and DC are equivalent, because two-dimensional vectors are distinguished only by length and direction. They are thus treated as equal i.e.  $\underline{AB} = \underline{DC}$ .

### 2.4.1 Linear Vector Space

The concept of linear vector space also commonly referred to simply as vector space is the focus of a number of literatures (Zybura, 2004; Nowak, 2010; Luenberger, 1969; van Hassel, 2009; and Kondratieva, 2008). A vector space V can summarily be defined as a set of vectors over the field F (such as real, complex, and natural numbers) which may be added together and multiplied (or scaled) by numbers referred to as scalars, such that, given that x, y,  $z \in V$ , the following eight axioms are satisfied:

i. Associativity of addition x + (y + z) = (x + y) + z;

(a) 
$$y = (x_2, y_2)$$
  
 $(x_1, y_1)$   
 $(x_1, y_1)$ 



Figure 2 (a): Vector in the plane (b): Two equivalent and parallel vectors (Lindeman, 2008; Kambites, 2014)

ii. Commutativity of addition

$$x + y = y + x;$$

- iii. Identity element of addition. There exists an element o ∈ V, called the zero vector, such that y + o = y for all y ∈ V;
- iv. Inverse element of addition. For every  $y \in V$ , there exists an element  $-y \in V$ , called the additive inverse of y, such that

- v. Compatibility of scalar multiplication with field multiplication. a(by) = (ab)y;
- vi. Identity element of scalar multiplication
  - 1y = y, where 1 denotes the multiplicative identity in F;
- vii. Distributivity of scalar multiplication with respect to vector addition. a(x + y) = ax + ay;
- viii. Distributivity of scalar multiplication with respect to field addition. (a + b)y = ay + by.

(Zybura, 2004; Nowak, 2010; Luenberger, 1969; van Hassel, 2009; and Kondratieva, 2008).

Depending on the literature, these expressions are sometimes compounded to give fewer axioms or expanded to give more axioms, though expressing the exact same concepts. Also, contrary to the common use of the word "linear" in linear vector space to connote vector space in a straight line, this study considers the word from the perspective of dimensionality. "Linear" is applied in this study to refer to "one-dimensional", as against "multidimensional".



### 2.4.2 NLP Applications of Vectors and Vector Spaces

The concept of vector space is considered in most linguistic and NLP literatures from the perspective of lexical and semantic distribution. Semantic vector space models of language make use of realvalued vectors to denote each word that are typically associated with a particular word. Words that typically occur together are assigned values that often depict their probability of occurring together in a sentence. They are used for a wide range of NLP operations including grammaticality evaluation and error detection. Detailed semantic and syntactic regularity have been successfully captured using vector space representations and vector arithmetic. The study of Pennington et al. (2014) came up with global vectors for word representation (GloVe). The studies focused on highlighting the properties that made the emergence of such captured regularities possible in word vectors. Schmid (2004) however focused on efficient parsing of highly ambiguous context-free grammars using bit vectors.

The study of Stolcke (1991) represented a formalism dubbed Vector Space Grammars (VSG) for deriving phrase structure categories that made use of structured samples of a context-free language. Using the connectionist approach, the entire training process made use of adaptation, competition and error back-propagation, all occurring in a continuous vector space. It advocates the use of vectors instead of symbols for the purpose of linguistic category labeling.

Vector Space Semantic Parsing (VSSP) presented in the work of Krishnamurthy and Mitchell (2013) is a framework for learning compositional models of vector space semantics. It applies Combinatory Categorial Grammar (CCG) to define the relationship between syntactic categories and semantic representations, taken as vectors and functions on vectors. Using CCG based semantic parser, texts are parsed into lambda calculus formulae that compute to equivalent vector space representations.

In general, vector space models make use of vectors and operations on vectors to represent the semantics of natural language expressions (Turney and Pantel, 2010). A number of other studies

including Coecke *et al.* (2013), Socher *et al.* (2011), Socher *et al.* (2012), Turney (2006), and Rapp (2003) focused on similar concepts. The studies achieved significant performances that corresponded well with human judgment.

Grammaticality is considered a vector concept within this literature, having both magnitude and direction. This is in contrast with scalar quantities that have only magnitude but no direction. The direction of grammaticality is either towards grammatical correctness, or away from grammatical correctness. Grammatical correctness is a state described as Grammatical Equilibrium (GE) within this study, and is ascribed a gradience value of zero (o).

On a general note, grammars are usually designed to express the state of grammatical equilibrium. For constraint based grammars (Johnson, 1988; Kay, 1979), each appropriate constraint within the grammar enforces the grammar towards being able express or determine sentences that are grammatically correct.

Furthermore, although grammaticality (Lau et al., 2014) is generally used to express the state of grammatical correctness or otherwise of a sentence, it is sometimes used strictly as a measure of grammatical correctness, especially when used alongside 'ungrammaticality'. From this perspective, grammaticality is used as a measure of grammatical correctness while ungrammaticality is used as a measure of grammatical incorrectness.

### 3. FORMULATION OF A LINEAR VECTOR SPACE FOR GRAMMATICALITY GRADING

The formulation of the vector space V of grammaticality vectors for grammaticality gradience is presented in this section. V defines the space of operation of grammaticality gradience for handcrafted grammars.

### 3.1 The Vector Space (V) of Grammaticality

Let x, y, z be weighted entities associated with grammaticality such as possible error categories including missing-word, extra-word, real-word spelling, verb-form, punctuation and agreement errors, which are uniquely identifiable within a sentence. Also let  $\{x, y, z, ...\} \in V$ . Like any other standard vector space, V is a set of vectors over the



field F (which in this case is the set of real numbers R) which may be added together and multiplied (scaled) by numbers referred to as scalars, such that the eight axioms listed in section 2.4.1 are satisfied.

These grammaticality vectors dubbed Mosesean vectors are strictly linear (one-dimensional) over the written plane. Therefore, the addition operation on the elements of V is performed by simple arithmetic summation ( $\Sigma$ ). Furthermore, ungrammatical elements within an input sentence are assigned negative values; while counter measures to correct such ungrammaticality are assigned positive values. Thus, the magnitude of grammaticality the Mosesean vectors introduce into the proposed system at any point in time is totally dependent on the magnitude of existing ungrammaticality. Figures 3, 4 and 5 illustrate these concepts.



Figure 3: Mosesean vectors on a real number line showing –ve Grammaticality Ungrammaticality



Figure 4: Resultants of Mosesean vectors not yet at equilibrium



Figure 5: Resultant of Mosesean vectors at equilibrium

#### 3.2 Resultants of Mosesean Vectors

The default value of zero (0) is assigned as gradience value to any sentence introduced into the proposed system of Mosesean vectors. At this default value, the sentence is at equilibrium, and is completely grammatical. This equilibrium is toppled when ungrammatical elements are identified within Grammatically the sentence. When a sentence is ungrammatical

by a certain magnitude, the proposed system attempts to find complementary grammaticality measures to pull the sentence back into equilibrium as depicted in Figure 5. Grammaticality (+g) is generated in response to Ungrammaticality (-g). +g can only be as large as to cancel out -g, resulting in equilibrium.

### 3.3 Derivation of Grammaticality Equations Using Mosesean Vectors

The default value of zero (o) is assigned as gradience value to any sentence introduced into the proposed system of Mosesean vectors. If the sentence is grammatically correct, the gradience remains unchanged at zero, and requires no further computation. However, if ungrammaticality (-g) is found within the sentence, the cumulative ungrammaticality (- $G_{sum}$ ) is the arithmetic sum of the individual ungrammaticality values.

Thus:

if the values assigned to ungrammaticality items in a sentence are  $-g_1, -g_2, -g_3, \dots, -g_n$ 

then

-Gsum = 
$$(-g_1) + (-g_2) + (-g_3) + \dots + (-g_n)$$
  
n  
127  $-g_i$   
 $i = 1$ 



(1)

-G<sub>sum</sub> =

Multiply both sides by -1 we get

$$G_{sum} = i = 1$$

n

Therefore, the cumulative grammaticality  $(G_{sum})$  required to bring the sentence into equilibrium is:

$$G_{sum} = \begin{bmatrix} |g_i| \\ i = 1 \end{bmatrix}$$
(2)

Furthermore in the proposed system, the grammaticality gradience of an ungrammatical sentence is computed by dividing the cumulative grammaticality ( $G_{sum}$ ) by the total number of leaf nodes in the sentence parse tree. The number of leaf nodes in a sentence parse tree is always equal to the number of words in the sentence, it therefore follows that grammaticality gradience ( $\mu$ G) for a sentence of length m is:

$$\mu G = \frac{\begin{vmatrix} n \\ |g_i| \\ \underline{i=1} \end{vmatrix}}{m}$$
(3)

In addition, the proposed system assigns a value of minus one (-1) to each ungrammaticality item in an input sentence. This follows that the cumulative grammaticality (Gsum) can at most be as large as the number (m) of words in the sentence.

Thus in equation (3): when n o:

m = m

when n ∞:

and µG 1

Hence, the grammaticality gradience  $\mu G$  is such that:

### $0 \le \mu G \le 1$

As grammaticality evaluation extends beyond the evaluation of a single sentence to the evaluation of multiple sentences, the gradience for each sentence is computed in the same manner, applying equation (2) to each of them. The arithmetic mean of the gradience(s) of the different sentences is then computed to give the gradience of all the sentences put together.

Thus: if there are q sentences with gradience values  $\mu G_1,\,\mu G_2,\,\mu G_3,\,...\,\,\mu G_q$ 

then

$$\mu G_{sum} = \mu G_1 + \mu G_2 + \mu G_3 + \dots + \mu G_q$$

$$q$$

$$\mu G_{sum} = \frac{\mu G_i}{i = 1}$$
(4)

and by extension, the mean grammaticality gradience  $\mu$ G for multiple sentences is:

$$\mu G = \frac{\underset{i=1}{\mu}G_{i}}{q}$$
(5)

### 3.4 Result

The formulations show that grammaticality gradience  $\mu$ G is such that  $0 \le \mu$ G  $\le 1$ . Therefore, based on the formulated vector space, grammaticality is evaluated on a graded scale between 0 and 1. While the grammaticality gradience value of 0 is assigned to a totally ungrammatical sentence, a grammaticality gradience value of 1 is assigned to a perfectly grammatical sentence.

### 4. CONCLUSION

This paper proposes the use of the derived equations (Equations 1 through 5) for computing the grammaticality gradience of natural language text. The gradience of grammaticality of a natural language text can be computed by inserting the derived equations into respective sections of a grammaticality evaluation algorithm. Subsequent phases of this research will focus on applying the derived equations in the automated grammaticality gradience computation for handcrafted grammar. Such automated implementation will provide more data for further validation of the proposed approach.

### **Future Work**



More work is still been done with respect to implementing these equations in an automated system for evaluating grammaticality gradience in user sentences. The effectiveness of these formulations will be tested extensively and results published.

#### ACKNOWLEDGEMENT

We hereby acknowledge and appreciate the Nigeria Computer Society (NCS) for the Postgraduate Scholarship awarded to Mr. Aregbesola M.K. in pursuance of this research work.

### REFERENCES

Baldridge, J., Chatterjee, S., Palmer, A., & Wing, B., 2007. DotCCG and VisCCG: Wiki and Programming paradigms for Improved Grammar Engineering with OpenCCG. Proceedings of the Grammar Engineering Across Frameworks (GEAF07) Workshop, pp. 5-25.

Bialystok, E., 1979. Explicit and implicit judgments of L2 grammaticality. *Language Learning*, 29(1): pp. 69-83.

Bialystok, E., 1981. The role of linguistic knowledge in second language use. Studies in Second Language Acquisition, 4(1): pp. 31-45.

Blache, P., 2005. Property Grammars: A Fully Constraint-Based Theory. in H. Christiansen & al. (eds), Constraint Solving and NLP, Lecture Notes in Computer Science, Springer.

Blache, P., 2006. A Robust and Efficient Parser for Non-Canonical Inputs. In proceedings of Robust Methods in Analysis of Natural Language Data, EACL workshop.

Blache, P., Hemforth, B., & Rauzy, S., 2005. Acceptability Prediction by Means of Grammaticality Quantification. Dec., COLING-ACL, <hal-00135422>, pp.57-64.

Bresnan, J., 1982. Control and complementation. In: J. Bresnan (ed.): The mental representation of grammatical relations. Cambridge, pp. 282-390.

Carnie, A., 2013. Syntax: A Generative Introduction. Third Edition. John Wiley & Sons, Inc.

Chomsky, N., 1957. Syntactic structures. Janua linguarum: Series minor 4, The Hague, Netherlands: Mouton.

Chomsky, N., 1975. The Logical Structure of Linguistic Theory. Plenum Press.

Chomsky, N., 1981. Lectures on Government and Binding. Studies in Generative Grammar 9.

Chu, H., 2003. Information representation and retrieval in the digital age. Medford, N. J.: Information Today, Inc. Published for the American Society for Information Science and Technology.

Clark, A., Giorgolo, G., & Lappin, S., 2013. Statistical Representation of Grammaticality Judgements: the Limits of N-Gram Models. Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics, pp. 28-36.

Coecke, B., Grefenstette, E., & Sadrzadeh, M. 2013. Lambek vs. Lambek: Functorial Vector Space Semantics and String Diagrams for Lambek Calculus. Department of Computer Science, University of Oxford.

Crocker, M. W., & Keller, F., 2006. Probabilistic grammars as models of gradience in language processing. In Gisbert Fanselow, Caroline F'ery, R. V. and Schlesewsky, M., editors, Gradience in Grammar: Generative Perspectives. Oxford University Press Inc., New York, USA., pp. 227–245.

Fillmore, C., 1998. Inversion and Contructional Inheritance. In Lexical and Constructional Aspects of Linguistic Explanation, Stanford University.

GASS, S., 1983. The Development of L2 Intuitions. The University of Michigan. TESOL QUARTERLY, Vol, 17, No.2.

Gazdar, G., 1981. Unbounded Dependency and Coordinate Structure. Linguistic Inquiry, 12, 2 - pp. 155-184.

Gazdar, G., Klein, E., Pullum, G., & Sag, I., 1985. Generalized Phrase Structure Grammar. Cambridge, Mass., Harvard University Press.

Johnson, M., 1988. Attribute-Value Logic and the Theory of Grammar. Stanford CLSI, CSLI Lecture Notes 16.

Kambites, M., 2014. Vectors and Matrices. Lecture notes, School of Mathematics, University of Manchester. Retrieved 05/01/2015 from



http://www.maths.manchester.ac.uk/~mkambites/t eaching/cs/lectures.pdf.

Karttunen, L., 1986. *Radical Lexicalism*. Artificial Intelligence Center at SRI and Center for the Study of Language and Information, Stanford.

Kay, M., 1979. Functional Grammar. In Proceedings of the fifth Annual Meeting of the Berkeley Linguistics Society. Berkeley, California.

KAY, M., 1982. Syntactic Models: Parsing in functional unification grammar. Academia.edu. Retrieved 07/01/2015 from https://www.academia.edu

/2803506/Parsing\_in\_functional\_unification\_gramm ar

Kay, P., & Fillmore, C., 1999. Grammatical Constructions and Linguistic Generalizations: the what's x doing y construction. Language.

Keller, F., 2000. Gradience in Grammar. Experimental and Computational Aspects of Degrees of Grammaticality. Phd Thesis, University of Edinburgh.

Kondratieva, M., 2008. Linear vector spaces and subspaces. Memorial University. Retrieved 06/01/2015 from http://www.math.mun.ca/~mkondra/

linalg2/la2lec1.pdf.

Krishnamurthy, J., & Mitchell, T. M., 2013. Vector Space Semantic Parsing: A Framework for Compositional Vector Space Models. In Proceedings of the Workshop on Continuous Vector Space Models and their Compositionality, Sofia, Bulgaria, pp. 1-10.

Lau, J. H., Clark, A., & Lappin, S., 2014. Measuring Gradience in Speakers' Grammaticality Judgements. In Proceedings of the 36th Annual Conference of the Cognitive Science Society, Quebec City, Canada, pp. 821-826.

Lin, C. J., 2007. Grammaticality and Parsability in Sentence Processing. International Workshop on Grammar and Evidence, Chiayi, pp. 13-15.

Lindeman, R. W., 2008. Computer Graphics: Points, Scalars and Vectors. Lecture Notes, Department of Computer Science, Worcester Polytechnic Institute.

Luenberger, D. G., 1969. Optimization by Vector Space Methods. Stanford University, Stanford, California. Retrieved on 06/01/2015 from http://math.oregonstate.edu/~show/old/142\_Luenb erger.pdf, John Wiley & Sons, Inc.

MacKinlay, A., 2012. Pushing the Boundaries of Deep Parsing. PhD Thesis submitted to the Department of Computing and Information Systems, The University of Melbourne, Melbourne, Australia.

Malchow, A., 1992. Syntax and Semantics of Construction. MA Thesis, Fachgruppe Sprachwissenschaft, Universität Konstanz.

Menzel, W., & Schroder, I., 1998. Decision procedures for dependency parsing using graded constraints. In S. Kahane & A. Polgu`ere (eds), Proc. ColingACL Workshop on Processing of Dependencybased Grammars.

Müller, S., 2013. German Grammar and General Linguistics. Symposium: The Semantics of Action, Göttingen, pp 1-25.

Nerbonne, J., 2002. Computer-assisted language learning and natural language processing. In Mitkov, R., editor, Handbook of Computational Linguistics, pp. 670–698, Oxford University Press.

Nowak, R., 2010. Review of Linear Algebra. Lecture Note: ECE 830 Fall 2010 Statistical Signal Processing. Lecture 3.

Parfitt, S., Tiilo, P., & Dorffner, G., 1999. Graded grammaticality in Prediction Fractal Machines. *In Proceedings of NIPS'*99, pp. 52-58.

Pauls, A., & Klein, D., 2012. Large-scale syntactic language modeling with treelets. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics, Jeju, Korea pp. 959–968.

Pennington, J., Socher, R., & Manning, C. D., 2014. *GloVe: Global Vectors for Word Representation*. Computer Science Department, Stanford University, Stanford, CA 94305.

Peters, P. S., & Ritchie, R. W., 1973. On the Generative Power of Transformational Grammars. Information Science, 6: pp.49-83.

Pollard, C., 1984. Generalized Phrase Structure Grammars, Head Grammars, and Natural Languages. Dissertation, Stanford University.

Pollard, C., & Sag, I. A., 1987. Information-Based Syntax and Semantics. Fundamentals Vol. 1. Lecture Notes 13. Stanford.



Rapp, R., 2003. Word sense discovery based on sense descriptor dissimilarity. In Proceedings of the Ninth Machine Translation Summit.

Sampson, G. R., 2007. Grammar without grammaticality. Corpus Linguistics and Linguistic Theory, 3(1): pp. 1–32.

Schmid, H., 2004. Efficient parsing of highly ambiguous context-free grammars with bit vectors. In Proceedings of the 18th International Conference on Computational Linguistics. Geneva, Switzerland. COLING, pp. 162–168.

Schroder, I., 2002. Natural Language Parsing with Graded Constraints. PhD Thesis, University of Hamburg.

Socher, R., Huval, B., Manning, C. D., & Ng, A. Y., 2012. Semantic Compositionality Through Recursive Matrix-Vector Spaces. In Proceedings of the 2012 Conference on Empirical Methods in Natural Language Processing (EMNLP).

Socher, R., Pennington, J., Huang, E. H., Andrew Y. Ng, & Manning, C. D., 2011. Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP).

Sorace, A., & Keller, F., 2005. Gradience in Linguistic Data. Lingua, pp. 115.

Stolcke, A., 1991. Syntactic Category Formation with Vector Space Grammars. In Proceedings of 13th annual conference of Cognitive Science Society, Chicago.

Taylor, J. R., 1995. Linguistic Categorization: Prototypes in linguistic theory. Claren don, Oxford.

Turney, P. D., 2006. Similarity of semantic relations. Computational Linguistics, 32(3).

Turney, P. D., & Pantel, P., 2010. From frequency to meaning: vector space models of semantics. *Journal of Artificial Intelligence Research*, 37(1).

Uszkoreit, H., 1986. Categorial Unification Grammars. In Proceedings of the International Conference on Computational Linguistics (COLIN), pp. 187-194.

van Hassel, R., 2009. Lecture Notes on Functional Analysis. Retrieved 06/01/2015 from http://www.win.tue.nl/~rvhassel/Onderwijs/2DE08-0809/OwnLectureNotesFA.pdf. Wagner, J., 2012. Detecting Grammaticality Errors with Treebank-Induced, Probabilistic Parsers. PhD Thesis, School of Computing, Dublin City University.

Witbrock, M. J., & Mittal, V. O., 1999. Ultrasummarization: A statistical approach to generating highly condensed non-extractive summaries. In Proceedings of the 22nd International Conference on Research and Development in Information Retrieval.

Zybura, K., 2004. *Linear Vector Space*. Retrieved o6/01/2015 from http://pirate.shu.edu/~wachsmut/ Teaching/MATH3912/Projects/papers/zybura\_vector \_spaces.pdf.



### A MODIFIED EXPECTATION CONFIRMATION THEORY WITH PERCEIVED TRUST ON INTERNET CAFES USE CONTINUANCE: A CONCEPTUAL FRAME WORK

### \*A. Lawan<sup>1</sup>, B. S. Galadanci<sup>2</sup> and A. A. Abdallah<sup>3</sup>

<sup>1</sup> Department of Information Technology, Bayero University, Kano <sup>2</sup>Department of Software Engineering, Bayero University, Kano <sup>3</sup> School of Computer Science and Information Technology, Linton University College, Malaysia

<sup>1</sup>abd\_wahhb@yahoo.com, <sup>2</sup>bashirshehugaladanci@yahoo.com, <sup>3</sup>dallahso8@yahoo.com

### ABSTRACT

The advent of Internet has propelled the dependency on Information Technology (IT) by enabling millions of Internet users to download software, music, upload digital pictures, buy goods, or search for information or services. Due to the Internet's openness and non-secured structured, trust in usage and services rendered by the Internet might be an issue. Prior research had suggested that the perceived trust has influence on user continuous intention in utilization of information systems platforms. Thus, in order to draw the support of users in exploring the issue that inspires them for continuous patronage of IS platforms, exploring the trust in users is critical for continuous usage of Internet cafe. Understanding the relationship between perceived trust and continuance intention based on a modified Bhattachee's model of expectation-confirmation theory (ECT) is the objective of this paper. An attempt is made in proposing a conceptual frame work of ECT Modification with perceived trust constructs towards Internet café continuance intention. The approached adopted has support from prior literatures. The integration of perceived trust constructs into the ECT model in context of Internet café is vital in understanding the antecedent factors of its continuance intention. Consequently, the proposed model is set forward as the basis of future empirical research with the purpose of validating the proposed framework.

Keywords: Perceived trust, Use Continuance, Internet cafe, Information systems, Digital divide.

### 1. INTRODUCTION

The origin of Internet could be traced to a proposal of the Advanced Research Project Agency (ARPA), the idea was to explore how computers could be connected to a network i.e. (ARPA-NET) to access information on research facilities and universities (Mishra, 2009). Mishra, (2009) reported that in 1969, four computers located at Stanford Research Institute, University of California Santa Barbara and the University of Utah were successfully connected. As time advances, other networks were connected such that by the end of 1971, ARPA-NET traversed the United States of America (USA), and subsequently the connection to Europe became possible. In Nigeria the most populace country in Africa, the Internet was first introduced in 1995 by



the UNESCO-sponsored project tagged Regional Informatics Network for Africa (RINAF) project (Mishra, 2009). Among the several workshop held by RINAF to disseminate the awareness of the Internet, the outcome gave rise to the formation of Nigeria Internet Group (NIG) formed as a nonprofit, non-governmental organization of the primary aim at promoting and facilitating access to the Internet in Nigeria (Mishra, 2009). The RINAF node for West Africa was established at the National Center for Technology Management (NACETEM) at Obafemi Awolowo University. By the end of the 1990s, several Internet service providers (ISPs) emerged, notable amongst the ISP's includes Link serve, Cyberspace, Hyperia, Infoweb, Skannet, and Steineng (Mishra, 2009).

However, the current trend of globalization has echoed the importance of the Internet as a tool for information and conspicuous communication technology (ICT) (Abdulwahab & Mijinyawa, 2012). Yet, Internet is not accessible to the majority of people in developing world most especially the rural inhabitants. The cost of connecting to the Internet has been a challenge (Ani, 2010; Aladeniyi & Fasae, 2013). In most of the developing world, the most noticeable access to Internet is by shared access to Internet café (Bjorn & Kristianson, 2005; Alam, et al., 2009).

Until recently, utilization of Internet services had been the preserve of private sectors in Nigeria (Adenike & Osunade, 2005). The earliest services offered was electronic mail (e-mail) using a dial up. But with development recorded in the technology, Internet services flourished to other sectors such that some organization and individuals get access. The services are provided through various means to end users. Among some of the technology use presently include VSAT, DSL, and Broadband. The VSAT option is mostly used by many organizations and institution due to ease of deployment. Internet service providers are allowed access directly to the internet backbone through VSATs. Recently some fixed wireless operators and GSM operators employed the use of fiber optics in deployments of Internet to customers. Accessibility to Internet is largely through cafés. Inadequate backbone access and capacity are the major obstacles to network growth, accessibility and quality of service. Most Internet connection backbones are set on VSAT with few internets broadband services.

Against the backdrop that Internet access requires telecommunication links and information technologies particularly networks and computer terminals prior research has suggested that communal access to Internet in developing world is much feasible than individual access (Abdulwahab Mijinyawa, & 2012; Alao & Folorunsho, 2008).

The emergence of Internet has increased dependency on IT by encouraging millions of Internet users to download software or music, upload digital pictures, buy goods, or seek information or services, due to the Internet's openness and non-secured structures, generally trust in the services rendered by the Internet became an issue. The continuous use of Internet have brought concerns regarding issue of trust and security of information over the Internet. Prior study has suggested that Trust plays a vital role in helping Internet users overcome perceptions of insecurity and risk (McKnight & Chevrvan, 2002). In surfacing Internet, only Trust makes users comfortable sharing personal information. A Study suggested that the perceived trust has gained growing attention in the context of research related to examining user' acceptance and adoption of innovative technologies in a rapidly changing digital world (Wiedmann, et al., 2010). Perceived trust in the reliability of technical elements and structures, as well as the fairness of other Internet user, is one major determinant of sustainable Internet usage among individuals (Wiedmann et al., 2010; Chang, et al., 2006). Thus, it could be argued that Trust plays a fundamental role in aiding users overcome perceptions of insecurity and risk associated with Internet cafes Usage. Gefen (2000) posited that trust is essential in the acceptance of Internet technologies. This paper proposed a theoretical extension of the Bhattachee's model of expectation-confirmation theory (ECT). The extended model aims at exploring the relationship between perceived trust and Internet café use continuance

### 2. LITERATURE REVIEW



like the theory of Prominent theories planned behavior (TPB) proposed by Ajzen (1991) and technology acceptance model (TAM) and TAM 3 (Davies, 1989; Venkatesh & Bala, 2008) have elaborated on the variables that inspire an individual to accept and use a new IS initiative. Particularly, the (TPB), was developed to supplement the TRA. The additional construct combined to measure intention in TPB is perceived behavioral control (Dillon & Morris, 1996). Particularly, in TPB intention is theorize as the direct determining factor of behavior. Moreover, attitude towards behavior, subjective norm and perceived behavioral control determined intention. Ajzen (1991) defined PBC as "the perceived easy or difficulty of performing a behavior". The TPB posited that the control people have over their behavior varies from behavior that can be done easily to those requiring effort, and resources.

In the same context, the unified theory of acceptance and use of technology (UTAUT) model was applauded to have portrayed a unified view of constructs that inspired the acceptance of IS initiatives (Venkatesh, et al., 2003; Venkatesh, et al., 2011)

Although initial acceptance of IS appears to be an important step toward realizing its success, Use continuance has been applauded to be an appropriate measure of IS success (Bhattacherjee, 2001) Bhattachee's, (2001) argued that long-term viability of an IS and its eventual success depends on its continued use than initial adoption. Accordingly, assessing use continuance is a crucial factor in appraising the IS success (Bhattacherjee, Barnes, 2011). Earlier, expectation 2001; confirmation theory (ECT) proposed by Oliver (1980) explored consumers' satisfaction and their re-purchasing decision behavior. Bhattachee's, (2001) modified the ECT to suit IS context. Researchers adopting the ECT to IS focuses on the use of information systems for operations, decision-making or other organizational activities (Barnes, 2011; Chen, et al., 2010). Although satisfaction as a prominent construct from ECT might play a crucial role in predicting users' continuance intentions, it is not known if satisfaction might be the main factor of Internet café use continuance, though Internet cafés may offer services often on shared manner, prior research has shown perceive trust has potential to attract or distract users (Chang et al., 2006).

Perceived trust is as an important indication of evaluating a user's decision on patronizing Internet café, specifically, users tend to consider the relevant link between perceived trust and their expectations by comparing it with their previous experiences of café usage (Chang et al., 2006). Wiedmann et al. (2010), considered different components of perceived trust in IT-ecosystems (ITEs), the multi-dimensional framework of trust effects proposed that include system-centric as well as user-centric determinants of trust. Based on their conceptual model, two sets of propositions were developed. The first covered technological drivers of trust in ITEs, while the second set of propositions considered individuals as well as social drivers of trust. Their finding suggested the existence of a variety of trust-related determinants in the context of personal interaction with and adoption of innovative technologies. On the one hand, there are aspects that are directly connected with the technology itself in terms of security, privacy, and functionality issues; its usability, including perceived usefulness, ease of use, and information quality, and the presentation of complex information. The researcher posits a better understanding of users' perceived trust and associated risks in interacting with ITE components and living in a digital environment may help to improve the diffusion and adoption of innovative solutions.

Chen et al. (2010) explored the role of Internet selfefficacy adapting the ECT, the finding suggested that the self-efficacy lacks significance on perceived usefulness and continuance intention. Atchariyachanvanich et al. (2006) extended expectation-confirmation theory by introducing a new factor of customer loyalty. The finding suggested that not only basic factors of satisfaction, confirmation , perceived incentives and perceived usefulness also a new factor, customer found loyalty, is to significantly influence the online customers 'intention to revisit the Internet. Fig 1 shows



# 12<sup>th</sup>

### International Conference

Bhattachee's (2001) model of expectation-confirmation theory.



Fig 1 A Post Acceptance Model of IS Continuance

Thus, in order to adapt ECT to a different context theoretical extension associated with the perceived trust is proposed. Therefore, this research intends to incorporate perceived trust construct into the expectation confirmation theory to enhance the understanding of continuance intention of user in perspective of Internet café based on the literature review (Wiedmann et al., 2010; Chang et al., 2006). The conceptual model is presented in Fig 2.



Fig 2 Conceptual Model

### 3.0 The Conceptual Framework

The ECT was adapted to include individual factors (IS continuance, satisfaction, perceives usefulness, confirmation and perceived trust) which could influence continuance intention of Internet café (Bhattacherjee, 2001; Barnes, 2011). Since continuous intention to use IS platform could not only be the factors of satisfaction, confirmation and perceived usefulness but also, user perception

regarding trust. The trust construct could be the main influencing factor of Internet café continuance intention, theoretically, these claims have not been tested empirically. The factors shown in the conceptual model are thus defined based on studies reported by (Bhattacherjee, 2001; Barnes, 2011; Oliver 1980; Chen et al., 2010).

### 3.1 Satisfaction

Satisfaction is the degree of contentment or obtained displeasure by comparing the performance of a product or service to its expected level (Chen et al., 2010). Bhattacherjee (2001) suggested that continuance intention is mainly affected by satisfaction of previous experience in the evaluation of an IS adoption. Satisfaction is increased by positive emotions and decreased by negative emotions that could be attributed to service failure (Chen et al., 2010). Previous study have shown that satisfaction is a determinant of IS continuance intention (Bhattacherjee, 2001). In the context of this study, satisfied users would form a continuance intention, while dissatisfied users discontinue subsequent use of the Internet cafe. Thus, the following hypothesis is proposed:

 $H_{\mbox{\scriptsize A1}}.$  Satisfaction has a positive impact on internet cafe continuance intention

### 3.2 Confirmation/ Disconfirmation

Confirmation is defined as the objective judgment of client on the difference in experience of preexpectation and actual experience (Oliver, 1980). ECT forwarded that satisfaction is affected by expectation and disconfirmation (Chen et al., 2010). Literarily, disconfirmation indicates the gap between expectations and perceived performance. Bhattacherjee (2001) argued that the confirmation is positively related to satisfaction with IS use because it implies realization of the benefits expected of IS use. whereas disconfirmation (perceived performance lagging expectation) indicates failure to achieve expectation. Furthermore, Bhattacherjee (2001) proposed that confirmation has a positive influence on satisfaction and perceived usefulness. Thus, the following hypotheses are proposed:



 $HA_2$ . Confirmation has a positive impact on satisfaction.

HA<sub>5</sub>. Confirmation has a positive impact on perceive usefulness

### 3.3 Perceived Usefulness

Perceived Usefulness is the degree in which an individual believes that using a particular IS would enhance his/her job performance (Davies, 1989). Usefulness is quantified into three; job performance, productivity and time saving. Accordingly, using IS enhances job performance, productivity and makes jobs to be accomplished on time (that is, by reducing the time to carry out a task or providing timely information (Davis, Bagozzi & Warshaw, 1989; Karahanna, Straub and Chervany, 1999). In the context of this study, the quality and variety of services offered by internet cafés operators will affect users' perceived usefulness. Thus the following hypotheses are proposed.

 $H_{A_3}$ . PU has a positive impact on satisfaction.

 $H_{A4}. \mbox{ PU}$  has a positive impact on continuance intention

#### 3.4 Perceived Trust

Perceived Trust is define as relying or depending on infrastructure systems like the Network or relying on specific information systems application software like Microsoft Excel<sup>TM</sup>((McKnight, & Chevrvan, 2002) . Trusting belief in an IT's competence means the IT is perceived to have the functionality or functional capability expected by the trustee to do some task (McKnight, & Chevrvan, 2002). Trust in information technology has interesting implication. Trust in IT influence use or adoption of a technology. For instance, unless one trusts a software product to be reliably

- Abdulwahab, L. and Mijinyawa, M.K. (2013). Cybercafés' Use Continuance : An Empirical Test of a Research Model . American Association of Scholarly Research Journal (AASRJ), 6(4), 91-103.
- Adenike, O., and Osunade, O. (2005). ICT Infrastructure available in Nigeria educational

otherwise, why would one continue using it? Thus the following hypotheses are proposed.

 $H_{\mbox{\scriptsize A6}}.$  PT has a positive impact on continuance intention

### 4. CONCLUTION

The continuous use of Internet café have yielded concerns for users regarding the trust in infrastructure and services that can make internet cafes clients to use the services rendered. Review of the previous literatures on IS continuance, reveals dearth of research that essentially addresses the issue of incorporation of perceived trust in post acceptance model of IS continuance. This paper under studies the Bhattachee's model of expectation-confirmation theory and proposed a modification particularly with regard to perceived trust, the need for the incorporation of the perceive trust construct in to Bhattachee's model in context of Internet cafés. Perceived trust could be a fundamental factor when accessing the use continuance of Internet cafés. This paper only proposed a conceptual frame work for measuring the proposed construct in context of Internet cafés. The research presented is limited being that, the proposed model is only based on literature review. Subsequent studies should focus on empirical validation of the conceptual model using the survey instruments.

### 5. ACKNOWLEDGEMENT

The author(s) acknowledged the support from Bayero University, Kano for sponsoring the research from the beginning to the present stage of presentation in a conference.

### 6. **REFERENCES**

and research institute. *Journal of Computer sciences*, 4(3), 56-75.

- Ajzen, I. (1991). Theory of planned behavior. Organizational Behavior and Human Decision Processes, 179-211.
- Aladeniyi, F. and Fasae, L.K. (2013). Use of cybercafé for internet access by the students of Rufus



## 12<sup>th</sup>

### International Conference

- Giwa Polytechnic, Owo, Nigeria", Program:. electronic library and information systems, 47(1), 4 - 14.
- Alam, S;, Abdullahi, Z. and Ahsan, N. (2009). Cyber Café usage in Malaysia: An exploratory study. Journal of Internet Banking and Commerce,, 14(1) , 1-13.
- Alao, A., and Folorunsho, A. L. (2008). The use of cybercafes in Ilorin, Nigeria. *The Electronic Library*, ., 26(2), 238 24, 2008.
- Ani, O. (2010). Internet access and use: A study of undergraduate students in the Nigerian universities. *The Electronic Library*, 28(4), 555-567.
- Atchariyachanvanich, K., Okada , H, and Sonehara, N. (2006). What Keeps Online Customers Repurchasing through the Internet? ACM SIGecom Exchanges, 6(2). 47-57.
- Barnes, S. (2011). Understanding use continuance in virtual worlds:An: empirical test of a research. *Information & Management*, 48, 313–319.
- Bhattacherjee, A. (2001). Understanding s continuance: an expectation-confirmation model,. *MIS Quarterly*, 25(3). 351-370.
- Bjorn, F., Kristianson, S, and Wahid, F. (2005). Information dissemination in a developing society: Internet cafe users in Indonesia. *EJISDC*, 22(3). 1-16.
- Chang, E., Dillon, T, and Hussain, F. (2006). Trust and reputation for service-oriented environments: Technologies for building business intelligence and consumer confidence. *Chichester: Wiley*.
- Chen, Y. Y., and Chen Lee, Y. (2010). Confirmation of Expectations and Satisfaction with the Internet Shopping: The Role of Internet Self-efficacy. *Computer and Information Science*, 3(3),. 14-22.
- Davies, F. (1989). Perceived usefulness, perceived ease of use and user acceptance of information technolog. *MIS Quarterly*, 13(3), 319-340.
- Davis, F., Bagozzi, R.P., and Warshaw, P.R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, 35(8), 982-1003.

- Dillon, A., and Morris, M. (1996). User acceptance of information technology theories and models. Medford NJc: Information Today, In
- Gefen, D. (2000). E-Commerce: The role of familiarity and trust. *omega*, 28(6), 725-737.
- Grace, A., and Lenny, R. (2008). The level of internet access and ICT training for health information professionals in sub-Saharan Africa. *Health Information and Libraries Journal*, 25(3), 175-185.
- Haseloff, A. (2005). Cybercafés and their potentials as community development tools in India . The Journal of Community Informatics, 1(3),29-46.
- Karahanna, E., Straub D.W, and Chervany, N. (1999). Information Technology Adoption across Time: A Cross-Sectional Comparison of Pre-Adoption and Post-Adoption Beliefs. *MIS Quarterly*, 23(2), 183-213.
- McKnight, D., and Chevrvan, N.L. (2002). What trust mean in E-Commerce customer relationship: An interdisciplinary conceptual typology. International Journal of Electronic Commerce, 6(2),35-39.
- Mishra, M. (2009). Use and importance of the Internet at the University of Maidugur Nigeria. *First Monday*, 14(3).
- Oliver, R. (1980). A Cognitive Model for the Antecedents and consequences of Satisfaction. Journal of Marketing Research, 17, .460-469.
- Venkatesh, V. S. (2011). Just what the Doctor ordered: A revised UTAUT for EMR system adoption and use by Doctors. .44th Hawaii International conference on system sciences, (pp. 1-10). Hawaii: hicss.
- Venkatesh, V., Morris, M.G, Davis, G.B , and Davis, F.D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27, 425,478.
- Venkatesh, V., and H, B. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39, 273-315.
- Wiedmann, K., Hennigs, N, ", Varelmann, D, and Reeh, M. (2010). Determinants of Consumers' Perceived Trust in IT-Ecosystems. Journal of Theoretical and Applied Electronic Commerce Research, 5(2),137-154.



- Zhu, W., and Fui-Hoon, N. (2002). Factors Influencing Adoption of Mobile Computing. Proceedings of the Conference on Issues and Trends of IT Management in Contemporary Organizations (IRMA). Hershey, PA: Idea Publishing Group.
- Zulkhairi, M., Nor Iadah, Y, Huda, I, Khairudin, M, and Zahur, M. (2011). Socio–economic benefits of Telecentre implementation in Peninsular Malaysia. in. 2nd Proceeding of the International Conf. 2009. Computing & Informatics, (pp. 374-376).



### PERFORMANCE EVALUATION OF APACHE WEB SERVER ON CLOUD COMPUTING ENVIRONMENTS

### \*Lawal Sulaiman<sup>1</sup> and Aminu Mohammed<sup>2</sup>

<sup>1,2</sup>Department of Mathematics, Computer Science Unit, Usmanu Danfodiyo University, P.M.B. 2346, Sokoto, Nigeria

<sup>1</sup>lawal.sulaiman@udusok.edu.ng, <sup>2</sup>maminuus@yahoo.com

#### ABSTRACT

The emergence of the cloud computing concept has changed the way information technology (IT) services are developed, deployed, used, maintained and paid for. Due to the increased in demand for IT resources and the cost of managing and maintaining IT infrastructure, companies providing IT services and users have started deploying their services to cloud environment in order to obtained data reliability and minimize their overhead. However, obtaining data reliability and minimization of this overhead depends on the appropriate selection of web server and cloud computing environment. Despite many attempts to evaluate the performance of these web servers, most of the performance evaluation carried out mainly focused on a single cloud environment. This paper evaluates the Performance of apache web server on 3 major cloud computing environments (i.e. HP Cloud, City Cloud and Rack space cloud computing) using execution time and throughput as performance metrics in order to examine in which of the 3 environments does Apache web server performs better. Results of the analysis have shown that Rack space outperform the other 2 environment (i.e. HP and City Cloud) in almost all the workloads because it has the lowest execution time and highest throughput than the 2 environments.

Keywords: Cloud computing, Apache web server, Execution time, Throughput, Performance evaluation

#### 1. INTRODUCTION

Cloud computing refers to the delivery of computing resources over the Internet. Instead of keeping data on your own hard drive or updating applications for your needs, you use a service over the Internet, at another location, to store your information or use its applications. Examples of cloud services include online file storage, social networking sites, webmail, and online business applications (Armbrust et al., 2009).

Peter & Timothy (2011) defined the concept as "a model for enabling 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" Cloud computing has grown up to a mature industry standard supported by many companies (Heena & Naghma, 2013). There are several major cloud computing providers including Rack Space, HP Cloud, Amazon, Google, City cloud, Microsoft Windows Azure etc that are providing cloud computing services. It provides access to resources like online files, data, program, online business



applications etc through web browsers via Internet. Cloud Computing describe applications that can be accessible through the Internet and for this reason, large data centres and powerful servers are used to host the web applications on web servers (Faisal & Majid, 2012).

The aim of this research is to evaluate the performance of Apache web server on cloud computing environment. While the specific objectives includes to measure and compare the average execution time and average throughput using sieve algorithm respectively in cloud environments.

### 2. RELATED WORK

The performance of Apache and Nginx web servers was analysed by (Dabkiewicz, 2010) using request per second as performance metric. Apache Benchmarking tool was used to test the performance of web servers using static, dynamic and image files. Even though, experimental results shows that Nginx outperforms apache on static files while Apache outperforms Nginx on dynamic files but the researcher conducted the experiment using benchmarking tool not on the real site and consider only one performance metric in the experiment (i.e request per second). Testing the performance on real website using other important metrics like CPU load, RAM usage will give more insight on the performance because increasing the system resources means increasing the performance of that system.

Analysis and evaluation of high performance web servers using CPU load and RAM usage as performance metrics was examined by Albert(2011) where Apache HTTP server benchmarking tool was used to test the performance of different web servers. Static test, dynamic test and load test were carried out in the experiment.

The author did a comprehensive work in evaluating the performance of various web servers using CPU load, RAM usage as performance metrics. However, conducting the research by testing other important metric like round trip times (RTT) is very important because RTT estimates the time required for a packet to travel from a specific source to a specific destination and back again, such estimates are used to ensure that data is reliably delivered. Chao, (2012) compared the performance evaluation of two popular cloud computing platforms (Google App Engine and Amazon) and traditional web servers using Round-trip time (RTT) as performance metrics. Three data sizes are chosen for the experiment, i.e small image (12kB), medium image (350kB) and large image (1MB).

From the experiments, result shows that cloud computing platforms perform better than traditional web servers. Even though only one performance metric (Round-trip time) was measured in the experiment. However, the experiment can be conducted using other cloud computing environment with different performance metrics like execution time.

Faisal & Majid, (2012) compared the performance of apache web server in cloud computing and local environment using execution time and throughput as performance metrics. The result shows that the execution time of the cloud environment took more time to accomplish than the local environment. Also the throughput of apache web server is better in local environment than in cloud environment.

Even though the experimental results shows that Apache web server performed better in local environment than cloud environment. However, the researchers conducted the research using only one cloud provider. Thus, the need for evaluating the web server using different cloud computing environments such as HP Cloud, Amazon, Windows Azure, Rackspace etc. .

### 3. METHODOLOGY

This section presents an overview of the methods employed in conducting the research. It shows the steps involved in evaluating the performance of apache web server as well as the tools used to conduct the experiment.

### **3.1 Experimental Procedures**

In order to measure the performance of Apache web server on cloud computing, the experiment was conducted using benchmarking method;


Benchmarking is the process of comparing the performance for two or more systems by measurements (Jain, 1991). The method was also adopted by other researchers like Faisal & Majid, (2012). Benchmarking was used to measure the performance of Apache web server on three cloud computing environments.

The three Clouds computing environment selected for the experiment includes HP, Rackspace and City cloud. An account was created on each of the three cloud environment after verifying the account a server was created, apache web server was installed on each environment, php application was hosted into apache web directory (www) to measure the execution time and throughput of Apache web server. Forty (40) experiments were conducted using different workloads and the result of each experiment was measured and recorded. Table 1 shows the specifications for the 3 cloud computing environments.

Description	Rackspace Cloud	HP Cloud	City Cloud
Memory	4 GB	4 GB	4 GB
Processor	2 cores CPU	2 cores	2 cores
		CPU	CPU
Hard Disk	60 GB	60 GB	60 GB
Operating	Ubuntu	Ubuntu	Ubuntu
System	Linux 12.04	Linux	Linux
	server	12.04	12.04
	version 64-	server	server
	bit	version	version
		64-bit	64-bit

#### 3.2 Experimental Setup

The experiment was conducted by setting up Apache web server on 3 cloud computing environments. Apache web server was considered Table 1: Specifications for the 3 cloud computing environments

for the experiments because it is currently one of the most popular web server technologies in use today; it is also used and adopted by many organizations and companies (Netcraft, 2014). Apache web server was installed in each of the three cloud computing environments (Rackspace, HP and City Cloud computing)

#### 3.2.1 Setting up Apache Web Server on 3 Cloud Environment

The following procedures were followed to setup Apache web server on 3 Cloud Computing Environments,

- (i) Account was Created on each of the three Cloud computing environment namely: Rackspace, HP and City Cloud.
- (ii) Billing information for verification of accounts was provided
- (iii) A server was created by selecting the required specifications i.e. the server name, operating system, memory, processor and hard disk. IP address was assigned to the server automatically.
- (iv) Putty software was installed on the client computer for remote login to the server.
- (v) After login into the server, Apache web server version 2.2.20 and PHP version 5.0 was installed on the server, other software installed includes dpkg-dev and php5-dev.
- (vi)Sieve algorithm written in PHP script was hosted into Apache web directory (www) and accessed via a web browser using the IP address of the server and the name of the script. (e.g. http:// 192.168.0.16/execution.php).

Rack space and HP Cloud Computing Environments were choosing for this experiment because their billing is cost effective and their procedure of configuration and testing is easier to understand compared to other cloud computing environment.

#### 3.3 Performance Metrics

Two performance metrics were employed to evaluate the performance of Apache web server on cloud computing environment. They are: Execution Time and Throughput

(i) Execution Time: Is the time spent by a web server in executing a task or program. (i.e. the actual time spent in executing or running a program). The time start (the time program started executing) and time end (the time program finished



## 12<sup>th</sup>

### International Conference

executing) were recorded and the execution time was computed using the expression in (1)

Execution Time = Time End - Time\_Start (1)

(ii) Throughput: Is the maximum number of simultaneous requests served successfully per second by a web server (Faisal & Majid, (2012)). Throughput can be computed as:

Throughput = No. of successful requests/ Total time. (2)

#### 3.5 Workload

Performance of a Computer system is generally evaluated and measured by using a specific workload. Workload is the amount of processing that the computer has been given to do at a given time. One of the popular workload to compare systems, microprocessors and high level languages is sieve kernel. It is based on Eratosthenes algorithm and is used to find all prime numbers below a given number n (Jain, 1991).

In order to measure the performance of Apache web server on cloud computing environment, a sieve algorithm written in PHP script was implemented. Sieve Algorithm is a simple, ancient algorithm for finding all prime numbers up to any given limit. The algorithm, in its manual form, consists of first writing down all integers from 1 to n and then striking out all multiples of k for k = 2,3,..., n (Jain, 1991).

Two scripts used by Faisal & Majid, (2012) was adopted in this research, one for execution time and the other one for throughput. The scripts calculate the time spent for executing prime numbers.In this experiment, five workloads (i.e., 10,000, 100,000, 1,000,000, 5,000,000 and 10,000,000) has been defined each of which serves as an input to measure execution time and throughput of Apache web server.

#### 4. PERFORMANCE EVALUATION

This section presented the results obtained from the experiment in a graphical format. Five workloads have been selected to serve as an input to measure the execution time and throughput of Apache web server.

#### 4.1 Execution Time

In order to measure the execution time of Apache web server, Apache web server was installed in each of the 3 cloud computing environments, Execution time's script was hosted into apache web directory (www) and access via a web browser using the IP address of the server and name of the script (e.g., 192.168.0.16/execution.php). The script was executed 40 times for each of the 5 workloads (10,000, 100,000, 1,000,000, 5,000,000 and 10,000,000). The execution time was computed using equation 1

The average execution time was presented in figure 1.



Figure 1: Execution Time for 3 cloud computing using different workloads

Figure 1 shows the execution time of 3 cloud environments for 5 workloads (i.e.10,000 to 10,000,000), it can be seen from the graph that for workload 10,000, HP cloud computing environment performed better than the entire environments because it took less time to execute i.e. 0.006474 seconds followed by Rack Space 0.006995 seconds and City Cloud 0.018089 seconds respectively.

For workload 100,000, it can be observed that Rack Space cloud environment performed better because it executes faster i.e. 0.128858 seconds followed by City Cloud 0.31395 seconds and HP Cloud took 0.353925 seconds respectively.



For workload 1,000,000, Rack Space execute faster because it took 2.802829 seconds to execute followed by City Cloud 3.360128 seconds and HP took 8.319734 seconds respectively.

For workloads 5,000,000, both Rack space and City cloud took almost the same time to execute (26.24685 and 26.7204 seconds respectively) followed by HP 76.43266 seconds.

For workloads 10,000,000, Rack space perform the best by taking less time to execute i.e. 68.75184 seconds followed by City Cloud 69.39657 seconds and HP took 152.8653 seconds respectively.

Thus, the result obtained indicated that Rack space performs better for about 80% across all workloads; therefore it is the best among the 3 cloud computing environments.

#### 4.2 Throughput

The throughput was measured by repeating the experiments 40 times and in each of the experiments the total execution time against each workload for each cloud were recorded and throughput was calculated and the average of all the 40 tests were computed and presented in Figure 4.2.



environment using different workloads

Figure 4.2 depicted the throughput of 3 cloud computing environments for workloads 10,000 to 10,000,000.

It can be seen from the graph that for workload of 10,000 HP cloud perform better than the rest of the

environments because it has the highest throughput of 1543.211 seconds followed by Rack space and City Cloud respectively while For workload 100,000 Rack space outperform with highest throughput of 76.85678 seconds followed by City cloud with 69.80531 seconds and HP with 27.69643 seconds.

For workload 1,000,000 Rack Space has the highest throughput of 3.525636 seconds followed by City Cloud and HP cloud respectively.

For workload 5,000,000 it can be observe that Rack Space also perform the best with highest throughput of 1.931868 seconds followed by city cloud and HP respectively.

For workload, 10,000,000, Also Rack Space outperform with highest throughput of 0.14712 seconds followed by city cloud and HP respectively. Thus, the result of the experiment shows that Rack space outperformed by 80% across all workloads because it has the highest throughput, and therefore is the best among the 3 cloud computing environments used in this study.

#### 5.0 CONCLUSION

Performance evaluation of Apache web server on cloud computing environment was presented in this paper using execution time and throughput as performance metrics. The experiment was conducted by Setting up Apache web server on three Cloud computing environment (HP Cloud, Rack Space and City Cloud). In this experiment, five workloads (10,000, 100,000, 1,000,000, 5,000,000 and 10,000,000) have been selected to serve as an input to measure the execution time and throughput of Apache web server.

A Sieve algorithm written in PHP script that was hosted into Apache web directory (www) was used to compute the prime numbers on 3 cloud environments. Forty (40) experiments were conducted for each workload and the execution time and throughput against each experiment was recorded and analysed.

The results analysis shows that Rack space cloud computing environment outperforms the other environments (i.e. HP and City cloud) in almost all the workloads because it execute faster than the 2



environment. It has also been observed that as the complexity (workloads) increases the Rackspace cloud environment continue to outperforms the other environment. For throughput, it has been observed that Rackspace cloud environment performs better than City Cloud and HP cloud in almost all the cases because Rackspace clouds has the highest throughput than the 3 environment. It has also been observed that the throughput decreases with the increase in workloads (i.e. the throughput for a less complex task is more than that of complex task).

#### **Future Work**

This research can be extended by conducting the same experiment to compare the performance of proprietary and open source web servers on cloud computing environment using other measurement techniques like benchmarking. Furthermoreother major providers can be consider using other performance metrics.

#### References

Albert, H. B. (2011):Analysis and evaluation of high performance web serversUnpublished master's dissertation. Department of Science and Telecommunication Engineering & Management, University of Catalonia.

Armbrust, M.et al. 2009,Above the Clouds: A Berkeley View of Cloud ComputingTechnical Report2009,viewed 7 November 2014, <<u>http://www.eecs.berkeley.edu/Pubs/TechRpts/200</u> <u>9/EECS-2009-28.pdf</u>>

Chao H. (2012): Performance Analysis based on two Leading Cloud Computing Platforms: Google App Engine and Amazon Web Service. [Online] Available at:http://www.cse.wustl.edu/~jain/cse567-

11/ftp/clouds/index.html.[Accessed: 27 December 2013].

Dabkiewicz, S. (2010): Web server performance analysis. [Online] Available at:.

#### https://www.os3.nl/\_media/2011-

2012/courses/lia/sebastian\_dabakiewicz\_-\_liareport-sebastian-dabkiewicz.pdf [Accessed 27April 2013]

Faisal, A. & Majid, K. (2012): Web Server Performance Evaluation in Cloud Computing and Local Environment (Unpublished master's dissertation) School of Computing Blekinge Institute of Technology Karlskrona, Sweden. pp. 1 – 74.

Heena, I. S., & Naghma A. B. (2013): Survey on Cloud Computing:. International Journal of Emerging Technology and Advanced Engineering. 3 (4), 308 – 312.

Jain, R. K (1991): The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling. New York: Wiley Computer Publishing, John Wiley & Sons, Inc. 1-714.

Netcraft. (2014): Web Server Survey. Available: http://news.netcraft.com/archives/2014/05/07/may-2014-web-server-survey.html. [Accessed 10 August 2014].

Peter, M. & Timothy, G. (2011):Definition of Cloud Computing:Recommendations of the National Institute of Standards and Technology.viewed 10 August 2014, http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf



### AN INVESTIGATION ON THE PERFORMANCE OF SOME NIGERIAN MOBILE BROADBAND NETWORK SERVICES

#### <sup>\*1</sup>Sirajo Shehu; <sup>2</sup>Ahmed Yusuf Tambuwal; and <sup>3</sup>Aminu Mohammed

Department of Mathematics, (Computer Science Unit) Usmanu Danfodiyo University Sokoto, P.M.B 2346, Sokoto, Nigeria

<sup>1</sup>siiraa2001@yahoo.com, ahmed\_tambuwal@yahoo.com, maminuus@yahoo.com

#### ABSTRACT

Many modern Internet applications require fast client-server interaction to maximize user interactivity. This need has caused many developed countries to invest in super-fast broadband access networks, with current speeds of 20Mbps to 80Mbps. This super fast broadband service is generally deployed using either wire lines or wireless infrastructure. However, wire line infrastructure deployment is quite expensive to set up. As such most cell phone operators provide this broadband service using wireless infrastructure. Nevertheless, there is little or no empirical data to evaluate the performance of most of these mobile broadband network services offered by the different operators. Our framework has been designed to produce one sample location for the three networks on 3G technology. We did this by setting of one test site in Usmanu Danfodiyo University Sokoto (UDUS) where 3G is available on MTN, Airtel, and Etisalat network. Data were captured from the Networx simulator on three systems and recorded manually which was further processed using Microsoft Excel. The performance evaluation of the three selected mobile broadband network (i.e. MTN, Etisalat and Airtel) within UDUS main campus was carried out using host-based empirical measurement tool "Networx Speed meter". The 'Network Speed meter' was used to monitor and collect the real-time Internet upload and download throughput at different times of the day. The main result shows that MTN is currently the best performing network, with highest average throughputs for both data uploads and downloads.

Keywords: Broadband, Download speed, Upload speed, Network applications

#### 1. Introduction

In modern academic institutions, there is a growing necessity for a range of high-bandwidth Internet applications such as E-learning, Online Web Seminars, Voice and Video Conferencing, E-medicine, and Desktop Sharing. Even commonly used applications such as Gmail, cloud computing and online chats now require broadband Internet, due to use of complex graphics, embedded audios and videos. Also the need for high speed internet connection has caused many developed countries to invest in super-fast broadband access networks, with current speeds of 20Mbps to 80Mbps. (Broadband Commission. (2013)). Broadband generally refers to "any high-speed Internet connection that offers



integrated access to voice, high-speed data, and interactive delivery services" (Dazhi, 2013).

Broadband within the Nigerian context is defined as "an Internet experience where the user can access the most

demanding content in real time at a minimum speed of 1.5Mbit/s"(Nigeria's Presidential Committee on Broadband, 2013). In practice however, so called "broadband access" is most times a highly frustrating experience for many Nigerian subscribers, with very few

data download or upload sessions being able to complete speedily, if at all. Nigerians living in semi-urban areas (such as many areas in Sokoto) are at a particular disadvantage in accessing broadband data and services, which prevents them from fully utilising the Internets potentials. On the other hand, Mobile broadband is "the marketing term for wireless Internet access delivered through mobile phone towers to computers, mobile phones (called "cell phones" in North America and South Africa), and other digital devices using portable modems"(Ergen, 2009).

Currently, Nigeria has six submarine cable landings, which provide up to 9Tbits/s of combined Internet access capacity to its shores. However, the fact that all the landings are in Lagos, seriously limits maximum transmission speed in other parts of the country due to distribution an inadequate network (Nigeria's Presidential Committee on Broadband, 2013). Unlike many countries where subscribers have a wide range of cable and wireless broadband access providers to choose from, most of over 136 million Nigerians rely on four national mobile network operators for broadband Internet services (NCC Operator data (2014)). This 'monopoly' situation presents a major problem because there is every tendency for the dominant operators to be negligent or even non-compliant by providing poor network services (below service contract) without any fear of loosing a significant market share. This has far reaching social implications because broadband is a basic

utility for societal transformation and development, necessary for all segments of society in the modern world. There are also economic costs such as loss of businesses and commercial revenues that would have been earned (e.g. through research activities), loss of savings due to transport costs that could otherwise be saved with broadband access, and also loss of job opportunities.

The rest of this paper is organised as follows: Section 2 reviews a set of recent literatures that are relevant to the work presented in this paper. Section 3 explains the methodology and tools used to carry out the work. Section 4 presents a summary of results with analysis and discussion. The paper concludes in Section 5 with presentation of future work.

#### 2. Related Works

Oborkhale et, al, (2012) evaluated the Quality of Service (QoS) of data service of Etisalat wireless cellular network in Nigeria using Ping and Trace route tools. Their study compared how well the Etisalat broadband network performs in Yola, Nigeria, at different periods of the day, using analytical hierarchy process method. They used jitter, latency, data loss and throughput as the basic metrics for evaluating the QoS.

Their results showed that the Etisalat network offers the best QoS at night followed by morning, evening and then afternoon. However, this study evaluated the enduser experience based on data service of a single network (Etisalat). The results would be more useful if the experiment is conducted for multiple networks, and using different performance metrics like download/upload speed, and data volume for download/upload.

Temidayo, et al, (2014) evaluated the performance of mobile data networks for MTN, Airtel and Globacom (GLO) using host-based experiment approach. They measured the Hypertext Transfer Protocol (HTTP) download/upload rates, download/upload volumes and daily average data volumes for a period of fourteen days. The main interest of this work was to evaluate impact of



rainfall on respective network performances. The study was carried out using the "Networx" tool.

Their results showed that while rainfall had little impact on performances for all the networks, MTN network service was more consistent during rainfall with highest peak download speed (22.2Kbits/s) compared to the other networks (GLO and Airtel). Although the authors conducted their experiment with modems of different technology and using Networx software tool, the experiment would be more valid if conducted using same cellular technology (such as HSPA+ modem) for all the networks.

Siwakoti (2014) evaluated the performance of four Norwegian mobile broadband networks under mobility using delay, packet loss, and connectivity as performance metrics. The experiment was conducted using "Nornet edge tool" to monitor/measure the performance of the four mobile data operators.

The result of the experiment showed that mobile broadband performance is affected by mobility, but the degree varies with respect to performance metrics. Also different mobile data operators perform differently. Even though three performance metrics (delay, packet loss and connectivity) were used in the experiments, the distance of the point of experiment from the base station for all the operators chosen for testing was not the same and this may affect integrity of the final results. Additionally, the experiment can be conducted using other tools to monitor /measure and evaluate the effect of mobility on mobile broadband performance using different performance metrics like: latency, web response time, download/upload speed at the same distance away from the base station of all the mobile data network operators.

#### 3. Material and Method

The work presented in this paper compares performances of three mobile broadband networks (MTN, Etisalat, and Airtel Nigeria) at UDUS. Host-based empirical measurement was employed, using Networx tool to simultaneously monitor the real-time Internet

upload and download throughput of the three networks. Our framework has been designed to produce one sample location for each network on 3G technology. We did this by setting of one test site in UDUS where 3G is available on MTN, Airtel, and Etisalat and data were captured from the Networx simulator and recorded manually. Networx was chosen because it permits close supervision of uploads and downloads and has a speedmeter feature that automatically keeps track of downloads/uploads via any chosen network interface. The Networx tool also provides a clear graphical display of the numeric average and peak throughput values. Significantly, it is free and does not contain any spyware/malware, thus can be used to monitor all types of network connections e.g. Dial-up, Cable modems, Asymmetric Digital Subscriber Lines (ADSL), and Ethernet network interface cards.

All the experiments in this work are stationary and carried out at a location that is centrally located between the base stations of the three networks. Each experiment was conducted by plugging each of the three 3G HSPA modems for MTN, Airtel & Etisalat Cat 7 (7.2Mbps Upload, 5.67Mbps Download speed) on three identical laptops (Intel(R) Core(TM) i3-3120M CPU@ 2.50GHz, 4GB Memory, 64bits System, Windows 8 OS). Each of the laptops was pre-installed with the network device's software, and Networx software version 5.3.3.0. The three network interfaces to be monitored were also preconfigured on the laptops respectively, namely Easynet, MTN F@stLink and Airtel Broadband.

To measure the performance of wireless mobile broadband network, the three laptops were connected to the Internet via the interfaces being monitored. The Networx software was loaded and its speed meter activated, which was used to measure the network device throughput for incoming and outgoing applications data. At this point, an upload was initiated via 4Shared and Gmail applications while a download was initiated via iLivid setup-r20-n-bc.exe version 5.0.2.4813 and YouTube application.



All the measurements were taken at Usmanu Danfodiyo University Sokoto main campus for a period of two weeks in March 2015. For each week, experiments to test the mobile broadband performance for data download and upload were conducted on four different days comprising three working days (between Monday to Friday) and one weekend day (Saturday or Sunday). Each experiment was carried out at four different periods of the day (morning, afternoon, evening and night). For each period, experiments were run within two hours, consisting of eight different upload and download sessions (i.e. thirty two upload and download sessions per day). For each session, the following data were recorded in Microsoft Exel: Average Throughput (Download/Upload Rate), Peak Throughput and Type of Application/Service which were later processed to hourly, daily and weekly average throughput. Finally, Microsoft Excel was used for computations and plotting of graphs.

#### 4. Results and Discussion

This section discusses results obtained from preliminary investigation of performances of MTN, Etisalat, and Airtel mobile broadband networks at UDUS. The key performance metrics measured were

- Average Throughputs (Download and Upload)
- Peak Rates (Download and Upload)

#### 4.1 Average Download Throughput

Figure 4.1 shows the average download rates of MTN, Etisalat, and Airtel at UDUS, Sokoto monitored over the two-week period in March 2015. As seen on the graph, MTN and Airtel have similar performances with average download throughputs of 1.86Mbps and 1.84Mbps respectively. Etisalat has the lowest average download throughput of 1.20Mbps, which is about 35% lower than both MTN and Airtel.



On closer inspection of the individual data readings (shown in Figure 4.2), it is clear that the performance of all the three networks fluctuates significantly at different times. For example, on few occasions, the average download throughput of MTN network exceeds 5Mbps, while it is less than 2Mbps at many other intervals. This inconsistency is quite frustrating for network users as it is not possible to rely on any of the networks for certain applications that require steadily-high throughput or low jitter. It is also noteworthy that Etisalat network performance appears to be particularly low during the first week of the experiment (readings 1-32) compared to the second week (readings 32-64). On the other hand, MTN exhibits almost the opposite behaviour. This will be further investigated.



Hours of the Experiments for two weeks(H) Fig.4.2 Showing Download Throughput at UDU Sokoto for two week



#### 4.2 Average Upload Throughput

Figure 4.3 shows the average upload rates of MTN, Etisalat, and Airtel at UDUS, Sokoto monitored over a two-week period in March 2015. In the upload case, MTN network clearly outperforms the other two networks with average upload rate of 1.7Mbps compared to 855.71Kbps and 695.13Kbps for Etisalat and Airtel networks respectively.

Interestingly, the measured average upload rate (1.7Mbps) for MTN network is only slightly lower than its average download rate (1.86Mbps), while for Airtel network, the average upload rate (695.13Kbps) is less than half of its average download rate (1.84Mbps). This result may be an indication of how the two networks allocate their wireless spectrum for uploads and downloads. It is quite common for mobile wireless networks to allocate larger share of bandwidth for data downloads than uploads, because the average person downloads more content than he uploads (Alcatel-Lucent, 2012).



Figure 4.4 shows that similar to the download case; MTN network upload performance has the highest fluctuation rate (varies between 500Kbps-3Mbps) compared to both Airtel (varies between 200Kbps-1.8Mbps) and Etisalat (varies between 500Kbps-1.4Mbps) networks. However when the download and upload rates are compared for all the three networks, the fluctuation rate during data uploads (between 200Kbps-4Mbps) is lower than when performing downloads (between 200Kbps-600Kbps).



#### 4.3 Performances at Different Times of Day

The final set of results compares performances of the three networks at different times of the day. These results are summarized in Table 4.1 and Table 4.2. The individual graphs are also shown in Figures 4.5 and 4.6.

For Etisalat and Airtel networks, there is not a significant performance difference seen for both data downloads and uploads. However, MTN performs much faster downloads at afternoon compared to all other times of the day. On the other hand, the MTN network performs much slower uploads in the morning compared to all other times of the day.

Table 4.1 Average Download rates in (Kb	ps	)
---	----	---

-	-		• • • •	
Network	Morning	Afternoon	Evening	Night
s				
Airtel	1079.2	1240.8	1139.2	1240.8
MTN	1710.4	2053.6	1780.8	1821.6
Etisalat	1473.5	1872.8	1638.4	2457.6

Table 4.2 Average Upload rates in (Kbps)

Network	Morning	Afternoon	Evenin	Night
s			g	
Etisalat	902.4	876	823.2	808
Airtel	621.6	768.8	750.4	640.8





Fig.4.6 Chart Showing the Average Upload Throughput for two weeks at UDUS for the 3 Networks

#### 5. Conclusions and Future Work

This paper has presented preliminary findings on performance of three mobile broadband networks at the main campus of Usmanu Danfodiyo University, Sokoto. Based on the results obtained, the following conclusions can be made:

• Two of the three networks providers namely MTN and Airtel have average download throughput of (1.86Mbps and 1.84Mbps respectively) this means they provide true broadband service within the Nigerian context (average speed of 1.5Mbps), but

far below values advertised by the mobile (7.2Mbps) broadband service providers of download speed, while Etisalat did not meet true broadband service as per definition within the Nigerian context (average speed of 1.5Mbps) and far below values advertised by the service providers (7.2Mbps) download speed. This has far reaching implications, as many Internet-based applications and services cannot function properly at the measured rates.

- During the study period, MTN was the best performing of the three networks with average download and upload throughputs of 1.86Mbps and 1.7Mbps respectively. However, MTN network performance was also seen to be highly unstable, in terms of throughput fluctuations.
- With an average download throughput of 1.8Mbps, Airtel outperformed Etisalat, which had an average download throughput of 1.2Mbps. However for data uploads, Etisalat outperformed Airtel with average throughputs of 855.71Kbps and 695.13Kbps respectively.

Our results also showed that MTN network has similar performances for both downloads and uploads, while the other two networks perform faster downloads compared to data uploads. Similarly, MTN network was seen to perform much faster downloads during the night period, while it was seen to be particularly slow for uploads in the morning period. While the reason for these behaviors is not clear, it may be an indicator of network policy in terms of how the different networks allocate bandwidth on the downstream and upstream at different times of the day. This will be further investigated in future.

Presently, the work is still ongoing and we are still generating more data for months and the data collection centers and experiment locations will be extended to other places/campuses within the university campus. Similarly, different mobile devices other than laptop computers will be used to see how those factors will affect the mobile broadband performance. However, we



recommend that a software be incorporated to capture large data during experiment from the simulation tool (Networx) on different systems which will be send to the central database for further processing.

#### References

Alcatel-Lucent (2012). Is Symmetrical Bandwidth a Myth or a Must? Retreived from <u>http://www2.alcatellucent.com/techzine/is-symmetrical-bandwidth-a-myth-ora-must/</u> on 18 April 2015.

Broadband Commission. (2013).The State of Broadband 2013:

Universalizing Broadband [White paper]. Retrieved from www.broadbandcommission.org/documents/bb-annualreport2013.pdf; On 16 January 13:30:15 GMT.

Dazhi, M.N.(2013). Understanding the relationship between customer and mobile broadband [Whitepaper]. Retrieved from http://www.cto.int/media/events/pstev/2013/CTOForum/; on 15 February 12:30:15 GMT.

Bergen, M. (2009). Mobile Broadband: including Imax and LTE.

Springer Publishing Company, Germany.

NCC Operator data (2014).Retrieved from http://www.ncc.gov.ng/index.php/index.php?option=com \_\_content&view=article&id=70&Itemid=67; On 5 March 13:20:20 GMT.

Nigeria's Presidential Committee on Broadband. (2013). *Nigerian National Broadband Plan* [White paper]. Retrieved from

http://www.researchictafrica.net/countries/nigeria/ Nigeria\_National\_Broadband\_Plan\_2013-2018.pdf; On 11 January 13:30:15 GMT.

Oborkhale, L., Salatian, A., and Onoh, G. (2012). Ranking the Quality Of Service of Wireless Celluar Network for Different Period of a Day.International Journal of Energy, Information and Communications, 3(1), 55-66.

Siwakoti, R.Y.(2014). Measuring Performance of Mobile Broadband Networks under Mobility. MSc Thesis, University of Oslo, Oslo, Norway. Retrieved from http://www.academia.edu/8055184/Measuring\_Performa nce\_of\_Mobile\_Broadband\_under\_Mobility; On On 15 January 12:40:15 GMT.

Temidayo, O.J., Stephen, U.N., Patric, E.I and Mathurine, G.(2014). Mobile Broadband Speed and Data Volume Analysis. Impact: International Journal of Research in Engineering & Technology, 2(5), 187-198.



### **EVALUATION OF FULL TEXT SEARCH RETRIEVAL SYSTEM**

#### \* K.D. Aruleba<sup>1</sup>, R.D. Aremu<sup>2</sup>, P.K. Oriogun<sup>3</sup>, K.K. Agbele<sup>4</sup>, A.O. Agho<sup>5</sup>

<sup>1,3,4,5</sup>Elizade University, Department of Mathematics and Computer Science, Ilara-Mokin, Ondo State, Nigeria

<sup>2</sup>University of Ilorin, Faculty of Communication and Information Science, Department of Computer Science, Ilorin, Kwara State, Nigeria

<sup>1</sup>kehinde.aruleba@elizadeuniversity.edu.ng, <sup>2</sup>draremu2006@gmail.com, <sup>3</sup>peter.oriogun@elizadeuniversity.edu.ng, <sup>4</sup>kehinde.agbele@elizadeuniversity.edu.ng,

<sup>5</sup>adrian.agho@elizadeuniversity.edu.ng,

#### ABSTRACT

With a number of search engines on the web and each with different indexing and ranking methods and different coverage, finding the one that gives the best results for a query becomes a bit challenging. The main problem however, that existing Search engines have to deal with is how to avoid irrelevant information and to retrieve the relevant ones. This current work presents a new approach for retrieving relevant information on the Web, by adopting breadth-First search algorithm. The implementation result of the retrieval system was analysed using recall and precision model for three departments at Elizade University. By learning from users' behaviour, the approach can return very high quality search results, with a strongly reduced computing load.

Keywords: Full-Text Retrieval System, Evaluation Approaches, Ir, Search Engines, Elizade University

#### 1. INTRODUCTION

Search engines are designed to help users to quickly find useful information on the web (Takakuwa, 2000). With a number of search engines on the web and each with different indexing/ranking methods and different coverage, finding the one that gives the best results for a query becomes a bit challenging. Previous studies shows that the performance of search engines depends on the performance measures used and the application domains. The performance of search engines can be evaluated using various measures such as precision, coverage, response time, recall and interface (Dong and Su, 1997). In this paper, we focus on recall and precision of search engines. The quality of searching for the right information accurately would be the precision value of the search engine. For example if we have Precision = 6 / 10 it implies that out of the 10 retrieved documents only 6 are relevant. Recall is the ability of a retrieval system to obtain all or most of the specifically relevant documents in the collection. For example Recall = 6 / 20 because there are 6 specifically relevant document out of the 20 documents retrieved.



#### 2. LITERATURE REVIEW

Recent efforts to create digital libraries have grown exponentially. A survey of the literature (Kreitz, 1996; Kreitz and Orgden, 1990) on digital libraries and initiatives offers definitions of digital library and challenges as well. This article focuses on electronic library resources within three Departments at Elizade University.

The importance of Information Retrieval (IR) keeps growing as the amount of digital information keeps expanding at an ever-increasing rate. Stored documents, photographs and contents of books, and billions of Web pages are useful only if they can be found when needed. Web search engines are the most common way to find such information. They are attracting more than 170 billion queries each month (Bonfils and Yandex, 2013). The field of IR also covers supporting users in browsing or filtering document collections or further processing a set of retrieved documents. Given a set of documents, clustering is the task of coming up with a good grouping of the documents based on their contents. It is similar to arranging books on a bookshelf according to their topic. Given a set of topics, standing information needs, or other categories (such as suitability of texts for different age groups), classification is the task of deciding which category, if any, each of a set of documents belongs to.

IR systems must have at least three different processes which are, representing the content of documents, representing a user's information need and comparing the two representations (Hiemstra, 2001). IR process begins when a user inputs a query into the retrieval system. Queries are formal statements (in declarative a formal language) of information needs, for example search strings in web search engines. In information retrieval a query does not uniquely identify a single object in the collection. Instead, several objects may match the query, perhaps with different degrees of relevancy. An object is an entity that is represented by information in the system database. User queries are matched against the database information.

Full text retrieval systems (FTRS) have become a popular way of providing support for text databases. In a full-text search, a search engine examines all of the words in every stored document as it tries to match search criteria for example (text specified by a user). The main components of a typical search engine according to (Brin and Lawrence, 1998) are: Web Crawler, Indexing and Ranking. Web Crawler according to (Sherman, 2002) are programs which traverse through the Web searching for the relevant information using algorithms that narrow down the search by finding out the most closer and relevant information. Indexing collects, parses, and stores data to facilitate fast and accurate IR. The main purpose of storing an index is to optimize speed and performance in finding relevant documents for a search query. Ranking is the medium a search engine use to determine which pages are more important than the others, and present them to individual users in order of relevance. The most famous one is the Page Rank Algorithm published by Google founders (Pavalam et al., 2012)

#### 3. METHODOLOGY

There are various search methods to traverse (visit all the nodes) of a graph systematically. A couple of these methods give us some information about graph structure (e.g. connectedness). The key idea behind graph traversal is to mark each vertex when we first visit it and keep track of what we have not yet completely explored. We describe some of the mechanics of these traversal algorithms here. Depth-First Search (DFS) is an algorithm for traversing a finite graph. DFS visits the child nodes before visiting the sibling nodes; that is, it traverses the depth of any particular path before exploring its breadth. A stack is generally used when implementing the algorithm. Breadth-First Search (BFS) uses a queue data structure and it is level by level traversal. Breadth First Search expands nodes in order of their distance from the root. It is a path finding algorithm that is capable of always finding a unique solution, if one exists.

#### 3.1 IR Evaluation Approaches

According to (Agbele, 2014) retrieval effectiveness can be quantitatively measured in a number of ways using a well-known metrics in the IR



community to enhance retrieval effectiveness. The most frequently and important basic measures for IR evaluation are precision and recall which are both used in this present study.

#### 3.1.1. Precision

After a search, the user is sometimes able to retrieve relevant information and sometimes able to retrieve irrelevant information. The quality of searching the right information accurately would be the precision value of the search engine.

$$Precision = \frac{relevant items retrieved}{total retrieved items} = \frac{RIR}{TRI}$$

#### 3.1.2. Recall

Recall is the ability of a retrieval system to obtain all or most of the relevant documents in the collection. Also, recall is the fraction of relevant items that are retrieved to relevant items in the database or the probability given that an item is relevant to the retrieved. For example for text search on a set of documents recall is the number of correct results divided by the number of results that should have been returned.

$$\operatorname{Recall} = \frac{\operatorname{relevant items retrieved}}{\operatorname{relevant items}} = \frac{RIR}{KI}$$

#### 3.1.3. 11- Point Average Precision

11-point average precision is a measure for representing performance with a single value. In 11-point average precision, we are looking at 11 recall levels (0.0, 0.1, 0.2,... 1.0) and finding the precision at each point. We average these scores across all of the different issued queries from the participants or information needs to validate the retrieval effectiveness of developed system.

#### 3.2. PROPOSED SYSTEM

Figure 1 depicted the architectural design for a Full Text Retrieval System (FTRS) proposed (Aruleba, 2015). This architectural design was implemented using Breadth First Search. The proposed system makes use of a Crawler to gather information from every document on the website and store this information in the index. The index is a structured system of storing the unstructured data returned by the Crawler.



Figure 1: Proposed Architecture for FTRS

#### 4. SYSTEM PERFORMANCE EVALUATION

This section presents the evaluation of the adopted document retrieval search algorithm. The aim of the section is to measure the effectiveness of the retrieval system. In order to test the effectiveness of the full text search system, three departments from Elizade University, that is Departments of and Computer Science. Mathematics Civil Engineering and English were considered. Recall and precision were the two performance parameters used for evaluating the search system. Department of Mathematics & Computer Science, Civil Engineering were chosen to test the computer skills of the users in query formulation and the use of keywords, also English was selected to see how users can construct sentence using keywords and to see how relevant the results of the retrieval system is. The various departments and the total number of participants used in the evaluation is as shown in Table 1.

Table1: De	partments	and Partici	pants
------------	-----------	-------------	-------

Department	Participants
Mathematics & Computer	15
Civil Engineering	15
English	15





Total

45

The evaluation had no fixed queries. The Users were asked to perform their daily book searches as usual, based on their daily information needs without any change. The only requirement was that they needed to focus mainly on using search terms related to their departments. The system was developed and implemented with PHP; before the system can be used some requirement (such as software and hardware requirement) must be met. Figure 2 depicts the sample snapshot search system.



Figure 2: Sample Search Screen

#### 5.1 RESULTS

During the evaluation, users were asked to rate their overall satisfaction with the search engine based on the retrieved results in facilitating their academic work. The results shows that the users in Mathematics & Computer science are more satisfied with the performance of search engines, while the opinions of the users in Civil Engineering and English appeared to be similar to one another as shown in Figure 3.

#### 5.2 DISCUSSION

According to Figures 3, it was observed that there is usually a trade-off between recall and precision i.e. at a high recall value, more documents containing a lot of junks was retrieved by the system and hereby reducing precision while at a high precision value, less but the most relevant documents were retrieved and thereby providing a low recall value. Another observation is that the system seems to perform well on one query than it does to another. This has to do with the query formulation skills of the individual user and how much knowledge a user had about the system content as illustrated in Figure 4.







Figure 3: Average 11-point r-p curve across 10 queries using Department of Mathematics and Computer Science, Civil Engineering, English.



Figure 4: Comparison of 11-point average of mathematics & computer science, civil engineering and English.

(NB: The curve closest to the upper right-hand corner of the graph indicates the best performance)

#### 6. CONCLUSION & FUTURE WORK

In conclusion, the analysis of results of the implemented information retrieval system shows that the users of the system find it very effective to

use. The implemented information retrieval system enables users to have access to latest learning facilities such as, articles, journals, textbooks, thesis, projects, newspapers, etc. without going



through the rigorous steps and routine in the conventional institution libraries. The field of information retrieval is a very interesting research area where improvements can always be made no matter how sophisticated your retrieval application looks. For the future, it remains to be seen whether novel algorithms which may use hybrid techniques and may outperform BFS and DFS individually.

#### 7. REFERENCES

- Takkakaw T., (2000) "Search engines worldwide" http://www.twics.com/takakuwa/search/search. html (Accessed 15<sup>th</sup> May 2015)
- Dong X., Su L. T. (1997), "Search engines on the World Wide Web and information retrieval from the Internet: A review and evaluation." Online and CDROM Review 21: pp. 67-81
- Kraits, P.A. (1996) "Role of the Library in the Scholar's Electronic Research Environment, Digital Libraries and Information Services for the 21st Century". Seoul, Korea: ICPR.
- Kreitz, P.A., and Ogden, A. (1990). "Job responsibilities and job satisfaction at the University of California Libraries, College & Research Libraries", 51, 4, 297.
- Bonfils M, Yandex (2013), "Just passed Bing to become 4th largest global search engine" 2013.

#### URL

http://searchenginewatch.com/article/2242374/ (Accessed 20<sup>th</sup> May 2015)

- Hiemstra D., (2001). "Using language models for information retrieval." Ph.D. Thesis, Centre for Telematics and Information Technology, 2001.
- Page L and Brin S. (1998). "The Anatomy of a Large-Scale Hypertextual Web Search Engine", Proceedings of the Seventh International Conference on World Wide Web 7, Brisbane, Australia, 30(1-7): 107 – 117, 1998.
- Sherman C. (2002). "Anatomy of a Search Engine: Inside Google." http://www.searchenginewatch. com/searchday/article.php/2161091. (Accessed 8<sup>th</sup> July 2015)
- Pavalam M., Jawahar M., Felix K., Akorli, S V Kashmir R. (2012). "Web Crawler in Mobile Systems" International Conference on Machine Learning (ICMLC 2011), Vol. 2, No. 4, pp 531- 534.
- Agbele K.K., (2014). "Context-Awareness for Adaptive Information Retrieval Systems" Unpublished PhD thesis University of Western Cape, South Africa, 2014.
- Aruleba K.D, (2015). "A full text retrieval system in a digital library environment" Unpublished MSc dissertation, University of Ilorin, 2015.



### AN ENHANCED HYBRID ALGORITHM FOR CELLULAR DATA OFFLOADING

#### A. A. Bichi<sup>1</sup>, and S. M. Zahid<sup>2</sup>

<sup>1</sup>Department of Computer Science Northwest University Kano <sup>2</sup>Faculty of Computing Universiti Teknologi Malaysia

<sup>1</sup>abdulkadirkut@yahoo.com, <sup>2</sup>soperi@utm.my

#### ABSTRACT

The cellular networks are increasingly facing the challenge of data explosion due to high demand of data by users. Several forecasts and analysis indicates that, in the near future the technology cannot cater for the demand of users. Many attempts to upgrade the technology resulted into another inefficient or expensive solution. Cellular data offloading to Wi-Fi is the most promising solution that solves the problem at an affordable cost. This paper proposed a technique to offload more traffics with reduce packet loss and latency respectively in a cellular network. In the proposed algorithm, whenever data request is received, the algorithm checks the availability and signal strengths of both the Wi-Fi and cellular and then classified the applications into real-time and non-real-time. The strategy always connects to Wi-Fi network if available. But when the Wi-Fi signal falls below the minimum threshold during the transfer, the strategy will complete the transfer using cellular immediately for real-time applications and wait till the end of the delay deadline in case of non-real-time applications before it switch to cellular. The performance of the proposed cellular data offloading algorithm was evaluated using a simulator. In the simulation the following input parameters were used: the value of threshold signal, network availability and application type. The followings metrics were determined for each simulation; rate of packet loss, transfer latency and fraction of data offloaded to Wi-Fi. To test the completion time of real time application, a small amount of data was used on VoIP application which can only tolerates delay in order of milliseconds. The simulation result shows that the enhanced algorithm achieved 59% of offloading efficiency similarly to the Wiffler algorithm but with 28% reduction in packet loss and 30% less completion time for real-time traffics compares to existing algorithm. It was concluded that the proposed approach may be adapted by the designer od cellular networks.

Keywords: on-the-spot strategy, delay-tolerance strategy, Wi-Fi, real-time applications

#### 1. INTRODUCTION

With the advancement in Information Technology (IT); users now demand connectivity anywhere and anytime. This evolution is driven by popularity of video, social media and Internet gaming across mobile devices such as smartphone and tablets

thus lead to explosion of cellular networks data traffic.



It was predicted that by 2014 an average mobile broadband user will consume 7 GB of traffic per

month (Wipro, 2012), and the number of mobile users was predicted to reach 5.2 billion by the year 2017 (Ericson, 2013). This will consequently, raised the amount of data traffic generated by mobile devices alone to 11.2 Exabyte per month by 2017, a 13-fold increase over 2012 (Cisco, 2013).

To address the problems and the issues associated with mobile traffic overload many solutions were proposed, the main proposed solutions to tackle the problem include:

a. **Upgrade:** upgrading the existing cellular technology with the higher technologies such as: long term evolution (LTE) by third generation partnership project (3GPP) and High Speed Packet Access (HSPA) that will provide more bandwidth (Siris and Kalyvas, 2012). But this cannot effectively solve the problem, since both the technologies has the same frequency band as the original cellular which is very scarce (Chandrasekhar et al, 2008) and expensive (Berg and Katsigiannis, 2012).

b. Adding Spectrum: additional spectrum will definitely enable operators to install more antennas and other equipment at a given cell site, which will eventually increase capacity without affecting the quality of the existing network. But this solution is not cost effective with the spectrum being a limited resource and governments and regulatory agencies value it a lot, therefore in many countries buying additional spectrum can never be an option (Berg and Katsigiannis, 2012).

c. **Cellular Data Offloading:** This is the use of other (mostly complementary) network technologies for delivering data originally plan to be transmitted via cellular network. The complementary networks work with the macro-cellular network as an adjunct or overlay network (Han et al, 2012).

In a cellular data offloading, other networks (e.g. Wi-Fi and Femtocell) are used to transfer data which originally intended for cellular network thus minimize the traffic in the cellular network. The other networks (e.g. Wi-Fi) usually offer less expensive method for delivering data services. Operators can utilize the complementary network to reduce traffic congestion on the main network and to cost-effectively increase network capacity at specific locations.

The main network technologies for cellular data offloading are Wi-Fi and Femtocell (Bichi, 2013). Wi-Fi is the general name given to any wireless local area network (WLAN) products which is based on IEEE 802.11 standards (Gass and Diot, 2011) while Femtocell refers to a small cellular base station, use mostly in a home or small business which connects to the service provider's network via broadband (Elleithy and Rao, 2011).

Using Wi-Fi Access Point (AP) is the most promising solution for cellular data offloading (Chandrasekhar et al, 2008) and is the only solution that can cater for the future continuous explosion of data traffic (Bulut and Szymanski, 2012). The solution contributes in higher bandwidth availability and reduced data services cost. And there is less interference since Wi-Fi operates in different spectrum with the cellular base station (Gass and Diot, 2011).

Other economic factors that contribute to the success of Wi-Fi APs in cellular offloading include the vast availability of APs, this simplifies the complexity and cost of both managing and deploying a new Wi-Fi AP; therefore reduce the overall cost of the network services (Gupta and Rohil, 2012). Since Wi-Fi operates in unlicensed spectrum, regulatory approval is not required for individual deployments (Bulut and Szymanski, 2012). Moreover, user's devices are already capable of communicating using Wi-Fi. The Figure 1 shows the typical offloading scenario using Wi-Fi:



## 12<sup>th</sup>

### **International Conference**



Figure 1: Cellular Data Offload Using Wi-Fi

#### 2. LITERATURE REVIEW

#### 2.1 Background

The two main strategies for cellular data traffic offloading to Wi-Fi Access Point (AP) are: On-thespot offloading (Hagos, 2012) and Delay tolerance (Lee et al, 2010).

#### a. On-the-Spot Offloading

In the strategy of the On-the-spot, mobile device switch the data transfer from cellular network to Wi-Fi AP immediately whenever is available; the transfer is complete using cellular network whenever user moves out of Wi-Fi coverage (Hagos, 2012). Majority of devices with Wi-Fi facility are configured by default to give higher priority to Wi-Fi over the cellular interface for data transmissions. Therefore, the strategy is the most commonly used whenever the Wi-Fi network is available. Offloading starts automatically whenever node move into Wi-Fi coverage area and continue the transfer using cellular networks when the Wi-Fi strength falls below the minimum threshold value.

#### b. Delay tolerance

The delay tolerance which uses the concept similar to that of delay tolerance networks (DTN) assigns a deadline to a data to be transfer, the data request will wait whenever a user moves out of Wi-Fi coverage and continue the transfer when user enters the Wi-Fi coverage area again until finally a maximum delay threshold is reached before it complete the transfer via cellular network (Lee et al, 2010). The delay can help to exploit Wi-Fi network more thus higher offloading percentage could be realized (Han et al, 2012).

The delay-tolerance strategy has higher offloading percentage compared to on-the-spot algorithm (Hagos, 2012), but the strategy treats application uniformly therefore result in high rate of data loss and lack of quality assurance for real-time applications such as video streaming and VoIP.

Real time data communication such as the one mention above requires low latency; therefore not suitable for the existing delay algorithm and the gain is insignificant with the on-the-spot strategy. In order to exploit Wi-Fi more while maintaining the quality of real time application. This research proposed a hybrid strategy for the cellular Data offloading. The proposed algorithm will apply delay tactics on non-real-time applications while for realtime the algorithm will offloads data on-the-spot. The algorithm will work by first classifying applications into, real-time and non-real-time.

#### 2.2 RELATED WORKS

Lee et al, 2010 and (Gupta and Rohil, 2012) compared the performance of both on-the-spot and delay-tolerance strategy by experimental and simulation techniques and the work showed that significant reduction of cellular congestion can be achieved by applying cellular data offloading. The work further confirms that the delay strategy has higher offloading percentage as compared to the strategy of on-the-spot.

In Siris and Kalyvas, 2012, a new delay-algorithm was proposed; the algorithm can predicts both the duration in Wi-Fi coverage and the time to complete the transfer of data. The simulation result of the proposed algorithm shows 60% cellular usage reduction.

Balasubramanian *et al.* (2010), introduces a method called Wiffler, which is an enhancement over the traditional delay tolerance option; Wiffler combined the strategies of delay tolerance and fast switching to cellular networks. The method works by predicting future Wi-Fi throughput, given any amount of data to transfer the algorithm predicts if



there is possibility of Wi-Fi occurrence within the limit of the delay tolerance; otherwise proceed with the transfer using cellular network.

The prediction of meeting AP is based on the historical probability of past occurrence of AP. The probability assumes that, if the device meets APs frequently, then it will likely meets the next AP within a short time interval. Similarly, if time interval between meetings is large, then the next AP will not be met at short interval. An analysis over large set of data shows that AP meetings obey the pattern.

However, the concept of offloading was proposed using social network (Han et al, 2012; Li et al, 2011). The research observes that MoSoNets connect large domain of people together; ranging from friends, work colleagues and family members who otherwise are hidden from these online services. It can also provide a platform to signal face-to-face interactions among nearby people who probably should know each other.

MoSoNets can be viewed as a marriage of traditional social networks with emerging opportunistic networks. They can exploit both types of communication to facilitate information dissemination in MoSoNets. Users can actively forward (push) information whenever they want, while mobile users that are in contact can also pull information from each other locally. Analysis shows that the proposed Heuristic algorithm of (Han et al, 2012) can offload cellular data traffic by up to 73.66 percent.

Kashihara *et al.* (2012), proposed the possibility of data offloading to public transport vehicles for sending/receiving large data objects that can accept some delay. In the proposed approach, they employ public transport vehicles, such as a bus or a train, as a communication medium between users and the Internet. A user sends a large data object to a public transport vehicle by using high-speed short-range wireless communication and the public transport vehicle forwards the object to the Interne. As for receiving a large data object, a

public transport vehicle receives a user's data from the Internet, and delivers it to the location (e.g., a bus stop or station) where the user desires to receive it. After that, the user receives the large data object using high-speed short-range wireless communication. Since such public transport vehicles are closely related to users' daily life (e.g. commute), users have frequent access to the reachable locations such as bus stops or train stations. A benefit of this approach is that the delivery time of the user's data can be estimated, because the public transport vehicles basically follow timetables. The paper proposes the basic design and performance for data offloading to public transport vehicles, and then considers technical issues for achieving it.

#### 3. PROPOSED APPROACH

In the proposed approach, the cellular data offloading algorithm work as follows: For a user to connect to cellular network, the algorithm checks Wi-Fi availability and the signal strength, if the Wi-Fi is available and the signal is above minimum threshold value the algorithm offload the data traffics to Wi-Fi else continue the transfer using cellular network. And for the user initially on Wi-Fi the algorithm will maintain connection unless user moves out of Wi-Fi or the signal fall below minimum threshold value then it will check whether the application is real-time or non-real-time. For realtime it immediately connect back to cellular while for non-real-time it will wait for tolerable delay period. If user enters Wi-Fi before the delay threshold then continue transfer using Wi-Fi else go back to cellular.

The on-the-spot strategy offloads less percentage of the traffic compares to delay-tolerance strategy. It is clear that, the delay-tolerance exploits Wi-Fi more, by waiting for some period before switching. The strategy will be good for non-real-time applications but affects the performance of realtime services. Some of the bulk data for example, personal videos, some scientific experiments data, and some uplink data created by sensors, that do not require real-time transmission can tolerate certain delay. But on the other hand, services like: video streaming, VoIP, mobile TV, video



conferencing, and tele-medicine, which require high network bandwidth and fast response time will be better using on-the-spot.

In the existing algorithm (Balasubramanian et al, 2010) the applications are only allow using cellular network after losing some packets therefore hurt the performance of the real-time services. And the algorithms only used on-the-spot in the upstream direction when forwarding the packets. There is another delay of counting the prediction before deciding the offload, these affects the performance of real-time applications.

To reduce the latency of packets transfer in realtime applications; in the proposed algorithm (Figure 2), the traffics from the real-time applications such as video streaming or VoIP will be transfer directly using cellular if Wi-Fi signal is not available or RSSI from all the AP's are below the minimum threshold for real-time communication. The delay of retransmission attempt over Wi-Fi and the delay of medium access have to be eliminated for real-time applications.

#### Variable:

T<sub>d</sub>: total delay M<sub>D</sub>: Maximum delay W<sub>s</sub>: Wi-Fi signal S<sub>th:</sub> Minimum Signal threshold **Algorithm:** 

**For** all AP's in the vicinity **Do** //collect RSSI information

Scan for RSSI

n= number RSSI received

For i=2 To n Do

If RSSI1 > RSSIi set RSSI = Ws

Increment i

If (W<sub>s</sub> >= Sth) then

Transfer data using Wi-Fi

Else

// check application

If application port address is known for

real-time application then

Transfer data using cellular

Else set T<sub>d</sub> =0

```
Do
If (W<sub>s</sub> >= Sth)
```

Transfer remaining data using

End If Advance T<sub>d</sub> While (T<sub>d</sub> < M<sub>D</sub>)

Exit loop

Transfer remaining data using

cellular

Wi-Fi

End if

End if

Fig 2: Proposed Algorithm



#### 4. RESULTS AND DISCUSSION

The performance of the proposed cellular data offloading algorithm was evaluated using a simulator. Variable parameters were used for each simulation. The inputs parameters include: the value of threshold signal, network availability and application type. The followings metrics were determined for each simulation; rate of packet loss, transfer latency and fraction of data offloaded to Wi-Fi.

To test the completion time of real time application, a small amount of data was used on VoIP application which can only tolerates delay in order of milliseconds. The Figure 3 compared the performance of the proposed cellular data



algorithm against the algorithms of delay-tolerance and on-the-spot strategies, and the result was positive on VoIP application which is an example real-time services.

Figure 3: Graph of Delay-deadline (vs.) Completion Time

As shown in Figure 4, the proposed algorithm has lower transfer latency as compare to delayalgorithm. The result also indicates that the proposed algorithm works like on-the-spot algorithms on real-time applications by eliminating the delay of packets retransmission and by avoiding the delay of transmission in both upstream and downstream as opposed to Wiffler algorithm (Balasubramanian et al, 2010).

The packets loss during offloading is directly promotional to the delay time. The graph in Figure 4 shows the percentage of packets lost in the proposed algorithm and two other existing algorithms for benchmarking when a video streaming of 5MB was applied.



Figure 4: Graph of Delay-deadline (vs.) Packet Loss

The simulation result (Figure 4) shows less packet loss in the proposed strategy even when high delay deadline is used. At the delay of 300s the proposed strategy lost only 30% of packets which is close to 20% of the on-the-spot and much lower than 58% of the delay-strategy.

The proposed strategy as opposed to other on-thespot algorithms can offload high percentage of the traffic from cellular network to Wi-Fi Access Point like the delay-strategy when a non-real-time applications that can naturally tolerate substantial delay is applied on the algorithm as in Figure 5.



## 12<sup>th</sup>

### International Conference



Figure 5: Graph of Delay-deadline (vs.) percentage of data offloads to Wi-Fi

From figure 5 the delay-strategy and the proposed strategy offloaded almost the same percentage of respective 60% and 59% at 50s delay deadline which is much higher compare to the percentage offloaded by on-the-spot strategy at the same point which is 29%. The finding shows higher offloading efficiency of new strategy compares to the on-the-spot when dealing with non-real-time applications that can tolerate high delay.

#### 5. CONLUSION

The proposed algorithm reduces the transfer latency and percentage of packets losses in the transfer of real-time applications. The enhanced algorithm have the advantageous features of both the on-the-spot and delay-tolerant strategies. The research conducted a simulation of all the three strategies using the same parameters, as described in the preceding figures. The proposed enhanced algorithm reduces the completion time of data transfer and percentage of packet losses in realtime applications. The enhanced algorithm as opposed to direct on-the-spot strategy like Oracle algorithm has lower transfer latency in real-time applications while achieving substantial percentage offloading. The simulation result shows 59% of offloading with 28% reduction of packet losses compared to existing delay-algorithm.

#### 7. REFERENCES

- Andrews J. G., Claussen H., Dohler M., Rangan S., Reed M. C. (2012). Femtocells: Past, Present, and Future, IEEE Journal on Selected Areas in Communications, Vol. 30, No. 3. IEEE.
- Balasubramanian, A, Mahajan, R. and Venkataramani, A (2010). Augmenting Mobile 3G Using Wi-Fi. The 8th Annual International Conference on Mobile Systems. San Francisco, California, USA: MobiSys. Pages 209–222
- Berg K. and M. Katsigiannis(2012). Optimal costbased strategies in mobile Network Offloading. 7th International ICST Conference on Cognitive Radio Oriented Wireless Network and Communication. pp. 95-100
- Bichi A. A. (2012). Hybrid Strategy for Cellular Data Offloading to Wi-Fi Access Point Based on the Type of Application. Master's thesis submitted to faculty of Computing, Universiti Teknologi Malaysia.
- Bulut E. and Szymanski B. K. (2012). Wi-Fi Access Point Deployment for Efficient Mobile Data Offloading. In a proceeding ACM workshop PINGEN at mobicom Istanbul Turkey.
- Calin D., Claussen H., and Uzunalioglu H. (2010). On femto deployment architectures and macrocell offloading benefits in joint macrofemto deployments. Communications Magazine, vol. 48, no. 1, pp. 26–32.
- Chandrasekhar V., Andrews J. G., and Gatherer A. (2008). Femtocell networks: A survey. Communications Magazine, vol. 46, no. 9, pp. 59-67. IEEE.
- Choi Y., H. Wook J, Park J., Kim H. and Silvester J. A. (2011). A 3W Network Strategy for Mobile Data Traffic Offloading. Communications Magazine pp. 118-123.
- Huang C., Lan K., and Tsai C. (2008). A Survey of Opportunistic Networks. 22<sup>nd</sup> International Conference on Advanced Information Networking and Applications. National Cheng Kung University, Tainan, Taiwan.
- Chuang Y. and Lin K. C. (2012). Cellular Traffic Offloading through Community-based Opportunistic Dissemination. Wireless communication and networking conference, pp 3220-3225.



- Cisco (2013). Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012 – 2017. Cisco
- Cuervo E, Balasubramanian A, Cho D, Wolman A, Saroiu S, Chandra R, Bahl P (2010). Making smartphones last longer with code offload. International conference on mobile systems, applications, and services, pp. 49–62
- Dimatteo S., Hui P., Han B. and Li V. O. K. (2011). Cellular Traffic Offloading through Wi-Fi Networks, in Proceeding of IEEE MASS
- Elleithy K. and V. Rao (2011) FemtoCells: Current Status and Future Directions. International Journal of Next-Generation Networks Vol.3, No.1, pp. 1-9. IJNGN
- Ericsson (2013). "Traffic and Market Data Report," Available:

ttp://www.ericsson.com/news/1561267

- Gupta V. and Rohil M. K. (2012). Mobile Data Offloading: Benefits, Issues, and Technological Solutions. Advances in Computer Science, Eng. & Appl., Vol167, pp. 73–80. Springer-Verlag Berlin Heidelberg.
- Hagos D. H., (2012). Performance of Wi-Fi Offload in LTE Networks. Master thesis, Department of Computer Science, Electrical and Space Engineering Lulea University of Technology Sweden.
- Han B., Hui P., Kumar V., Marathe M. V., Pei G., and Srinivasan A. (2010) "Cellular Traffic Offloading through Opportunistic Communications: A Case Study," in ACM CHANTS

- Kashihara S., Sanadidi M. Y. and Gerla M. (2012). Mobile, Personal Data Offloading to Public Transport Vehicles. International Conference on Mobile Computing and Ubiquitous Networking, pp142-147. 2012.
- Lee K., Rhee I., Lee J., Chong S., and Yi Y., Song C. (2010). Mobile Data Offloading: How Much Can Wi-Fi Deliver ACM p425-426. New York
- Li Y., G. Su, P. Hui, D. Jin, L. Su and L. Zeng (2011). Multiple Mobile Data Offloading Through Delay Tolerant Networks. CHANTS '11 Proceedings of the 6th ACM workshop on Challenged networks pp. 43-45.
- Pan H, Kumar, A, Madhav V. Marathe, J. S, Aravind S (2012). Mobile Data Offloading through Opportunistic Communications and Social Participation, MD 20742, USA.
- Pasandideh M. R. (2011). Automatic Planning of 3G UMTS All-IP Release 4 Networks with Realistic Traffic. Master's thesis. Carleton University Ottawa, Ontario, Canada
- Siris V. A. and Kalyvas D. (2012). Enhancing Mobile Data Offloading with Mobility Prediction and Prefetching. . MobiArch'12, Istanbul, Turkey.
- Wipro Council for Industrial Research (2012). Data Offloads Approaches for Mobile Operators. Wipro, Bangalore
- Zhuo X., Gao W., Cao G., and Dai Y. (2011). Win-Coupon: An Incentive Framework for 3G Traffic Offloading, Proceeding of IEEE ICNP.



### **MULTI-FACTOR AUTHENTICATION FOR SOCIAL MEDIA**

#### <sup>\*</sup>O. O. Ayannuga (PhD)<sup>1</sup>, O.N. Lawal, O. Oladipo<sup>3</sup>, and W.A. Salau<sup>4</sup>

<sup>1,2,3,4</sup>Yaba College of Technology, Yaba, Lagos.

<sup>1</sup>ayannugalanre@yahoo.com, <sup>2</sup>kingwales@gmail.com, <sup>3</sup> segun.oladipo@yabatech.edu.ng, <sup>4</sup>ayotunde.salau@yabatech.edu.ng

#### ABSTRACT

Presently, user authentication on social media depends on username and password, which is text-based. The accounts of many users have been hacked because text-based password can be guessed or hacked. However, nobody can hack a user's biological feature. Multi-factor authentication (MFA) is a method of computer access control, which a user can pass by successfully presenting several, separate authentication stages. This paper presents an application model with the conventional username/password and voice biometrics authentication system to further secure social media access. The developed system has three modules: the voice enrollment, profile acquisition, and multi-factor matching stage. The application model was developed and run using visual studio 2010 development environment and the .net system's speech library.

Keywords: Authentication, Multi-Factor, Security, Social-Media

#### 1. INTRODUCTION

The last decade has witnessed a tremendous evolution and expansion of the Web, which has led to the introduction of additional technologies such as social networks, blogs, content sharing and video sharing. These technologies allow dynamic user-generated contents, the publishing of consumer feedbacks forming global online communities at real time (Saw et al., 2012). Social networking is becoming the preferred (by endusers) way to manage personal data. It is an area where people take an active interest in how their personal information is managed and displayed rather than being passive account. With social networks representing the world's largest body of personal data collection, it is becoming a trend for users to give their account passwords to others to help in managing their various profiles, and this leads to various users falling into the hands of social networks miscreants, and in most cases, it was discovered that personal data was being compromised.

There is the need to see tools for more finegrained delegation of authorization to help solve the problem of identity theft and data compromise by unauthorized users.

Social networking sites are essentially websites designed for human interaction. They enable users to meet others; keep in touch with them; and share experiences, feelings, and opinions (Sancho, 2009). They are all built on a similar foundation 'Trust' - the user builds a network of contacts bound by an element of trust. The user then creates content for his/her friends and, in turn, accesses the content they have created. This content can include diverse things as holiday pictures, interesting links, latest news, opinions, comments, and mood updates. There are potentials for mischief and malicious activities when one or more of your contacts or social aggregators breaks your trust. Some of the



things that can go wrong when such happens may include, but not limited to, the following:

- Your account might be compromised and somebody else is using it.
- Insufficient use of privacy controls caused you to share data with people you never intended.
- Your account might be used by someone else to perpetrate fraud.
- Loss of vital information to cyber miscreants.

Despite well-known security issues, passwords are still the most popular method of end-user authentication. Guessing and offline dictionary attacks on user-generated passwords are often possible due to their limited entropy (Emiliano, et al., 2014). Over the years, a number of policies have been introduced in order to increase security in the use of social media such as enforcing a minimum characters, inclusion number of of nonalphanumeric symbols, or frequent password expiration. But in reality, this often creates an undesirable conflict between security and usability as highlighted in the context of password selection (Egelman et al, 2013), management and composition (Zezschwitz, et al, 2013) and in turn drives users to find the easiest password that is policy-compliant (Adams, 1999).

Multi-factor authentication has emerged as an alternative way to improve security by requiring the user to provide more than one authentication factor, as opposed to only a password (Ayannuga and Lawal, 2012; Emiliano et al., 2014).

This study presents a Multi-factor application model for social media, using the Username/password and Voice biometric authentication systems.

#### 2. LITERATURE REVIEW

Authenticating is the process of verifying the identity of a certain person. This has been used for many different purposes throughout history e.g., scouts and other messengers needed to authenticate themselves to city guards and sentries in the past before they were given access to different areas (Richard, 2001). In the modern electronically wired information society, authentication has gained an even more important role, where user authentication is used to grant access control for many computer systems (UNIX, Linux and Windows). User authentication on these

operating systems is usually based on password authentication.

Xiaoyun, et al., (2005) divide authentication methods into three categories: token-based authentication; biometric-based authentication; and knowledge-based authentication. Ayannuga (2012) identified a few threats to authentication, which include: brute force, shoulder surfing, hashing, spyware, social engineering, phishing, spoofing, guessing, and dictionary attack. However, new threats are being discovered regularly. These authentication threats have necessitated the need for multifactor authentication, which implements a combination of at least two categories of the authentication methods.

Multi factor authentication is an extra layer of authentication added to the conventional single factor authentication to an account login, which requires users to have additional information before access to a system is granted (Gonzalez, 1998). The traditional method of authentication requires the user to enter only a username and password before being granted access to a closed software, whereas multi-factor authentication requires the user to have additional information available only to the user before access to the system is granted (Al-Fairuz, 2011). The information required to authenticate users includes one of the following methods listed below. Knowledge based i.e. a piece of information the user knows such as a Personal Identification Number (PIN) or password. The Biometric factor; the physiological or biological components of the user such as fingerprint, face image and voice pattern. Things within the possession of the user, such as a hardware or software token could also be used but not as safe as biometric factor.

Ayannuga and Lawal (2012) and Emiliano et al. (2014) worked on the usability study of multi-factor authentication system with emphasis on two-factor systems. These studies inferred that two-factor technologies are overall perceived as usable, regardless of motivation and/or context of use. The studies also presented an exploratory factor analysis, highlighting that three metrics –ease-ofuse, required cognitive efforts, and trustworthiness – are enough to capture key factors affecting twofactor authentication usability.

Marise and Micheal (2010) also worked on the adoption of Single Sign-On (SSO) and Multifactor Authentication in Organizations (MFA). The



findings in the study suggested a number of technology, organization and environment factors both positively and negatively that affect organizational adoption of SSO and MFA. It clearly stated a number of key benefits gained from adopting SSO and MFA, such as increased corporate security and reduced organizational costs of managing access control. There are also a number of key challenges to be overcome by organizations adopting SSO and MFA. These include the ability to accommodate the complexity of multiple heterogeneous systems and to be resilient to new information security threats thereby allowing a SSO and MFA solution to deliver improved and secure access control to information systems both within and across organizations.

Ayannuga and Lawal (2012) observed that almost all the existing usable authentication schemes have one or more noteworthy security shortcomings for which shoulder surfing is a major player. The schemes do not provide mechanisms that will, to a large extent, protect user against shoulder surfing. Some of these schemes do not suggest good usability as they are difficult for users to memorize and adapt to, especially when a user needs to use different passwords for different online transactions; since it is unsafe for a user to use the same password for all online transactions.

An important goal of all usable authentication schemes is to ensure a usable yet secure system for user authentication. Sometimes, this was done too faithfully, compromising parts of the system security. Other times, the schemes failed to consider certain important aspects of usability such as: ease of use, ease of adapting, ease of memorizing, and several others. To ensure that the goal of usable authentication schemes are achieved, developers should put the user into consideration, while ensuring high level security of user's authentication details (Ayannuga and Lawal, 2012). The use of multifactor authentication, which combines the regular username/password with voice biometric authentication system is an attempt to achieve this laudable goal. This is what the paper is set out to achieve.

#### 3. RESEARCH METHODOLOGY

Conventional authentication system uses one of knowledge factor (password), possession factor such as identification cards or inheritance factor (Biometric factor) for access control system. This is defective since the circumvention of anyone of the factors used could breach the security system of the application. This calls for further research to ensure extra layer of security for authentication system. This study implements the conventional password factor and voice biometrics to authenticate social media user. The application used a multitier development paradigm of programming and different .dll from the System.Speech library to simulate the authentication system.

#### 3.1 Objective of the Developed System

The objectives of the developed system include: representing voice pattern using discriminate features, storing extracted features into a relational database, matching voice pattern against a database of enrolled features, adding extra layer of password authentication to ensure that the password given matches with the identity provided by the matching system and ultimately ensure and enforce the social media user owns the identity he/she claims.

#### 3.2 System Design

The system was designed to proceed in two stages; the enrollment and the matching stages. The enrollment stage was modeled to enroll different users, acquire their voice pattern, and assign inputted profile and passwords to the acquired pattern. The enrolled data is permanently stored in the database. The matching stage requires that a user pronounces the same word acquired during the enrollment. The application makes effort to match the voice pattern against those initially stored in the database. The multifactor layer of the application then picks up the user's identity if found and prompt to enter password. If the password matches with the identity fetched by the biometric layer the application displays access granted and if it does not match, it displays invalid entry. The flowchart in figure 1 shows the processes involved in the enrollment stage while the flowchart in figure 2 describes the processes involved in the matching stage. The matching stage is presented as a web module which authenticates users before giving access to the social media.



#### 3.3 Database Relational Entity Diagram

Two tables were synchronized to design an effective multifactor authentication system. The tables are: the Userprofile (i.e. the profile of the social media users), Users table (i.e. all valid users table, which stores voice templates and usernames) and secQuestion Table (i.e. Table with security question answer and the affiliated username). The "Userprofile" typifies a valid user and the assigned password, while the users table has the voice biometric templates of valid users created during the enrollment process. The users table has an "Id" field (i.e. user's identification number) as the primary key to ensure that the table can accept incremental voice templates of a user described by same username at different times in case of wrong identification of users. Such users are made to re-enrolled using a 'my voice do not match link' after providing answer to the security question. The Users table keeps all the enrolled templates for all the username. Figure 3 shows the entity relationship diagram.



# 12<sup>th</sup>

### **International Conference**







# 12<sup>th</sup>

### International Conference



Figure 2: Matching Stage Flowchart







Figure 3: Entity Relational Diagram

#### 3.4 Simulation Tool

The simulation tools used include:

- Visual studio 2010
- System.Speech in .net
- Microsoft Sql server 2008

#### 4. IMPLEMENTATION

The web based application to model a multifactor authentication system was developed and in Visual Studio 2010 development run environment. The interfaces at different levels of implementation are shown in this section. The application proceeded in three stages, namely, the enrollment stage, where the diction, speech engine and pronunciation adjustment is done; profile acquisition and management module; and the matching module. Figures 4 and 5 show the interface during diction and speech engine selection.



Figure 4: Starting Diction Selection Process



Figure 5: Interface to Select Diction

In the profile acquisition module, basic information of social media users is captured. The voice biometric obtained is matched to the username stated in this module. Figure 6 showed the interface where the information is acquired.



Figure 6: Data Acquisition Interface

The matching module presents the multi-factor authentication system, as it matches the identity provided by the voice biometric matching module with that provided when username and password are entered. Figure 7 and 8 showed the username and password interface and voice biometric matching interface, while figure 9 shows the user valid dialog box.



Figure 7: Username and password Authentication page



Figure 8: Authentication system showing "invalid user dialog box"





Figure 9: Authentication system showing "valid user dialog box"

#### 5. CONCLUSION

This study presents a multi-factor authentication system toward improving social media security. The developed system would help improve social media security system by adding extra layer of biometric based security on it and help avert intrusion and data theft, which occur as a result of social media profile compromise.

#### 6. **REFERENCES**

- Adams, A. and Sasse, M.A., 1999. Users are not the enemy of Communications. In Communications of the ACM, Vol. 42, Issue 12, pp. 40-46.
- Al-Fairuz, M. A. S., 2011. "An investigation into the usability and acceptability of multi-channel authentication to online banking users in Oman". University of Glasgow, Scotland.
- Ayannuga, O. O., 2012. A Usable Knowledge-Based Authentication System (UKBAS). A PhD Thesis submitted to the Department of Computer Science, College of Natural Sciences, Federal University of Agriculture, Abeokuta, Nigera.
- Ayannuga, O. O. and Lawal, O. N., 2012. Usable Authentication Schemes: A Critique. In IEEE

African Journal of Computing and ICT, Vol.5, No. 6, 86-92.

- Egelman, S., Sotirakopoulos, A., Muslukhov, I., Beznosov, K and Herley, C., 2013. Does my password go up to eleven?: the impact of password meters on password selection. In CHI '13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, USA, pp 2379-2388
- Gonzalez, E.P., 2008. Content Authentication and Access Control in Pure Peer-to-Peer Networks. Universidad Carlos Iii De Madrid, Spain.
- Marise, M. and Micheal, L., 2010. The Adoption of Single Sign-On and Multifactor Authentication in Organizations – A Critical Evaluation Using TOE Framework. Issues in Informing Science and Information Technology, Vol. 7(1), pp161-189.
- Richard E. S. 2001. Authentication: From Passwords to Public Keys. Addison-Wesley Professional, Boston.
- Saw, G., Abbott, W., and Dinghy J., 2012. Social media for international students – it's not all about Facebook. Proceedings of the IATUL conferences. Nanyang, Singapore, pp156 - 174.
- Zezschwitz, V.E., Luca, A.D., and Hussmann, H., 2013. Survival of the Shortest: A Retrospective Analysis of Influencing Factors on Password Composition. *Human-Computer Interaction* – *INTERACT* 2013, Volume 8119, 2013, pp 460-467.
- Sancho, D., 2009. Security Guide to Social Networks. Trend Micro Incorporated, Califonia.
- Xiaoyun S., Ying, Z., and Scott, O., 2005. Graphical Passwords: A Survey. In the proceeding of 21<sup>st</sup> Annual Computer Security Applications Conference (ACSAC'05).



### A CRYPTOGRAPHIC ANTI-PHISHING SCHEME

#### \*A. A Orunsolu<sup>1</sup>, A.S Sodiya<sup>2</sup>, I.M Fadeyi<sup>3</sup> and S.A Onashoga<sup>2</sup>

<sup>1</sup>Department of Computer Science Moshood Abiola Polytechnic, Abeokuta, Ogun State <sup>2</sup>Department of Computer Science Federal University of Agriculture, Abeokuta, Ogun State <sup>3</sup>Data Analysis Department, Neuropsychiatric Hospital Aro, Abeokuta, Ogun State

<sup>1</sup>orunsoluabiodun@yahoo.com; <sup>2</sup>sinaronke@yahoo.co.uk; <sup>3</sup>mayordman@yahoo.com

#### ABSTRACT

In recent times, numerous types of attacks have been reported behind HTTP connections. One of such attacks that have generated serious concerns for both service providers and online users is phishing attacks. Phishers continue evolving new patterns despite the existence of various countermeasures. Sequel to this, this research proposed a Cryptographic Anti-Phishing Scheme (CAPS) that mitigates Phishing attacks on the basis of integrity checking and which is achieved through mutual authentication between a user and the server. The proposed approach allows website user to differentiate between a legitimate website from a counterfeit one by simply verifying the signature of the website with its server's unique public key. In the same way, the server authentication. The computational efficient Elliptic Curve Digital Signature Scheme (ECDSS) was employed as a building block for the construction. The security of the proposed scheme relies on the Elliptic Curve Discrete Logarithm Problem (ECDLP) on which the security of Elliptic Curve Cryptosystem (ECC) is based. Finally, we present the security analysis of the scheme and discuss possible attack forms with their countermeasures.

Keywords: Authentication, Cryptography, Elliptic Curve, Key Management, Phishing

#### 1. INTRODUCTION

The increasing acceptability of digital option for many critical applications and services is legendary as the whole world attained the status of global village. Examples of such applications are financial services, retail services, on-line news channels and digital libraries (Topkara et al,2005). This has led to aggressive colonization of manual methods by the Internet-based technology which has created a lot of opportunity in terms of automatic availability of services, global coverage, efficiency, reliability and zero-delay of service delivery. Despite these noble contributions of Internet to service delivery and management, the security issue of the cyberspace has become a key concern to the stakeholders: service providers, online companies and users who patronize these services (Kumar and Kumar, 2014). The Web has become a hotbed for supporting a wide range of criminal enterprises such as spamadvertised commerce, financial fraud and as a vector for propagating malware. Numerous types of attacks are hidden behind Hyper Text Transfer Protocol (HTTP) connections such as phishing, cross-site scripting and botnet attacks (Cao et al., 2013). The incidence of Internet security reached alltime high with the advent of phishing, a costly



attack that threaten stakeholders' confidence in the usability of Internet for driving major services (Abdelhamid et al. ,2014). Phishing is a major security threat to the online community and one of the fastest growing scams on the Internet.

Phishing is a web-based criminal act that uses social engineering and technical subterfuge to exploit Internet users and acquire sensitive data (Gowtham and Krishnamurthi, 2014). Social engineering schemes use spoofed emails to lead internet users to counterfeit websites designed to trick recipients into divulging financial or personal data such as credit card numbers, answer to security questions, passwords, PIN etc. On the other hand, technical subterfuge schemes plant crime-ware onto PCs to steal credentials directly, often using Trojan keylogger spyware. One of the motivations that keep Phishers in this malicious business is the prevalent use of phishing toolkits that can create a phishing page in a very short time (Xiang et. al., 2012). Another incentive for phisher is their ability to maintain anonymity despites the huge financial losses occasioned by their activities(Longe and Wada, 2012),

Despite the existence of various countermeasures, phisher continues evolving new patterns and the problems of zero-day attack, high true positive rate, low false positive rate and high computational overhead still possess challenges for anti-phishing schemes. In addition, the server/client-side approaches employed in most works have been recently criticized as inadequate. The challenges of client-side method (e.g. browser plug-in, toolbars etc.) were identified to include the use of specific browser (e.g. spoof Guard on Mozilla), user intensive administration (configuration, installation and upgrade), high browser exploits and vulnerabilities of over 79% (Ofuonye and Miller, 2013). The server filters is another popular alternative in most anti-phishing schemes which suffers from the challenges of trust and third parties involvement (e.g. SSL certificate etc.).

To this end, a cryptographic anti-phishing defense scheme (CAPS) that mitigates Phishing attacks on the basis of integrity checking and which is achieved through mutual authentication between a user and the server where a particular website resides was proposed in this paper. As against the conventional authentication mechanism in which only the server verifies user's identity, mutual authentication implies that both the user and server verify each other's identity before a session begins proper. The computational efficient Elliptic Curve Digital Signature Scheme (ECDSS) was employed as a building block for our construction. The security of the proposed scheme relies on the Elliptic Curve Discrete Logarithm Problem (ECDLP) on which the security of Elliptic Curve Cryptosystem (ECC) is based. Ultimately and generally, ECC based protocols exhibit a greater and better performance in terms of computation and security with smaller key sizes when compared with other available asymmetric cryptosystems (such as RSA, DSA and Elgamal), hence the choice of ECDSS in our construction.

The rest of the paper is organized as follows: In section 2, literature review and preliminaries to the study are presented. In section 3, the proposed scheme was discussed in detail, while section 4 discussed the security analysis of the scheme and possible attack forms with their countermeasures and finally section 5 conclude the paper.

#### 2. RELATED WORKS

Anti-phishing research has attracted a lot of interests from security experts from both academics and IT industries. Governments are contributing their own quota by enacting laws that ensure the prosecution of cyber criminals. The legal implications of these laws have been an intense subject of debates by legal practitioners. This section provides overview of these existing antiphishing studies and techniques.

A number of studies have examined the reasons that people fall for phishing attacks. For instance, Dhamija et al 2006 identified lack of computer system knowledge, lack of knowledge of security and security indicators, visual deception and bounded attention. The authors further showed that a large number of people cannot differentiate between legitimate and phishing web sites, even when they are made aware that their ability to identify phishing attacks are being tested.

The instruments of Law in the race against phishing were considered in Larson (2010) and Lovet (2009). Lovet (2009) examined the judicial challenges and recommended for expedient



international cooperation and harmonization of cyber-criminal offences amongst legal systems beyond borders. Larson (2010) recommended that courts should consider either large-scale damages against individual phishers or secondary liability against Internet Service Providers (ISP) under the areas of either intellectual property (IP) or unfair competition law.

An efficient phishing webpage detector, which uses a combination of search engine to detect status of a webpage base on deviations from the sites prescribed identity using extracted heuristics, was proposed (He et al., 2011). The approach does not suffer from client side vulnerabilities and the cost of deployment is low, which increases the practicability of the proposed solution. On the downside, the solution fails to detect form-based anomalies, multi-page phishing and XSS attacks.

The use of an automated individual white-list (AIWL) to protect users' online credentials using Naïve Bayesian classifier to automatically maintain an individual white-list of a user was proposed in (Han et al.,2012). AIWL could effectively defend users dynamic pharming attacks and is an efficient tool for protecting web digital identities. However, the system does not prevent pharming attacks. Again, there is possibility of new login problem in which a user is unnecessarily warned when interacting with new system.

A new method based on heuristics and machine learning which depends on profiling phishing attacks has offered significant progress in the quest against phishing. A three-tier classification approach was investigated (Islam and Abawajy, 2013) to detect phishing emails where accuracy of detection is up to 97%. Though interesting, this technique suffers from lengthy training time and complexity of analysis before classification decisions are made.

A novel methodology to detect phishing attacks and discover the entity/organization that the attackers impersonate using Conditional Random Field and Latent Dirichlet Allocation was presented (Ramanathan and Wechsler, 2013). The proposed solution was found to have a discovery rate of 88.1%. It also helps the legitimate organization to take down the offending phishing site. Unfortunately, the associated problems with the approach are that of scalability, need for robust parser and inadequate for online transaction due to lack of real-time detection.

Recently, Gowtham and Krishnamurthi (2014) came up with an architecture which detects phishing based on features extracted from web pages. These heuristics were fed as an input to a trained learning algorithm to detect phishing sites. To reduce computational overhead the framework introduced screening modules consisting of preapproved site identifier and Login Form Finder. The strengths of the framework include protection of users from pharming and multipage phishing attacks as well as being a stateless approach because it does not record any personal data of the user. However, the framework involves user intensive administration (configuration, installation and upgrade) and high browser exploits. Also, high categorization time associated with the scheme results in high overhead.

#### 2.1 Elliptic Curve Cryptosystem

The Elliptic Curve Cryptosystem (ECC) was independently proposed in Miller (1987) and Koblitz (1987). Since then, ECC has received intense scrutiny from Cryptographers, Mathematicians and Computer Scientists around the world to the extent that today, ECC is not just practical but it is the most efficient public-key cryptosystem known. The fact that no significant weaknesses have been found has led to high confidence in the security of ECC. The equation for elliptic curve over the field of real numbers is generally expressed as:

$$y^2 = x^3 + ax + b$$
 (1)

for some fixed values of the domain parameters *a* and *b*.

The addition of points on elliptic curve forms a group operation that forms the basis of constructing ECC. Given any two points and their coordinates, say  $P_1=(x_y, y_1)$  and  $P_2=(x_2, y_2)$  on the curve, the following show how a point  $P_3=(x_3, y_3)$  can be obtained:

$$x_3 = s^2 - x_1 - x_2$$
  
 $y_3 = s (x_1 - x_3) - y_1$ 

where slope s=

 $(y_2 - y_1)/(x_2 - x_1)$ ; if  $P_1 = P_2$  (point addition)

$$(3x_1^2+a)/(2y_1)$$
; if  $P_1 = P_2$  (point doubling)

It should be noted that:

$$P_1 + P_2 = \mathbf{O}$$
 (if  $P_2$  is inverse of  $P_1$  denoted as  $-P_1$ )

$$P_1 + \mathbf{O} = \mathbf{O} + P_1 = P_1$$
 (**O** denotes point at infinity

178


Figure 1 depicts how two different points are added on elliptic curve while figure 2 illustrates point doubling.









Moreover, the security of ECC relies on the difficulty of solving the Elliptic Curve Discrete Logarithm Problem (ECDLP). That is, given points P and Q, it is hard to find a number k such that:

$$Q = kP$$
 (2)

if k is sufficiently large. Again, several methods have been used to encrypt and decrypt using elliptic curves. The commonest one is to simulate the Elgamal cryptosystem using an elliptic curve over GF(p) or  $GF(2^m)$ . Thus, the major components of ECC are key generation, encryption, decryption and digital signature.

#### 3. THE PROPOSED SCHEME

Cryptographic techniques have offered various security services that mitigate the influence of cyber crooks in online communities. To this end, we present an antiphishing scheme tagged CAPS based on the Elliptic Curve Digital Signature Scheme (ECDSS). The ECDSS is a digital signature component of ECC and a variant of Digital Signature Algorithm (DSA). The proposed solution uses an elliptic curve over Galois field of prime numbers denoted as GF(p). The elliptic curve over GF(p) is discussed below.

#### 3.1 Elliptic Curve over GF (P)

The field GF(p) is generally used in cryptographic applications because operations over the field of real numbers are slow and inaccurate due to roundoff error. Also, elliptic curve over the field of real numbers generates smooth curves that are easier to attack in polynomial time. The equation for elliptic curve over GF(p) with characteristic p > 3 is given by:

$$y^{2} = (x^{3} + ax + b) \mod p$$
 (3)

The notation Ep(a, b) represents all the points (x, y) that obey equation (3), Ep(a, b) also includes the point at infinity denoted by **O** that serves as the identity for elliptic curve group operations. The points in Ep(a, b) are set of coordinates (x, y) such



that the relationship for non-singularity of elliptic curve points holds .i.e.

 $(4a^3 + 27b^2) \mod p \neq 0$  (4)

The implication of using a singular elliptic curve is that the elliptic curve discrete logarithm problem on such curve is computed in polynomial time. A cryptographic algorithm with polynomial time complexity is said to be insecure because the algorithm can be easily broken without much computational difficulty.

#### 3.2 The Scheme (CAPS)

Assuming we have a set  $W=\{w_1, w_2, w_3, \dots, w_n\}$  of websites residing respectively on server set  $S=\{s_n\}$  $s_2$ ,  $s_3$ ,.., $s_n$ }, the challenge usually faced by users of a particular website  $w_i \in W$ (i=1,2,3,...n) is diffrentiating between website  $w_i \in W$  and another similar website  $w_i \in W$ . Failure to distinguih between a genuine website and a counterfit one by a user usually results in information leakage. This happens when the innocent user unknowngly supplies cruicial and confidential information to a fake web page. In order to guide against this form of phishing attack, we hereby propose the use of mutual authentication between a user and server. This implies that both user and server verifies each other's identity before any online transaction takes place. The phases of CAPS are as follows:

#### 3.2.1 Key Generation

Prior to commencement of usage of a website  $w_i \in W$  residing on a server  $Si \in S$ , the web designer (Admin) chooses a non-singular elliptic curve E(a, b) over GF(p). Also, a base point  $G = (x_G, y_G)$  on the curve and a prime number q (the field size) are selected. The web designer then uses these parameters to generate private/public key pair for his website  $w_i$  using the elliptic curve key generation procedure as follows:

**Step 1:** Admin chooses a private integer  $d_{wi}$  such that  $0 < d_{wi} < q$ .

**Step 2:** Admin calculates  $e_{wi} = d_{wi} \times G$ 

**Step 3:** Admin then publishes  $(e_{wiv}, p)$  as  $w_i$  public key and keeps the private key  $d_{wi}$  a secret.

The private/public key pair generated are preconfigured onto the server  $s_i$  where the website  $w_i$ resides. It should be noted that  $w_i$  also shares Ep(a, b), G and q with every user of the website.

Furthermore, when a new user *U* logs on to the website  $w_i$  which resides on server  $s_i$  for the first time, he chooses a user name *Uid* and the server  $s_i$  completes the registration process by generating private/public key pair for the user as follows:

**Step 1:**  $S_i$  chooses a private integer  $d_u$  such that  $o < d_u < q$ .

**Step 2:**  $S_i$  calculates  $e_u = d_u \times G$ 

**Step 3**:  $S_i$  sends  $(e_w, p)$  and  $d_u$  to U as his public key and private key respectively.

On receiving private/public key pair from the server,

U announces  $(e_u, p)$  and keeps  $d_u$  a secret.

#### 3.2.2 Entity Authentication

This is a process by which a registered website user U and the server  $S_i$  hosting the website  $w_i$  verify each other's identity. The proposed Elliptic Curve Digital Signature based authentication protocol is a two-way scheme (mutual) and involves the following steps:

(i) The protocol is started by a user  $U \log on$  to website  $w_i$  and connecting to server  $S_i$  through a user authentication interface on the website.

(ii)  $w_i$  prompts U to supply his username Uid and public key ( $e_w$  p).



(iii)  $S_i$  compares the response with the one registered during registration of U for a match.

(iv)  $S_i$  validates U's identity if a match is found.

(v)  $S_i$  then generates two random number  $R_{Si}$  and  $k_{Si}$ , signs the numbers with its own private key  $d_{wi}$  and sends the resulting tuple  $(R, r_{xi}, s)_{Si}$  to U for verification. This is to ensure that U is communicating with the right server and not a counterfeit one before proceeding with his actual transaction on the website.

(vi) U verifies the signature with  $S_i$ 's public key  $e_{wi}$  and terminates his transaction if the signature is invalid.

(vii) For a valid server signature, U proceeds by generating his own two random number  $R_u$  and  $k_u$ , signs the numbers with his private key  $d_u$  and sends the resulting tuple  $(R, r_{xy}, s)_u$  through  $w_i$  to server  $S_i$  for verification.

(viii)  $S_i$  verifies the signature with U's public key  $e_u$ , allows U to continue his main transaction if the signature is valid and deny him access if otherwise. The proposed authentication protocol is detailed in figure 3.







Figure 3: Server-user authentication flow in CAPS

The proposed mutual authentication protocol as shown in figure 3 serves two major purposes:

(i) It ensures that a registered user of a particular website can continually verify the authenticity of the website using the website's unique public key. With this, the user can easily identify counterfeit websites (assuming he has at least once used the original website before).

(ii) It also safeguards the user from being impersonated by an attacker who intends to use the user's credentials to commit cybercrimes. This will continue to hold as long as the user does not divulge his private key to anyone.

**Correctness:** The correctness of the signature verification by either side (user or server) is of

**Proof:** To show that  $v_{x2} = r_{x1} \pmod{q}$  in the signature verification process, we first show that  $v = r_1 \pmod{p}$ . This is because  $v_{x2}$  was extracted from v while  $r_{x1}$  was extracted from  $r_1$ . Therefore, the proof goes as thus:

 $v r_1 \pmod{p}$  (Assumption)

Recall that  $v [(u_1 \times G) + (u_2 \times e_{si})] \mod p$ 

But  $u_1$  (R × w) mod q and  $u_2$  ( $r_{x1}$ ×w) mod q

- v  $((R \times w \times G) + (r_{x1} \times w \times e_{si}) \mod p) \mod q$
- v  $(w((R \times G) + (r_{x1} \times e_{si})) \mod p) \mod q$



Note that from key generation,  $e_{si} = d_{si} \times G$ 

(i.e. 
$$e_{wi} = d_{wi} \times G$$
)

- v  $(w((R \times G) + (r_{x1} \times d_{si} \times G)) \mod p) \mod q$
- v  $(w \times G(R + (r_{x1} \times d_{si})) \mod p) \mod q$  (\*)

Recall that s  $[k^{-1}(R+(d_{si} \times r_{x1}))] \mod q$ 

Making k subject of the equation, we have

 $k [s^{-1}(R+(d_{si} \times r_{x1}))] \mod q$  (\*\*)

Recall also that w s<sup>-1</sup> mod q

Substituting w for  $s^{-1}$  in (\*\*), we have

 $k \quad [w (R + (d_{si} \times r_{x1}))] \mod q$ 

Substituting for k in (\*), we have

 $v k \times G \pmod{p} r_1$ 

Hence,  $v_{x_2}$  (k × G mod p) mod q  $r_{x_1}$ 

If  $v_{x2} \neq r_{x1}$ , then a fake website (who does not know the private key associated with the public key of the legitimate website) may have attempted to forge the signature. The user simply terminates his transaction in case of this suspicious occurrence.

#### 4. SECURITY ANALYSIS

Assuming there is a counterfeit similar website  $w_j \in W$  with private and public keys  $d_{wj}$  and  $(e_{wj}, p_{wj})$  to the legitimate website  $w_i \in W$  having private and public keys  $d_{wi}$  and  $(e_{wi}, p_{wi})$ , we show that the proposed scheme allows a registered website user U to distinguish between a genuine website  $w_i$  and imitated website  $w_j$  by proving that the signature verification described in the last section will not be valid if  $w_j$  (without knowing wi's private key  $d_{wj}$ ) has forged  $w_i$ 's signature with his own private key  $d_{wj}$  as follows:

**Proof:** Recall that from figure 3, U computes the following to verify website wi's signature

1. U computes  $w = s^{-1} \mod q$ 

2. U computes  $u_1$  (R × w) mod q and  $u_2$  ( $r_{x1}$ × w) mod q

3. U computes  $v = (u_1 \times G) + (u_2 \times e_{Si}) = (x_2, y_2)$  [ $w_i$  resides on  $s_i$ , thus  $e_{wi} = e_{Si}$ ]

4. U calculates  $v_{x_2}$   $x_2 \mod q = r_{x_1}$ 

Let us now assume that  $w_i$  has used its private key  $d_{wj}$  to forge wi's signature in an attempt to imitate  $w_i$  and launch phishing attack on user U, the verification process by U (using  $w_i$ 's public key  $e_{wi}$ ) will not yield a valid signature as shown below:

 $v r_1 \pmod{p}$  (Assumption)

Recall that  $v [(u_1 \times G) + (u_2 \times e_{si})] \mod p$ 

But  $u_1$  (R × w) mod q and  $u_2$  ( $r_{x1}$ ×w) mod q

v ((R × w× G) + ( $r_{x1}$ ×w× $e_{si}$ ) mod p) mod q

v  $(w((R \times G) + (r_{x1} \times e_{si})) \mod p) \mod q$ 

Note that  $e_{si} = d_{si} \times G$ ,  $e_{sj} = d_{sj} \times G$  and  $d_{si} \neq d_{sj}$ 

v  $(w((R \times G) + (r_{x1} \times d_{sj} \times G)) \mod p) \mod q$ 

v 
$$(w \times G(R + (r_{x1} \times d_{sj})) \mod p) \mod q$$
 (#)

Recall that s  $[k^{-1}(R+(d_{sj} \times r_{x1}))] \mod q$ 

Making *k* subject of the equation, we have

 $k [s^{-1}(R+(d_{sj} \times r_{x1}))] \mod q (\#)$ 

Recall also that  $w = s^{-1} \mod q$ 

Substituting w for  $s^{-1}$  in (##) above, we have

 $k [w (R + (d_{sj} \times r_{x1}))] \mod q$ 

Substituting for k in (#) above, we have

Since  $d_{si} \neq d_{sj}$  then  $v \quad k \times G \pmod{p} \neq r_1$ Hence,  $v_{x_2} \quad (k \times G \mod p) \mod q \neq r_{x_1}$ 



The signature verification by *U* is invalid because he used the legitimate website  $w_i$ 's public key  $e_{wi}$  to verify  $w_j$ 's signature that has been computed with  $w_j$ 's private key  $d_{wj}$ , which has no connection with  $w_i$ 's public key  $e_{wj}$ . To successfully forge  $w_i$ 's signature, designer of website  $w_j$  has to obtain private key  $d_{si}$  from  $e_{si} = d_{si} \times G$  which is ECDLP. Alternatively, the designer of  $w_j$  may decide to obtain  $d_{si}$  from the signature equation:

s  $[k^{-1}(R+(d_{si} \times r_{x1}))] \mod q$  (5)

However, using different value for k per signature makes this even a more complicated approach for such phisher to determine  $d_{si}$ . Also, using different values of k per signature increases the complexity of *brute force* attack on  $w_i$ 's signature.

Generally, the larger the prime number p in  $E_p(a,b)$ , the more number of points on the elliptic curve and as a result the harder for an attacker to break the cryptosystem. Therefore, we suggest that a website Administrator should carefully choose the curve parameters for his server/website in order to prevent guess, exhaustive search and forgery attacks by imitators.

#### 5. CONCLUSION AND FUTURE WORK

As technology advances with unprecedented reliance on Internet by business and individuals, the Phishing techniques being used are also getting advanced and hence it demands that anti-phishing techniques are continuously upgraded to meet with new challenges. In this paper, a Cryptographic Anti-Phishing Scheme based on the Elliptic Curve Digital Signature Scheme (ECDSS) was proposed. The proposed scheme allows a website user to differentiate between a legitimate website from a counterfeit one by simply verifying the signature of the website with its server's unique public key. The scheme also allows the website to authenticate the user with the user's unique public key instead of the commonly used password-based authentication. The ECDSS based mutual authentication protocol increases the complexity of dictionary attack as well as exhaustive search attack that are usually associated with the traditional password-based authentication. In future, we hope to devise more effective and efficient anti-phishing system based on other areas of information security other than Cryptography. This is to eliminate the challenge of having to store private keys proportional to the number of websites a user is registered.

#### 6. REFERENCES

Abdelhamid N., Ayesh A. and Thabtah F. Phishing detection based Associative Classification data mining. Expert System with Applications, 2014.

Cao L., Probst T. and Kompella R. PhishLive. A View of Phishing and Malware Attacks from an Edge Router. Springer, 2013.

Dhamija, R., Tygar, J.D. and Hearst M., Why phishing works, Proceedings of the IGCHI Conference on Human Factors in Computing Systems, ACM ,pp. 581-90, 2006.

Gowtham R., Krishnamurthi I., A comprehensive and efficacious architecture for detecting phishing webpages, Journal of Computers and Security, 2014.

Han W., Cao Y., Bertino E. and Yong J. Using automated individual white-list to protect web digital identities, Expert Systems with Applications, 2012.

He M., Horng S., Fan P., Khan M.K., Run R., Lai J., Chen R., and Sutanto, A. An efficient phishing webpage detector", Expert Systems with Applications, vol. 38 2011.

Islam R. and Abawajy J., A multi-tier phishing detection and filtering approach, Journal of Network and Computer Applications, 2013.



Koblitz N. Elliptic curve cryptosystems, Math. Comp., 48(177), pp. 203-209, 1987.

Kumar G. and Kumar K. Network security – an updated perspective. Systems Science and Control Engineering, An Open Access Journal, (2014).

Larson J.S. Enforcing intellectual property rights to deter phishing, Intellectual Property & Technology Law Journal, vol. 22, I, pp. 1-8, 2010.

Longe and Wada. Hunter Hunted, Journal of Computer Applications, 2012.

Lovet G. Fighting cybercrime: technical, juridical and ethical challenges, Proceedings of the Virus Bulletin Conference, pp. 63-76, 2009. Miller V. Use of EllipticCurves in Cryptography, In Advances in Cryptology –CRYPTO '86, LNCS 263, Springer -Verlag, pp. 417- 426, 1987.

Ofuonye E. and Miller J. Securing web-clients with instrumented code and dynamic runtime monitoring, Journal of Systems and Software 2013.

Ramanathan V. and Wechsler H. Phishing detection and impersonated entity discovery using Conditional Random Field and Latent Dirichlet Allocation, Journal of Computers and Security 2013.

Topkara M., Kamra A., Atallah M. and Nita-Rotaru C. ViWiD: Visible Watermaking based Defense against Phishing. 2005

Xiang G., Hong J., Rose C. P. and Cranor L. Cantina+: A feature-rich machine learning framework for detecting phishing websites, ACM Transactions on Information and System Security (TISSEC), 14(2), 2– 21, 2012.



### DESIGN AND DEVELOPMENT OF A MULTI-MODAL BIOMETRICS SYSTEM WITH DE-DUPLICATION FUNCTIONALITY

#### Chris C Nwannenna

Condata Systems Limited, 4, Omoighodalo Street Ogudu G.R.A Lagos

condatasystems@live.com

#### ABSTRACT

Biometrics has been embraced widely by both civil and military formations in security and defense applications for the purposes of tracking persons and identifying them uniquely. The accuracy of a biometricsenabled database however depends on the tools and technologies used to develop it. In this work we design and develop a 2-modal biometrics systemthat is very efficient not only for authenticating persons uniquely but also for eliminating any possibility of double registration.

Keywords: Access Control, Authentication, De-Duplication, Biometrics Modalities, Identification.

#### 1.0 INTRODUCTION

Biometrics is a Technology that measures the intrinsic physical or behavioral characteristics of human beings. A typical physical characteristic is fingerprints which are permanent and different for each individual even for identical twins. Because these natural characteristics are different for each individual and are also measurable for each individual the Technology of biometrics relies on these features and factors to develop systems that automatically identify people uniquely.

The uniqueness of identification that biometrics provides is making the latter a powerful tool for tracking criminals in forensicsand crime scene investigations. Areas of use in civil applications include election control systems, Access Control- including Passenger Arrival and Departure monitoring systems at aviation terminals.

#### 1.1 FINGERPRINTS

Thespecific features that characterize a person's fingerprints in such a way that it is unique to him and different from those of others has been discussed<sup>1</sup>. These features include how the **Friction Ridge Skin** (**FRS**)in the person's finger is composed of. The FRS has ridge units that are aligned and fused differently for each individual to form his unique fingerprints.

#### 1.2. BIOMETRICS ALGORITHMS

These are mathematical functions derived from modeling the unique features of fingerprints. These functions or **Algorithms** are used to recognize the unique features of fingerprints.



The Biometrics Algorithm provides an Application Programming Interface (API) which programmers adopt for integration to database applications.

#### 1.3 MODALITIES

Biometrics has advanced from its earlier adoption which centered only on fingerprints to more areas that incorporate other human physical characteristics such as Face, Iris, Palm, Voice, Gait, etc. Each of these physical characteristic is referred to as a Modality in biometrics parlance. The performance of any of these modalities in any identification process is a function of the superiority or inferiority of the algorithm that is used to model, characterise and implement it.

#### 2.0. ENROLLMENT

The process of capturing and storing images of biometric traits (such as fingerprints, faces irises, etc.) in a database is known as Enrollment. The biometric system provides basic framework of functions to: i) to talk to a sensor device, such as a fingerprint scanner, camera, etc, ii) capture an image (of face, fingerprint, iris, palm, etc), and iii) extract the unique minutiae data from the image.

The minutiae data, known as *template,* is what is stored in the database.

#### 2.1. AUTHENTICATION/VERIFICATION

Theessence of extracting biometrics information from a subject, (or enrollee) is to either store it in the database or to compare it to existing information in the database. The comparison of two sets of extracted templates is done during Authentication and Verification processes. In an Authentication process an enrollee's biometric data is once again acquired and the system uses this to identify who the enrollee is. This involves comparing the acquired biometric information against templates corresponding to all enrollees in the database. Verification on the other hand involves the comparison with only those templates corresponding to the claimed identity. Thus while Verification is a 1:1 comparison Authentication is a 1: N comparison where N>1.

#### 3.0. UNIMODAL BIOMETRICS

From our experience most of the biometrics systems in use in the country currently are built around fingerprint modality. Known in the industry as an AFIS(Automatic Fingerprint Identification System) the Enrollment Process is typical of what is given in Figure 1. A photograph of an enrollee, when it is part of the input requirements of the system, is incorporated manually using a camera (webcam, for example).

The level of confidence implicit in this type of system is very low. The reasons for this arise from a number of factors.

i) By capturing photographs manually it is easy

touse an image cut from newspapers to createa false record (spoof). A typical example of this occurred some years back when thephotograph of the famous boxer, Mike Tyson, was used to create a record in the registration database of an important government agency in Nigeria, INEC, to be precise.

ii) The use of one modality makes it possible for people to introduce multiple registrations into the database. Since fingerprint is the only requirement in terms of biometrics anybody who was able to fake a fingerprint will take advantage and rubbish the system.





#### 4.0. MULTI – MODAL BIOMETRICS

As discussed above the use of a single biometrics indicator to build enterprise biometrics systems is not ideal because, from experience, such a system is more susceptible to biometrics spoofing and therefore suffers from unacceptably high error rates. One of the key objectives of our proposed system therefore is to build a biometrics system that is robust enough to eliminate multiple errors in the areas of authentication, and verification. To this end we will adopt a scheme that features multimodal biometrics indicators. Specifically we will advocate the implementation of a system comprising of one of:

- i) face and fingerprint,
- ii) fingerprint and iris,
- iii) fingerprint and voice,
- iv) palmprint and face.

The system to be developed here will be based on (i), that is, face and fingerprint modalities.

Our multimodal system will seek to alleviate most of the drawbacks mentioned in (7) above. This it will do by providing multiple evidences (face and fingerprint) characteristics of the same person. Multiple evidences will minimise spoofing since it will be more difficult for an impostor to spoof multiple biometrics traits of a genuine user simultaneously. In summary the key features of our proposed system will include the following:

i) Fingerprint and face images captured simultaneously.

ii) Face image is captured by a biometrics engine via a camera/webcam.

iii) 'Liveness' Detection: A mechanism that will ensure that the system will respond only if it detects that a 'live' enrollee is indeed present at the point of data acquisition.

iv) Face and fingerprint images of an enrollee must be captured before his registration could succeed.

v) The registration of an enrollee succeeds only when the Matcher engine confirms that a record of the enrollee is not already present in the biometrics database.

vi) The biometrics Algorithm to be used for integration must satisfy all industry standards (NIST, FBI, ISO, BioAPI 2.0, WSQ data format compliant).

Our proposed solution constitutes what is known in the industry as an ABIS (Automatic Biometrics Identification System). With our ABIS in place, most, if not all, the errors inherent in the single modality AFIS solution, prevalent in the country and discussed earlier (see 3.0), will be eliminated.

4.1 Multimodal Biometrics Enrolment.

The architecture of our proposed biometric system is given in Fig 2 below.



12<sup>th</sup>

### International Conference



The architecture is self-explanatory. The functionality of the system is split into components with each component providing complementary services to the others. The major components available consist of the biometric engines viz:

i) The Fingerprint Extractor: Biometrics systems do not work with raw imagesextracted by devices such as scanners. Rather they work with biometric templates extracted from the fingerprint image by the Fingerprint Extractor. Templates are in a data format amenable to biometrics processing by the Matcher (see later) and records in the database are stored in this format.

The Fingerprint Extractor itself is actually an engine built from an algorithm developed by using the characteristic patterns of a fingerprint.

ii) The Face Extractor: This is similar to the Fingerprint Extractor but performs functions related to extraction of face templates from face images captured using cameras.

iii) The Fusion Engine: The fingerprint and face templates extracted by the Fingerprint and Face extractors are fused together by the Fusion engine to form a single multimodal biometric template which is stored in the database.

iv) The Matcher engine: The Matcher compares the fused template with those already stored in the biometrics database with the aim to check if it belongs to the same person.

The biometric engines are indicated in the figure as A, B, C, and D. The performance of our multimodal system will largely depend on them: that is, the ability of the two Extractors to generate reliable templates and the strength of the Fusion engine to wrap up everything well. The Matcher complements these functions by deciding correctly when to reject or accept a new enrollee. We adopted a biometrics algorithm that delivered excellently well on all key indices of our system.

### 5.0 STRATEGIES FOR MINIMISING SPOOFING

We will elaborate on some of the features listed in (4) above and add other specific requirements that are built into the system to minimise spoofing. These strategies include:



i) Technology: use of robust biometrics algorithm,

ii) User control: control of access to the system,

### 5.1 Technology Features: 'Liveness' Detection.

This feature has been mentioned earlier. The biometrics algorithm has advanced features that were implemented to ensure that the system will respond only if it detects that a 'live' enrollee is indeed present at the point of data acquisition. The data capture device, that is fingerprint scanner, complemented this feature. It too is Live-Finger-Detection compliant (see below).

#### 5.2 Technology Features: Capture Devices

The type of device used to capture a biometrics modality, such as fingerprint, plays a key role in the quality of the specimen (image). Generally the two major capture devices are fingerprint Scanners and Cameras.

#### 5.2.1 Cameras.

The recommended cameras are those that capture ISO/IEC 19794-5 compliant face images.

#### 5.2.2 Fingerprint Scanners.

The recommended fingerprint scanners are those: i) that are certified by FBI to deliver fingerprint image quality of PIV-071006. Such scanners use advanced CMOS technology and precise optical system.

ii) have special built in electronic circuit to do Live Finger Detection(LFD). This LFD feature is selected from the scanner when it is also present and implemented in the biometrics algorithm.

iii) The fingerprint scanning window should be at least 16.26 x 24.38 mm.

iv) The fingerprint image resolution should be at least 480 x 320 pixel and 500DPI.

The LFD feature will ensure that only a live fingerprint is scanned into the system. Fake fingers made from, for example, silicone rubber will be rejected.

#### 5.3 User Control

These are features that control user access to the system. These include:

 i) Login: Users are authenticated by a biometrics Login module. Current Login module is based on Fingerprint biometrics but can be extended to face and fingerprint biometrics.

ii) User trace: Login and Logout periods and activities undertaken.

#### 6.0 IMPLEMENTATION

Thetwo-modal biometrics system described above has been implemented using Microsoft .NET WinForms. The solution is currently based on three distinct modules:

i)A Biometrics Login module: featuring fingerprint modalit

ii)A Biometrics Data Capture Module (to capture 1-10 fingerprints and face images of an enrollee).

iii)A Demographic Data Capture Module(to capture name, date of birth, address, Location, etc.). The Location Information is based on State, LGA, Registration Area, and Polling Unit. This is derived from the 36 states plus FCT and the 774 LGAs as defined in the Nigerian Constitution. It is similar to that used by INEC.

Other features of the solution include:

iv)Face and Fingerprints: must be captured for each enrollee before registration can proceed beyond Biometrics Data Capture Module.

v) De-duplication: Biometrics matching and de-duplication is strictly enforced in the Biometrics Data Capture Module. The registration of an enrolleeterminates in



this module if his data is already in the database otherwise registration proceeds to the Demographic Data Capture Module. vi) Algorithms: The solution implemented face and fingerprint algorithms that have "liveness" detection functionality to

ensure that the system responds only if it detects that a 'live' enrollee is indeed present at the point of data acquisition. The algorithms are also WSQ compliant thus ensuring interoperability of templates with templates of algorithms of other providers. vii)COTS Capture Devices: The solution implements open technology that allows the use of Commercial-Off-The-Shelf capture devices. It allows for the deployment of a wide range of Fingerprint scanners and Face cameras.

In our case we have deployed two types of FBIcompliant fingerprint scanners which we use interchangeably.

a)Futronic FS 80 (1x10 fingerprint capture system), b)DactyScan84C (4+4+2, OR 1x10 fingerprint capture system).

For camera we use Logitech HD 720p Autofocus webcam.

viii) Scalability: Though we have implemented Face and Fingerprint modalities thesolution is scalable for future expansion. For example, we can implement the multibiometric functionality to enable the use of fingerprint, face, iris, palmprint or voiceprint biometrics in any combination. Costs will be an important factor in this regard.

ix) Enrollee's ID: This is generated uniquely by concatenating data relating to the enrollee's location information.

#### 7. SCREEN SHOTS OF BIOMETRICS DATA CAPTURE MODULE AND DEMOGRAPHICS DATA CAPTURE MODULE



fingerprint and face fusion center Demographics

#### 8EXTENSION OF WORK

Thecurrent work implements de-duplication at the local workstation level. Miscreantsmay of course wish to



defraud by registering several times at different local workstations.

De-duplication in this case will require that all the workstations be tied to a central biometrics database hosted by a network server, for example.Work is now ongoing to incorporate this functionality by using a distributed computing environment.

#### 9. CONCLUSION

In this work we have designed and developed an Automatic Biometric Identification System (ABIS) using face and fingerprint modalities. The ABIS systemwas tailored towards minimising, if not entirely eliminating, the penchant for Nigerians to spoof and defraud biometric registration systems in the country. The system achieved its objectives by ensuring that introducing "ghost" registration through "spoofing" is eliminated by implementing "liveness" detection at various stages of data capture. In addition uniqueness ofrecords is ensured by eliminating any possibility of duplicate registration.

#### 10. **REFERENCES:**

Arun, Ross etal, Sept 2003, Information Fusion in Biometrics. Dept of Computer Science Michigan State University,

INEC, 2006, INEC(Independent National Electoral Commission) PIC Report 2006.

Nwannenna, C. C, November 2008, Building Efficient Voters' Registration Database: Software Considerations.Proceedings of A Joint Retreat by INEC & Nigeria Computer Society Asaba, Nigeria, November2008. Pp74-79. NIST National Institute of Standards and Technology.



### A SMART HOME SIM300 GSM MODULE CONTROLLER

#### D.N, Nnakwuzie<sup>1</sup>, \*C.I, Akobundu<sup>2</sup>, and M.U, Onuh<sup>3</sup>.

<sup>1</sup>Department of Mathematics/Computer Science/Statistics /Informatics, Federal University Ndufu Alike Ikwo, Abakaliki, Ebonyi State.

<sup>2</sup>Computer Software Engineering Department, Rocana Institute of Technology, Emene, Enugu State.

<sup>3</sup>Information Communication Technology Department, Interoil Investment Nigeria Limited Oil and Gas Free Zone Onne, Rivers State.

<sup>1</sup>Okafordoris49@yahoo.com, <sup>2</sup>Onah.mases@gmail.com, <sup>3</sup>Chinyere\_i@yahoo.com.

#### ABSTRACT

With advancement in technology things are becoming simpler and easier for mankind. Automatic systems are being preferred over manual system. This research work presents the development of GSM-based control of home and office appliances. Remotely, the system allows the user to effectively monitor electrical equipment via the mobile phone set by sending commands in the form of SMS messages and receiving the appliance status, as the feedback. The main aim of this prototype is to reduce electricity wastage, provide security and replace human work force, which is prone to errors to automated system. The system will be integrated with microcontroller and SIM300 GSM module network interface using assembly language. It is activated when a user sends the SMS to the controller, on receiving the SMS command, the microcontroller unit then automatically controls the electrical appliances by switching "ON" or "OFF" the device according to the user instruction. The design technique used in this research is top down decomposition; the approach is best-suited when the problem and its environment are well defined.

**Keywords:** E-EPROM, SIM300GSM module, Microcontroller, Embedded system.

#### 1.0 INTRODUCTION

Throughout the developing world, there is a strong focus on improving information technology. In this view, many believe that the use of Information and Communication Technology (ICT) is a key that can help to address certain problems that required constant attention, like monitoring, controls, security and other commercial businesses. In early 1940's computers were used only to solve problem such as calculation and finding of trigonometric problems: As time went on the use of computer extended to things like data processing applications and result oriented programs.

Today Information technology has advanced so much in the last decade or two that it has made

life more comfortable and efficient. The comfort of being able to control devices from one particular location has become vital as it saves a lot of time and effort. With the adoption of this system, control can be gained over certain things that required constant attention. Their existing

technologies that are widely used for such transfer of information/control from one place to another are as followed:-



Bluetooth Technology, Zigbee Technology, Wi-Fi Technology, GPRS (General Packet Radio Service). SIM300 GSM Module implementation of Smart Home is based on application of the GSM technology. Using GSM Network, a control system has been proposed that will act as an embedded system which can monitor and control appliances and enable home security against intrusion in the absence of home owner.

Communication and networking technologies have advanced tremendously over the past decade. The application of mobile phone cannot be restricted to sending SMS or starting conversations. The following are the goals of the researchers:

- i. To co-ordinate appliances and other devices required for fighting fire during fire outbreak case without risking life at very high temperature through short message services (SMS) send by ordinary mobile phone.
- ii. There is no way; power consumption can be minimized specifically now prepaid meter is used in various homes and offices.
- iii. The life span of electrical home and office appliance are reduced to the lack of devices that will help in the monitoring and controlling of the electrical devices on weather change while in office or market.
- iv. Human power are been wasted, in monitoring and control appliance in home and industries where usage of electrical appliance are highly required just to ON or OFF the equipment.
- v. It is difficult to operate machine and gadget that are hazardous to operate, so many people lost their lives in performing task that are beyond human capability size and speed in the existing system (example, the work of the fire services people)

The main aim of this research is to develop an interface between the GSM and remote electrical appliances that will be able to:

- i. Co-ordinate appliances and other devices through short message services (SMS).
- ii. Effectively receive and transmit data via SMS.
- iii. Design a circuit that can automatically switch 'ON' and 'OFF' the home appliances.

#### 2.0 Literature Review.

In order to come up with the best design of the control process of the project, the researchers review previous articles, book and journal on the subject matter. The information gathered from the literature will go a long way not only in shaping the direction of the project but also in sharpening. Faisal et.al,2012 develop a system that uses Android OS based mobile giving voice command, the mobile application convert the voice into text to control the electrical device. Amit, et.al, 2011 sign remote control based on GSM mobile technology, using SMS message that is generated by the web. It is implemented based on micro-controller that receives SMS and commands from a cellular phone over the network. The microcontroller then carries out the issued commands and then communicate the status of a given applicant or device back to the cellular. Yuan Xin Lin et.al, 2014 implemented smart home using ZigBee technology, the system used a low power cost CC2430 processor as central controller. Touch screen interface allows the ZigBee technology to achieve wireless monitor-ing of home device. Pandikumar et.al, 2013 proposes architecture to enable the users to control and monitor smart devices through Internet. It creates an interface between users and smart home by using GSM and Internet technologies. In this architecture the users give commands through web then the users inputs are converted into GSM-SMS commands. These commands are sent to embedded system module (embedded system directly connect with devices) through GSM network, and finally the user commands are executed by microcontroller to control any electronic objects like home appliances and lights and it sends the acknowledgement. The embedded system module can place anywhere in the world and it will controlled by IoT, (Internet of Things) Agent through GSM network. Jayashri et.al, 2013 developed two methods for home security system. The first system uses web camera, whenever there is a motion in front of the camera it gives security alert in terms of sound and a mail is sends to the owner. The second methods send, SMS which uses GSM GPS Module (Sim 548c) and Atmega664p microcontroller, relays and buzzers. Das. C. K et.al, proposed a method which enables users to control their home appliances and system from



anywhere using cell phone. To access the control unit, then user should send an authentication code (DTMF) along with the required /desired function to his/her home control system via GSM.

#### 2.1 The Proposed System.

In our proposed system, we developed a software module that can control many ports, send back report on the condition of the appliances to the owner through the user's phone, provides security to the user electrical appliances and it strictly can be controlled by only person with password. This system can be monitored or controlled using any type of phone.

#### 3.0 Design Analysis

This section expresses the actual design analysis and the construction of this project. Design has to do with plans, sketching of circuit diagram that will suit the desired specification of the project.

#### 3.1 The Block Diagram of the System Designed

At this point each block was developed into circuits to achieve the respective peculiar function of each of them. This project circuit diagram has seven modules.



#### 3.2 Detail Design

This section of this work expresses the detailed entire modules with the respective components and devices that make up the module.

#### 3.2.1 POWER SUPPLY MODULE

This module is very important part, which none of the remaining module cannot do without. The moment this module is off, every other parts of the entire system will turn off. The power supply module has two current form, the alternate current (AC) and direct current (DC). The alternate current is used to power and monitor devices under control by the project furthermore, it is also used to power the power supply unit of the designed system while the direct current is the output power from the power supply unit that rectify and regulate to specified DC voltage level from the AC input of the system. This module consists of transformer, Capacitor, bridge rectifier, voltage regulators and other components.

### 3.2.2 Data Trans-Receiver Module (TX/RX Module)

This is the unit that serves as the gateway of the system. It receives information from the user mobile phone station and feed into Data processing module on control mode; it also receives feedback information from the processing module back to user mobile phone station on monitoring mode that it has full duplex function. This module is made up of Phone Fbus and Mbus, SIM card reader, registered and MTN SIM card.

#### 3.2.2 PROCESSING MODULE

This module is the brain behind the operation of this system. It controls both the input and output information of the entire system, processes it and gives out instruction to appropriate module for action. When the system is in control mode it receives information from the trans-receiver



module, process it and send instruction to the appropriate device the information is made for through driving module then indicating module, process it and send SMS message to the user mobile phone via trans-receivers module. This module is made up microcontroller, capacitors, crystal oscillator and resistors.

### 3.2.3 PROGRAMMING LANGUAGE PLATFORM.

It is most important to select the best platform that suits a particular application. The program that makes up this system is design using an assembly programming language. The choice of assembles programming language in this research work implement is due to its enormous advantage over other identified programming languages such as it is easier to understand and saves a lot of time and effort of the programmer, easier to correct errors and modify program instructions. Assembly language has the same efficiency of execution as the machine level language; it is one to one translator between program assembly language and its corresponding machine language program. Lastly it reduces size of the code, increase speed; use tricks to prevent disassembling debugging.

#### 3.2.4 DRIVING MODULE

This module drives the device connected to this system; it receives information from the processing module as instruction and turns the connected device on/off depending on the processing module instruction. This module is made up of NPN transistor, relays, diodes and resistors.

#### 3.2.5 INDICATION MODULE

This unit serves as the output unit of the system; it is made up of bulbs, device under control and also the user mobile phone. It is on to indicate that the appliance is "ON" and "off" to indicate the appliance status. Another indication component the appliance itself goes "ON" /"off" to indicate its operation state.

### 3.2.6 SOFTWARE/ HARDWARE INTEGRATION MODULE

This is the combination of software and hardware part required for the system to work. After the programming, compiling, simulation and debugging of the software part, the researcher integrate the software into machine language and transfer machine language into the microcontroller with the help of E-EPROM (Electrically Erasable Programmable Read-Only Memory) it is a type of non-volatile memory used in computers and other electronic devices to store information into devices. Fig 2 and fig 3 displays the hardware internal integration and external integrated part of the system respectively





Fig 2: Hardware Internal Integrated of the system



Fig 3: The External Integrate Gadget.

#### 3.3 Object-Oriented Analysis of the System.

The UML (Unified Modeling Language) is the most widely known and used notation for objectoriented analysis and design. To be an analysis or design method it must include guidelines for using the notation and methodological principles. The most useful standard UML diagrams are: - Use Case Diagram, class diagram, sequence diagram, state-chart diagram activity diagram. But the researchers explain only Use Case Diagram and state diagram.

#### 3.3.1 State Chart Diagram

The state chart diagram defines the states of a component and these state changes are dynamic in nature, so its specific purpose is to define state changes triggered by events. State chart diagram are used to model states and also events operation on the system. When implementing a system it is very important to clarify different states of an object during its life time and state chart diagrams are used for this purpose. Therefore, the main usage of state chart diagram is: to model object states of a system, to model reactive system, to identify events responsible for state changes.



12<sup>th</sup>

## International Conference



Fig 4: State Chart Diagram for Smart Home.





#### Fig 5: Use case Diagram for Smart Home

#### 3.3.2 Use-Case Diagram

A Use Case illustrates a unit of functionality provided by the system. The main purpose of the Use-Case diagram is to help development teams visualize the function of the system including the relationship of the "ACTORS" to essential processes, as well as the relationships among different use case. Below is a diagram of use Case for G.S.M control for Home and Office appliance.

#### 4.0 RESULTS AND DISCUSSION

#### 4.1 Input Specification

In the proposed system, GSM is used to send the message to the gadget, this service's as the user input form, to the system. The input depends on the user request at the particular time. The table 1 shows the type of message that will be sending to the GSM interface to "On" a particular electrical device or electrical equipment.



#### Table 1: Input Analysis Table

Password (the phone	Input Message	Status report				
number)	to the system.	to user mobile.				
08142788247	A=1	Turn A On				
07032193487	B=1	Turn B On				
08142788247	C=1	Turn C On				
08142788247	D=1	Turn D On				
08142788247	E=1	Turn E On				
08142788247	F=1	Turn F On				
208142788247	G=1	Turn G On				
08142788247	H=1	Turn H On				
08142788247	l=1	Turn I On				
08142788247	J=1	Turn J On				
08142788247	K=1	Turn K On				



Fig 7: Feedback Report to 'ON' Device



#### Fig 6: Input User Interface to 'ON' a Device.

Figure 6 shows the user interface screen. The message which is "A=1" Simple means that device A is being turn "On". When this message is received by the gadget and processed it will automatically turn the device ON. Fig 7 shows the user, feedback message received from the gadget; on the status of the device this makes the system user friendly.



Fig 8: Input interface to off Device





#### Fig 9: The Feedback on Turning the Device "OFF"

The command in the interface of figure 8 phone screen will be send to the gadget via user phone to 'off' the device.

### The Feedback Massage on Turning the Device OFF

The message which is "A=0" Simple means that device A is being turn off. The fig 9 shows the device A which is Television is turn **"OFF".** 

Fable 2	Input	Analysis	Table
---------	-------	----------	-------

Password	Output message	Status report to user
	to the	mobile.
	system.	
08142788247	A=o	A Off
08142788247	B=0	B Off
08142788247	C=0	C Off
08142788247	D=0	D Off
08142788247	E=o	E Off
08142788247	F=o	F Off
08142788247	G=0	G Off
08142788247	H=o	H Off
08142788247	l=o	I Off
08142788247	J=0	J Off
08142788247	K=o	K Off

#### 4.1.1 Output specification of the system

#### The expected output includes:

- i. The acknowledgement the user gets from a particular request on the condition of the device.
- ii. Action expected to perform on the appliances which may be 'ON' or 'OFF'.

#### 4.2 Testing of the Individual Component

The testing at this stage is carried out by the use of AVO- meter. In this stage the researchers tested the components in respect to their respective features after necessary inter and intra-module connections as specified in the diagram, (Fig: 1 The System Block Diagram of the System).



The table 2 shows the type of message that will be sending to the GSM interface to "Off" a particular electrical device or electrical equipment.

#### 4.2.1 System Testing.

At this stage all modules were linked together and we SMS are sent command from mobile phone, and the phone network service provider MTN indicates that SMS command is delivered to the system. The particular relay switch which the SMS command is made turn on, the SMS command was repeated by instructing the relay to 'OFF' the message delivered again to the processing unit, which in turn, turns off the relay switch. After this confirmation the researcher then integrate the modules by soldering them on the board.

#### 4.2.2 Expected Versus Actual Test Result

There was a little variation in the power supply voltage when the test was carried out. The change observations are negligible since it falls between the actual values. The system works perfectly at range of voltages of 3.5 volt and 12 volts.

The fig: 10 shows the expected result at the end of the test, which was represented with electrical bulls.



#### CONCLUSION

This research work that was undertaken has helped to gain a better perspective on various aspects related to our area of study as well as practical knowledge of electronic equipment and communication. The extensive capabilities of this system are what make it so interesting. From the convenience of a simple cell phone, a user is able to control and monitor virtually any electrical devices. This makes it possible for users to rest assured that their belongings are secure and that the television and other electrical appliances was not left running when they left the house to just list a few of the many uses of this system.

#### RECOMMENDATION

Remote operation of weapons in controlling of insurgency is still a difficult task to embark on in Nigeria, thus, many soldiers die daily in North East of Nigeria, hence, this automation system is recommended to soldiers and fire service men for monitoring and controlling gadget that are hazardous in nature.



#### REFERENCE

- Amit et.al, 2011, SMS based remote control system," IJCSMS international Journal of computer science and management studies, vol.11. No. 02,
- Das C.K, et.al, 2007."Development of a cell phone based remote control system," International Journal of Electrical & Computer Sciences IJECS Volume 9 No. 10.
- Faisal B, et.al, 2012 "Controlling home appliances remotely through voice command": International Journal of computer Application. Volumes 48, No.17,

2437-2440.

- Jayashri .B. and Arvind .S, 2013, "Design and implementation of security systems for smart home based ON GSM technology" : International Journal of Smart home. Vol.7 No. 6,
- Pandikumar .S. and Vetrivel .S, 2013, "Internet of things Based Architecture of web and smart home interface using GSM," International Journal of Innovative Research in search, Engineering and technology. Vol. 3 No. 3.
- Yuanxin lin Rui Kong et.al, **2014** "Design and Implementation of Smart home intranet Based on ZigBee," Research Journal of Applied science, Engineering and technology. 7(12)



### SECURING DATA IN THE CLOUD USING RSA HOMOMORPHIC ENCRYPTION

#### O. B. Alaba<sup>1</sup>, W. O. Adesanya<sup>2</sup>, <sup>\*</sup>A. S. Ogunbanwo<sup>3</sup>, and O.O. Olusanya<sup>4</sup>.

<sup>1,2,3,4</sup>Department of Computer Science, Tai Solarin University of Education, Ijebu-Ode, Nigeria

<sup>1</sup>olumuyiwaalaba@gmail.com, <sup>2</sup>sanyalanre2003@yahoo.com, <sup>3</sup>ogunbanwoafolakemi@yahoo.com, <sup>4</sup>yinka\_olusanya@yahoo.com.

#### ABSTRACT

Security is one of the major challenges that people are facing all over the world in every aspect of their lives. Similarly, security in electronic world has a great significance. This is an area of substantial interest in database because the use of database is becoming very important in this world of technology to an organization. Databases contain information that is majorly enterprise asset. In cloud computing this database is exposed to the third party (service provider) who has privilege to access the data and hijackers. The vulnerability of the data due to undue access makes security a major concern to the organization. This work takes a precursor look at a prototype that can secure data in cloud using RSA Partial Homomorphic Encryption (PHE) schemes to encrypt and decrypt the data as it goes in and comes out of the database. The prototype that will be designed provides information confidentiality, integrity, accountability and prevents abuse of special privileges because any attack on the database by any unauthorized person will only access obfuscated or meaningless cipher texts.

Keywords: Homomorphic, Security, Data, Encryption, Obfuscated.

#### 1. INTRODUCTION

Information or data is a valuable asset in any organization whether it is social, governmental, or educational. Organizations have become highly dependent on database for their daily operations because almost all have their information system and other operational functions automated, this automation makes the database to be more valuable assets in today's economy. Database form the basis for web applications such as e-commerce, e-banking, e-leaning, e-portal, e-registration, eexam, e-government, etc. Due to the sensitivity of the information or data stored, there emerges the need to protect the information. Securing the information or data has become a top priority and a major challenge for the organization. Protecting the confidentiality and sensitivity of data stored in the repository is referred to as database security. Database security is the system, process and procedure that protects data in the database from unintended activities and unauthorized access. The database security is to prevent undesired information disclosure and modification of data while ensuring the availability of the necessary service.

Outsourcing database means using an external company to manage an organization's database. This can also be referred to as database stored in the cloud. Cloud computing is the idea that data



and programs can be stored centrally, in the cloud, and accessed anytime from anywhere through thin clients and lightweight mobile devices Mark (2013). Shared platforms such as database server hardware and software are provided to host multiple outsourced databases.

Security and privacy are major challenges in cloud computing because both the hardware and software are under the physical control of the service providers that host the database. For databases deployed into a cloud database service, the service providers have the privilege to access the databases and the database might be accessed improperly by the service providers accidentally or intentionally. The potential danger of such improper accesses causes the concern of users about the privacy of their outsourced databases and the privacy of the outsourced databases can also be at risk of being breached by the attackers. To protect data in cloud databases, a straightforward approach to database security must be put in place because no matter the degree of the security that was put in place, sensitive data in database are still vulnerable to attack.

In general, encryption is a useful tool for protecting the confidentiality of sensitive data so that even if a database is compromised by an intruder, the data remains protected even after the database has been successfully attacked or stolen. Security in cloud is one of the major areas of research. The survey shows that, the researchers are focusing on efficient algorithms and encryption techniques to enhance the data security in cloud. This work focuses on encrypting data using RSA homomorphic encryption scheme before they are stored. By this way, the service providers or attackers can access only meaningless ciphertexts. RSA is a multiplicative partially Homomorphic encryption scheme. It was first discovered by Ron Rivest, Adi Shamir and Leonard Adleman, the letter RSA was derived from the first letter of their surnames.

#### 2. LITERATURE REVIEW

Kadhem, et al (2009) studied security of the databases shared between many parties from a cryptographic perspective. They proposed Mixed Cryptography Database (MCDB), to encrypt databases over untrusted networks in a mixed form using many keys owned by different parties. The addressed the issue of data framework confidentiality, privacy and integrity when the database is shared between many parties. To achieve this, the author grouped the framework into four steps. First, it introduces new data classification based on who owns the data. Second, it proposes a mixed cryptography database based on data classification methods. Third, it illustrates a query management system over a mixed encryption database. Finally, it analyzes the security of data storage and data transmission in the new cryptography framework. The proposed framework is very useful in strengthening the protection of sensitive data even if the database server is attacked at multiple points from the inside or outside but it has its own limitations in the area of performance of gueries and security analysis and access control.

Neha, et al (2012) in their paper titled Implementing DES Algorithm in Cloud for Data Security described Data security system implemented into cloud computing using DES algorithm. This framework considered security at both clients and server side. The security architecture of the system was designed by using DES cipher block chaining, which eliminated the fraud. In order to secure the system the communication between modules was encrypted using symmetric key. The author proposed that the cloud data security must be considered to analyze the data security risk, the data security requirements, deployment of security functions and the data security process through encryption. This framework was acceptably secured but it has its own limitation in the area of key distribution, scalability and authentication.

Joshi, et al (2011) argued that Asymmetric Cryptography Algorithms and Digital Signature techniques are reliable and efficient to provide more security user's data in Cloud Computing. This paper proposed a more effective and flexible



distributed verification scheme to address the data storage security issue in cloud computing, as it relied on the cryptography algorithms [RSA] and digital signature techniques, for protecting user data which include encryption prior to storage, user authentication procedures prior to storage or retrieval, and building secure channels for data transmission. This method achieved the availability, reliability and integrity of erasure in coded data and simultaneously identifies misbehaving servers i.e. whenever data corruptions will occur during the storage correctness verification. Through detailed performance analysis, it showed that the scheme should provide more security to user's data in cloud computing against failure, unauthorized data modification attacks and even server colluding attacks

The potentiality of their paper is that, they have seeded the idea of using two different keys algorithms. However, they failed to give the model or methodology that shows how and where these algorithms should be implemented.

#### 3. METHODS

From the literatures reviewed, different methods of encryption techniques have being applied to address the issue of security. However, these methods were deficient and could not adequately secure the data outsourced or stored in the cloud. Due to these shortcomings or weaknesses of reviewed existing technique, we proposed RSA homomorphic encryption technique. This method is a much more secured technique because it provides confidentiality, authentication, integrity, has better scalability, both private and public keys are kept by the client, and third party can perform computation without knowing the content. This section shows the framework to implement partial Homomorphic Encryption Scheme (PHES) in securing student's data in an outsourced database. RSA HES will be used for privacy-preservation using the iterative model to analyze and design the framework. This designed model encrypts the user's data first before it is stored in the cloud to provide security of the data, only the authorized users have the secret key to decrypt the data in other to view it correctly. RSA is a block cipher, in which every message is mapped to an integer. RSA consists of Public -Key and Private-Key. In our Cloud environment, Pubic-Key is known to all, whereas Private-Key is known only to the user who originally owns the data. Thus, encryption is done by the Cloud service provider and decryption is done by the Cloud user or consumer. Once the data is encrypted with the Public-Key, it can be decrypted with the corresponding Private-Key only.

Design architecture: In order to achieve the system flexibility, robustness, and resistance to potential change, the popular three-tier (layer) architecture is deployed. The architecture is composed of three layers namely, the user interface layer, the application logic layer and the database layer. The three-layer architecture aim is to solve a number of recurring design and development problems, and make the application development work more easily and efficiently.

Three-layer architecture:

Client/User Interface Layer: The first layer is the user interface layer. It offers the user a friendly and convenient entry to communicate with the system as it manages the input/output data and their display. The users are allowed to access the system by using any existing web browser software. The user interface layer contains HTML components needed to collect incoming information and to display information received from the application logic layers. The clients communicate with the web server via application protocols, such as HTTP and SSL, sending requests and receiving replies. Clients of the system are administrator and students.

Application Logic Layer: The application logic layer is the middle layer, which bridges the gap between the user interfaces and the underlying database, hiding technical details from the users. Components in this layer receive requests coming from the interface layer and interpret the requests into a proposed action controlled by the defined work flow in accordance with certain pre-defined rules. Application logic layer consists of a control module and three functional modules. The control module controls the flow of functions execution and transferring required information between them and the database layer. The functional modules are authentication module, result module and homormophic module, Figure1.

Database Layer: The database layer is responsible for modeling and storing information



needed for the system and for optimizing the data access. Data needed by the application logic layer are retrieved from the database, and then the results computations produced by the application logic layer are stored back in the database, Figure2. Since data are one of the most complex aspects of many existing information systems, it is essential in structuring the system. Both the facts and rules captured during data modeling and processing are important to ensure the data integrity.



Figure 1: Architecture of the Proposed System.



Figure 2: Use Case Diagram for Proposed System

#### 4. RESULTS AND DISCUSSION

This section presents some actual results from the implemented application. We ran a few examples in the actual application showing the final look and feel of the Graphical User Interface (GUI) and Use Dialogs including real results from the transactions made to the database located with the mySQL server.

First Step is to make the server active in order to start the program.

To start the program, these three steps were followed.

i. Ensure that the server is on, by double clicking the WampServer icon on the desktop



Figure 3

- Right click the green wampserver icon shown on the tool bar and choose Local Host (Figure 3)
- iii. To view the web page, click on RSAEncryption (Figure 4).





The web page shown in Figure 5 below would be displayed immediately after the three steps. The page serves as welcome page and the main GUI and is the first interaction with the user after starting the program. As discussed and shown in the architecture of the proposed system, the Client Interface Layer has two clients, the student and administrator.





STUDENT LOGIN:

i. Click on student to display the page below (Figure 6), the page asks for the student's matriculation number to be entered. Figure 7

- The reply to the command would be in two forms:
  - i. In case the user's matriculation number cannot be found in the database, the web page in Figure 8 is displayed.

Contraction (Contraction)	÷.		Internet Street Street Street	10mds		1						A TO BE A
6 ftaføls semilit		27.5 Q/A0ca0	A & C. Mad & C Allake	1991						me law	- 14	/ I Dr assor C
15.Bernyich			50-mm	pheSyste	e 11	έ.						duse
Metore Sciences I	Search Result by make No	Serr Jamai	ien:	Nel: Terri	ome enternori							
_						itarra (	H	nę:	м	2.4		
Figure 6			10.0444									

 Example: Type the matriculation number and click search result (Figure 7) Figure 8

ii. In the case where user matriculation number is found in the database, it displayed the web page of all the page of all the results tag with the matriculation number (Figure 9)



E diversion	nik	**				125 3, 118	kr.	2 6 D' +	t Balantining	2 (Alternative Section 2.1.1.2)
i tana Tanat, politicat Tan	Wei Ser	COME Death, Last N								Flæse sign in
	34	Didint d	:tat	Gart .	itai :	Gasi	Lants			82.00
		20062401			38 34	-004 -002				Palant
		CONCERCT -				69011				
and the second se									Figure 11	

iii.

#### Figure 9

#### ADMINISTRATOR LOGIN:

 To login in as an administrator, click on Admin Login at the right corner of the web page (Figure 10) This web page collects the user/login name and the password, which is hidden from view by using special characters as the password (Figure 12).

			100000 BCA.9/7				1. P22 #20
			• Shake Verifier In		(66)	121/940m	5 1 0- UN- 0
e then	11	0- +JA- D					
FS4 Henenetyha Sych	n de se	-		f.			
ener Soltan Sol Lu Lu Lu Lu	RSA Homomorphic Encryption System Minimum Visitarian Statement Statement Statement of the All Statement Statement Statement Statement Statem Statement Statement State			Pleas Inter Inter Inter	e sign in		
tanie X-			Figure 12				

#### Figure 10

ii. Immediately the Admin Login is pressed, the web page below will be displayed (Figure 11), requiring the administrator to supply user name and password as a measure of security to authenticates the authorized user. iv. In the case the user enters invalid information, which cannot be validated by the mySQL server, the user will be denied access to the administrator web page. The web page in Figure 13 will still be displayed, requesting for correct user's name and password.





Figure 13

v. Once the password is correct, the web page below is displayed (Figure 14).



Figure 14

vi. To upload, click browse to bring in the file from a CSV file then click upload to send to the database (Figure 15)



Figure 15

vii. Figure 16 below shows how the result was stored in database. During the uploading, the software encrypts the students' result before it was stored in the database.

F.Salutinian	way to all the it to	A REPORT OF			11116	d bariner	F. B. D. Sole D
AND AND A	pater juste pare game						
1000 1013	7 Sergevi 2-4	raa Qaynal ADitrae					elettromente i ben
	iter i	nai dalipto and 3 o salatitati dali	12.0	88	Ng unte	9	
	1.7-1.000		1.04	1000	141	100	
	5 / X 100	CHARLES 1929-24 194	21094	110216	2010 CO.	W HERITS	
	C.P. N	ROAD STORE OF THE	21094	27348	DEMISTR	AN ADDRESS OF	
	5./X 1m	THE REPORT OF THE	210000	10101	70000 m	al traines	
	12X 100		2210045	W	TANKER	ALCOHOL: N	
	52X 10	THE R STATUTO AND A	24004	334067	2010/01/1	and the second second	
	127110	THE REPORT OF	222234	373548	DESIGNATION OF	IN ISSUES	
	1 / X 1 HD	CONTRACTOR NO.	21091	STYLE	200011	NUMBER OF STREET	
	27. 1.80	COMPANY OF THE OWNER.	21391	<b>BOOM</b>	THE OWNER OF	2761523 IN	
	DZX FR	10441072281818	26005	DOM: N	2006111	WHERE W	
	CAN NOT	THE R. LEWIS CO., LANSING MICH.	20005	201314	四根長!	NO DECEMBER	
	127 100	10414 (03599) 134	210385	<b>DINES</b>	TRENT 1	IN HISTORY	
	278 = M	THE R STREET WE	2122005	100	JEHICT .	IN SCOTHES.	
	578 D H	10100-00200-008	22000	175301	20003	N ICONS	
	CZX IN	10010320301140	22/00095	201216	NHEL	W IDINE	
	57X 840	1011010-00110-001	2004	101941	201617.1	N 1734KB	
	DZXXX	CONTRACTOR OF STREET, STATE	2110165	STATES	10000110	NI STOOTS	
	TXX TH	101010-0070201-0070	20005	1000	19930-011	ALCOHOL: N	
	1- 2 X - 3 H	COLUMN DESIGNATION.	212201	Distant.	COMPANY.	A 1011 (	

Figure 16



viii.

To view result, click view all result. The web page will display the decrypted result of the entire students (Figure 17)

Calegier.		41				hard and a second
+ Carded				111 Acres	Pa Dani	A B E total D
RSA Honomophe Sy	utern (	an				i per
-	We	come				
Contraction of the local division of the loc		(antere)	 (Det)	 lien	100	
		100000				
1991		140 Call				
0.000						
INCOME NO.		Comment.				
		10000				
198						
		Depart				
		SASSAS				



#### 5. CONCLUSION

Data to any organization is a valuable property. Security of sensitive data is always a big challenge for an organization at any level. In today's technological world, database is vulnerable to hosts of attacks. Due to these security issues the scope and promises of homomorphic encryption scheme to protect the data cannot be ignored. Researchers all over the world are taking great interest in recent times to develop homomorphic that can be deployed practically. This work described a complete designed scheme and evaluation of a provable secured students' data stored in database. It makes use of RSA homomorphic encryption schemes to encrypt and decrypt the data as it goes and comes out of the database to prevent information leakage and protects information integrity. Any attack to the database by unauthorized person will only access meaningless ciphertexts.

#### 6. **REFERENCES**

- Alfred J.Menezes, Paul C.Van OOrschot & Scott A. Vanstone (1996). Handbook of Applied Cryptography.
- Amichai Shulman (2006). Top ten database security threats. Imperia Application Defense Center.
- Bharath K. Samanthula, Yousef Elmehdwi, Gerry Howser & Sanjay Madria (2013). A secure data sharing and query processing framework via federation of cloud computing. Department of Computer Science, Missouri University of Science and Technology, 500 West 15<sup>th</sup> Street, Rolla, MO65401, United States.
- Cong Wang, Qian Wang, Kui Ren, & Wenjing Lou (2009). Ensuring Data Storage Security in Cloud Computing. In Proceedings of the 17th International Workshop on Quality of Service.
- Craig Gentry (2009). A Fully Homomorphic Encryption Scheme Ph.D Thesis Stanford University. Stanford.
- Damgard, I., Jurik M., & Nielsen, J. (2010, December 6). A Generalization of Paillier's Public-Key System with Applications to Electronic Voting. International Journal on Information Security (IJIS), Special Issues on Special Purpose Protocol, Vol 9, , pp. 371-385, Springer-Verlag, Heidelberg, Berlin, Germany.
- Dan Boneh, Eu-Jin Goh & Kobbi Nissim (2005). Evaluating 2-DNF formulas on ciphertexts. In Theory of Cryptography Conference, TCC'2005, volume 3378 of Lecture Notes in Computer Science, pages 325-341.
- Dongxi Liu D. Liu (2014). Securing Outsourced Databases in the Cloud. CSIRO Computational Informatics, Marsfield, NSW 2122, Australia e-mail: dongxi.liu@csiro.auBologna, Italy, pp. 139-149.
- Fontaine Caroline and Fabien Galand (2007). A Survey of Homomorphic Encryption for Nonspecialists. EURASIP Journal on Information Security, pages 1 to15.
- Groth, J. (2004). Rerandomizable and Replayable Adaptive Chosen Ciphertext Attack Secure Cryptosystems. In: Proceedings of the 1st Theory of Cryptography Conference (TCC'04), Lecture Notes in Computer Science (LNCS), Vol 2951, Springer-Verlag, pp. 152-170.
- Harshavardhan Kayarkar (2009). Classification of Various Security Techniques in Databases and



their Comparative Analysis. M.G.M's College of Engineering and Technology, Navi Mumbai, India.

- Iqra Basharat, Farooque Azam & Abdul Wahab Muzaffar (2012). Database Security and Encryption: A Survey Study. International Journal of Computer Applications (0975 – 888) Volume 47– No.12.
- Jayesh M. Tanna & Mital P. Sanghani (2012). Designing of a Symmetric Key Algorithm. International Journal of Application or innovation in Engineering & Management, (IJAIEM) Vol 1.
- John Fairfield (2004). Databases ANAESTHESIA AND INTENSIVE CARE MEDICINE. The Medicine Publishing Company Ltd.
- Joshi Ashay Mukundrao & Galande Prankish Vikram (2011). Enhancing Security in Cloud Computing, Information and Knowledge Management, International Institute for Science, Technology and Education (iiste) ISSN 2224-5758 (Paper) ISSN 2224-896X (Online) Vol 1, No.1
- Kadhem, H. Amagasa & T. Kitagawa, (2009). A Novel Framework for Database Security based on Mixed Cryptography. Internet and Web Applications and Services, Fourth International Conference Page(s): 163–170.
- Ken. (2011). Encryption Scheme. KenFinalpaper.
- Kerchoffs, A. (1883). La Cryptographic militaire(part 1). Journal des Sciences Militares, 5-38.
- Maha TEBAA, Saïd EL HAJJI & Abdellatif EL GHAZI (2012, July 4-6). Homomorphic Encryption Applied to the Cloud Computing Security. Proceedings of the World Congress on Engineering 2012 Vol I WCE London, U.K.
- MAIMUŢ Diana-Ștefania, PĂTRAȘCU Alecsandru & SIMION Emil (2012). Homomorphic Encryption Schemes and Applications for a Secure Digital World. Journal of Mobile, Embedded and Distributed Systems, vol. IV, no. 4

- Mark D. Ryan (2013). Cloud computing security: The scientific challenge and a survey of solutions. *The Journal of Systems and Software vol-86*.
- Morgan Kaufmann (2009). Relational Database Design and Implementation.
- MS. Parin V. Patel, Mr. Hitesh D. Patel & Prof. Pinal J. Patel (2013). A Secure Cloud using Homomorphic Encryption Scheme. International Journal of Computer Science Research & Technology (JJCSRT) Vol. 1 Issue 1.
- Naveed Islam, William Puech, Khizar Hayat, Robert Brouzet & Elsevier B. V. (2011). Application of homomorphism to secure image sharing.
- Neha Jain & Gurpreet Kaur(2012). Implementing DES Algorithm in Cloud for Data Security. VSRD International Journal of computer Science & Information Technology VSRD – IJSIT Vol. 2(4), 316 -321.
- Nitin Jain, Saibail K. Pal, & Dhananjay K. Upadhyay (2012). Implementation and Analysis of Homomorphic Encryption Schemes. International Journal on Cryptography and Information Security, 2, 1-18.
- Sachindra K. Chavan & Manoj L. Bangare (2013). Secure Data Storage in Cloud Service using RC5 Algorithm. International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-2, Issue-5.
- Youssef, G., Mouhcine, G., Zouhair, G., & Khalil, E. K. (2013). Security Internet Applications Using Homomorphic Encryption Schemes. *Journal of Theoretical and Applied Information Technology.*
- Zongkai Yang, Samba Sesay, Jingwen Chen & Du Xu (2004). A Secure Database Encryption Scheme. American Journal of Applied Sciences 1 (4): 327-331, 2004 ISSN 1546-9239 © Asian Network for Scientific Information



### TOWARDS AN ADAPTIVE AND SCALABLE ACCESS CONTROL MODELFOR A CLOUD-BASED ENVIRONMENT

#### <sup>1</sup>N.A Azeez,<sup>2</sup>O.B Okunoye, <sup>3</sup>F.AOladeji,<sup>4</sup>E.OEdafeadjeke

#### <sup>1,2,3,4</sup>Department of Computer Sciences, University of Lagos, Nigeria

<sup>1</sup>nurayhn@yahoo.ca, <sup>2</sup>bokunoye@unilag.edu.ng, <sup>3</sup>foladeji@unilag.edu.ng, <sup>4</sup>edafe4all@yahoo.com

#### ABSTRACT

This paper reviews previous access control models and proposes an adaptive task role-based access control model with federated identity as a gatekeeper for cloud-based services so that fine-grain access can be achieved. The motivation behind this architecture stems from the fact that cloud computing, though a growing field in information technology is not widely being implemented by enterprises due to security challenges. Information Technology strategists and decision-makers need to balance the gains of cloud computing, be it on the public or private model with the security needed to safeguard information assets and ensure regulatory compliance. Hence, this means one security policy model cannot cater for every IT infrastructural design model. A quantitative methodology is used to elicit access scenarios on the issuance of security tokens for task and processes. A policy-based markup language is being used to implement the model. This model can be applied in a public, private or hybrid cloud environment.

Keywords: model, security, architecture, markup language, identity, MAC, DAC

#### 1.0 INTRODUCTION

Cloud computing is considered one of the most dominant paradigms in the Information Technology (IT) industry today. It offers new cost effective services on-demand such as Software as a Service (SaaS), Infrastructure as a Service (IaaS) and Platform as a Service (PaaS). This form of distributed system offers solutions to the challenges presented by device proliferation, the exponential growth in volumes of data, and the accelerated changes demanded by the increasingly fast pace of businesses. However, with all of these services promising facilities and benefits, there are still a number of challenges associated with cloud computing. It may suffer from conventional distributed systems' security attacks such as malicious code (Viruses, Trojan Horses), back door, Man-in the Middle attack, Distributed Denial-Of-Service (DOS) attack [1], insecure application programming interface, abuse and infamous use of cloud computing, and malicious insiders (current or



former employee, contractor, or other business partner who has or had authorized access to organization's network, system, or data and intentionally exceeded or misused that access in a manner that can negatively affect the confidentiality, integrity, or availability of the organization's information information or systems)[2][3].

Cloud services could be inaccessible due to these attacks or security threats of increasing sophistication that represent a new generation of complex advanced persistent threats (APTs) such as Stuxnet and Hydrag (a piece of malware that uses a method of gathering system and network information initially, and then steals user names and passwords before collating all this information into a 'config' file)[4]. These bizarre names are among the latest in what has been a series of concerns regarding the unintended exposure of information and the exploitation of resources that depend on IT-managed services. It is an important and primary requirement for cloud service providers to ensure its services are fully usable and available at all times [1].

A security policy model is a concise statement of the protection properties that a system or generic type of system must have. An access control system is a pool of components and methods that determine the right admission to activities by valid users centered upon preconfigured access permissions and privileges defined in the access security policy [5].

Access control mechanism is the logical component that serves to receive the access request from the subject to decide and to enforce the access decision. Access control is a crucial requirement to ensure information security. It is also used as a critical system component to filter out inappropriate data, to provide customized views, and to improve effectiveness in information management [6].

#### 2.0 RELATED WORKS

The basic goal of any access control system is to restrict a user to exactly what it should be ableto do and as well protect information from unauthorized access. There is an array of methods, models. technologies and administrative capabilities used to propose and design access control systems. Thus, each access control system has its own attributes, methods and functions, which derive from either a policy or a set of policies [7]. Generally, there are three fundamental access control methods; Mandatory Access Control (MAC), Discretionary Access control (DAC) and Role-based Access Control (RBAC) [8].

#### 2.1 Mandatory Access Control (MAC) Model

Mandatory access control model is a set of rules by institution rather than individuals, used for authorization to read or write files in a system. It is loosely defined as any access control model that enforces security policies independent of user operations [9]. MAC allows for a system-wide access control policy which determines who is allowed to have access. Mandatory Access Control is usually associated with the 1973 Bell-LaPadula Model of multi-level security which uses only two properties, no-read-up and no-write-down, and the Biba integrity model.An individual is only given a read permission to a document D if (i) there is a reason to trust that person would not divulge the secrets in D (ii) if the contents of D are relevant to the person's job.Nevertheless, the reasons why this model is not suitable for cloud is its inability to supporttime and location constraint and dynamically activated access right for certain tasks.

#### 2.2 Discretionary Access Control

Discretionary access control model provide considerable flexibility by evolving many


protection decisions about object to the subject who created it. DAC empowers the subject to specify for each object what type of access can be made and by whom. This access is referred to as privilege.

DAC model is generally less secure than mandatory access control model, so it is used in environment that do not require a high level of protection. It is a very flexible and easier to implement than other models. Hence, DAC model is the most used access control model in commercial operating systems (WINDOWS UNIX base and operating systems). There are two ways to implement a discretionary access control model, this can be accomplished using identity based access control or by means of an access control matrix (Access Control List (ACL) or capabilities) [7]. However, DAC violates the integrity and confidentiality of objects by allowing users to pass rights to another user. It is not scalable for cloud computing.

#### 2.3 Role-Based Access Control Model

RBAC method can be seen as an independent component of access control, coexisting with MAC and DAC when appropriate. Early work on rolebased access control goes back to 1988, when Lochovsky and Woo defined roles and organized them into a hierarchy. RBAC has three principal rules, namely, Role assignment, Role authorization, and Permission authorization. In the first rule, if an individual has been assigned a role, he can use a permission. In the second rule, a person's active role must be authorized for that person. Finally, in permission authorization rule, a man can get permission only if the permission is authorized for the man's active role. This rule ensures that users could get only permission for which they are authorized.

The RBAC model has many benefits compared to DAC and MAC models, yet it has its own limitations and problems when it is deployed in the real-world scenario [7]. Defining the roles in a different context is difficult and it may result into large role

definition. Sometimes it produces more roles than users. Also, it does not separate tasks from roles and active responsibilities of staff is not supported.Moreover, it does not support the delegation principle. Although an extension of the RBAC model called Open Architecture for Secure Internetworking Services (OASIS) does support both the delegation and revocation principle[10].

#### 2.4 Existing Solutions with RBAC

The MIRE report of 2004 recommends focus on risk and proposes three guidelines;

- i. Measuring Risk
- ii. Establishing acceptable levels of risk
- iii. Ensuring that information is accessible at acceptable risk level

The concept of Risk-Adaptive Access Control Model (RAdAC) was introduced in 2009 wherein the limitations of traditional access control approach to sharing information and the main conceptual characteristic of RAdAC were discussed, but without articulation of a precise formal model.

RAdAC brings the concept of risk from the entire component used for access into the access control decision process. It also brings the concept of operational need (need-to-know) and situational factors (external or environmental conditions under which the access decision is being made) into the access control decision process. RAdAC principles of real-time, adaptable and risk-based; addresses real world scenarios where risk is an important factor in making access decisions [5][11].

In 2011, a paper by [12] proposes a more robust RBAC model to deal with the shortcomings of applying conventional RBAC model in an enterprise environment. It inherits many features from RBAC and distributed RBAC (dRBAC) such as dRBAC's domain [7].

Cloud optimized Role-based access control model seeks to solve some of these limitations, one of which the certification space of RBAC model



becomes complicated when facing a large number of users. Although conventional RBAC model provide the scalability to deal with users that are frequently changed, coRBAC combines the dRBAC's distributed authentication services together and expands the function of the certificate Authority.

In this improved model, the certificate authority also assigns the domain for the enterprise and organization with which they can manage their role and user in their own internal network. coRBAC also improves the overall efficiency of the access control system by adding hierarchical caches.Attribute is defined in different ways in the ABAC model. An attribute can be a subject's work start date, the location of a subject, the job role of a subject or all of them. Attributes are considered discrete therefore they may or may not be related to each other. The discrete values of attributes are compared against set of values by a policy decision point to determine whether to give or deny access.

Moreover, subjects need only to authenticate itself to the system and provide its attributes. The system does not require prior knowledge of any subject.ABAC systems can enable Risk-Adaptable access control solutions with risk values expressed as attributes.

In the ABAC model, subjects request access to objects, access control mechanism evaluates rules, subject attributes, object attributes, and environment conditions to compute a decision and Subjects are given access to objects if authorized.

ABAC solves the user role assignment problem which is present in RBAC and instead of focusing on the roles it focuses on the attributes of a user. It gives higher flexibility in a distributed, open, sharable and dynamic environment where the numbers of users are very high. Therefore ABAC is very flexible model for administration purposes and it works very well than RBAC.ABAC also support the global agreement of attributes so that attributes which are provided in one domain can be forward to the other domain at the point of domain to domain interaction.

However, due to the diversity of user information, complexity is increased therefore to solve this, it requires the central database having all attributes in the same format. On the other hand, if the multiple organizations decide common set of standardized attributes, this would raise the problem of low expressiveness for representing the subjects and objects, therefore it loses the advantages of the flexibility and dynamism of the ABAC functionality [15].

#### 3.0 OVERVIEW OF THE PROPOSED ACCESS CONTROL MODEL

The aim of this research is to review existing access control models and propose an adaptive and scalable access control model that is suitable for private public or hybrid cloud models. It is adaptive because attempt was made to set up a risk engine that keeps track of user's behaviours and takes certain action base on previous behaviours. Although the core of the model would rely heavily on the task-role based model, the use of the task and role concept makes it expandable to meet future needs or other cloud base infrastructural need. This feature can successfully deal with large number of users and administrators.

The proposed model facilitates the role and task principles as shown in Figure. 1. In the model, users are classified according to their actual jobs. Thus, users will be located on a security domain that relates to their roles. Every role within the model will be assigned a set of the most relevant and needed tasks for practicing this role. Every task will have a security classification for accessing the data or assets, and the exact permissions needed for accomplishing this task. A risk engine is utilized to deal with dynamic and random behaviours of users; it credits consumers according to their



access behaviours. A security token engine is also utilized for issuing security tags in semi (having the characteristics of both a private and public cloud environment) or untrusted environments and processes.

The model can secure access to data or assets by marking the data and assets with security labels. Any attempt to access the data has to ensure task classifications dominate the data or assets security labels. In this model, effort was made to utilize security tags in some circumstances according to the level of trust and security used within the environment. The security tag proposed consists of a user role, classification, permissions, the current location, issued time and a random unique number as shown in Figure 1.



#### 4.0 ACCESS CONTROL FRAMEWORK

The choice of a formal language for specifying a security policy is based on the capabilities, suitability and the richness of this language on one handand the requirements of the intended applicationon the other hand. In order to specify security policies, there is need to express norms, i.e. rules which say what must, may or must not be done.

1) The access control framework for the proposed model would have the following components:

- Domain (Do) is a set of Domains in a trusted environment
- Users (U) is a set of users identified by the IMS
- Roles (R) is a set of Roles
- Tasks (T) is a set of tasks.
- Sessions (S) is a set of sessions controlled by the IMS.
- Permissions (P) is a set of permissions.
- Data (D) is set of data.



### International Conference

- User Assignment (UA) is a subset of intersection between U and R.
- Role Assignment (RA) is a subset of intersection between R and T.
- Permission Assignment (PA) is a subset of intersection between P and T.
- Constraints (Con) is a set of constraints used in the system such as separation of duties and delegation.
- Classifications (Cla) is a set of security classifications utilized to classify tasks in the model.
- Sensitivity labels (SL) is a set of sensitivity labels used to restrict access to data according to its sensitivity.
- Security tokens (ST) is a set of security tokens
- 2) Every user U in the model can have an outlined number of roles {r<sub>1</sub>.....r<sub>n</sub>}, and every role can have a defined number of tasks {t<sub>1</sub>.....t<sub>n</sub>}. Every task t is assigned the exact needed permissions {p<sub>1</sub>.....p<sub>n</sub>} to accomplish its job, and a classification cla to access the targeted data or asset.
- 3) The relationships in the model are manyto-many except user-to-session (a user can have one session) and task-toclassification (every task has а classification) relationships which are oneto-one. For each task to be executed, a security token is generated and assigned to it.
  - ∀ u ∈ U → s<sub>t</sub> ∈ S : Where u cannot activate r outside s or Do but can use multiple roles in one domain.
  - $\forall t \in T \rightarrow c_1 \in C$

- ∀r ∈ R has a maximum number
   of authorized users and
   activation at one time
- U ⊆ U × R : a many to many mapping of user to role assignments
- *R* ⊆ *R* × *T* : a many to many mapping of role to task assignments
- P ⊆ P × T : a many to many mapping of permission to task assignments
- 4) This access control model supports the following constraints:
  - Least privilege principle: It is the idea of granting a subject *s* the only needed permissions *p* to accomplish its task t∀ *u* ∈ *U* → {*r*<sub>1</sub> .....*r*<sub>n</sub>}, ∀ *r* ∈ *R* →
    - $\{t_t \dots t_n\}, \forall t \in T \rightarrow$
    - $\{p_1 .... p_n\}$

U assigns  $r_{
m f}$  which uses  $t_{
m f}$ 

 $\mathbf{L}_{t}$  Only requires permission pto accomplish its actions

Delegation of capabilities: For flexibility and dynamism, users are allowed to collaborate to accomplish general tasks.However, delegation of tasks can happen only under the risk engine control, and between two users which have equivalent roles and workwithin the same domain. $u_t, u_j \in U, t_t \in T, r_t \in$ Ru<sub>r</sub> is assigned t<sub>r</sub> but cannot finish it. Hence, an admin can delegate  $t_i \ t_i \ u_j \leftrightarrow u_i, \ u_j \in r_i$ and in the same domain.

Separation of duties: The goal is to partition tasks and permissions linked to roles in other to prevent



## 12<sup>th</sup>

### International Conference

assigning too much authority to one user. It also ensures roles and interest do not conflict.  $u_t \in$ U;  $r_t$ ,  $r_j \in R$ ;  $t_t$ ,  $t_j \in T$ ; where  $r_t \neq r_j$  as  $t_t \neq$  $t_j$   $u_t$  activates  $r_t$  as  $r_j \leftrightarrow r_t \cap r_j =$  $\emptyset$   $r_t$  is assigned to  $u_t$ ,  $u_t$  activates  $t_t$  as  $t_t \leftrightarrow t_t \cap$ 

$$t_1 = \emptyset$$

5) Classifications are given to tasks when they try to access data or a system's assets. They have to control an object's security label before accessing it.

 $\forall t \in T \to c_{-t} \in C \quad , \quad \forall d \in D \to s \in S$ 

6) Sensitivity labels are attached to data or a system's assets to restrict access according to sensitivity degree.In this model, a hierarchically ordered set of security labels isutilized, which are highly

8) 
$$\forall p \in P \rightarrow \{\overline{R} (R), W (W), E \ D (D)\}$$

9) Every task is distinctive and can either be executed alone or requires one or more actions to be triggered to fulfill its purpose. Nevertheless, situations arise where sequences of tasks are used to complete a job or a task relies on another task to finish its work. In such cases, the task that is accessing the data or assets directly has to allot a security token, in order to prevent privilege leakages or denied access. There is no delegation of security tokens in this model, that is, an issued token for a task being delegated to another task.

#### 5.0 TASKS AND SECURITY TOKENS

The following scenarios illustrate how a security token is employed by a task accessing data:

 $L \xrightarrow{s} d$ .....(1)

classified (HC) >Classified (CL) > Confidential (C) > Unclassified (U).

 Security tokens will be used when a task tries to access data from another domain or employed processes.

 $\forall s \in S \\ \rightarrow \{r, c, u_c \mid l, r, R \}$ 

Wherest is a security tag assigned by the security token engine RUN (Random Unique Number) is used to ensure every security token is unique

(E) a

A security tokenst is generated for every task that tries to access data base on the task classification.

 $t_{\rm f} \rightarrow t_{\rm j} \stackrel{s}{\longrightarrow} d$ .....(2)

A task attempting to access data indirectly or through another task, a security token is generated by the actual task accessing the data.

$$t_i \stackrel{s_i}{\to} p \in p \qquad \stackrel{s_i}{\to} d.....(3)$$

Any process that requires access to a particular data must request a security token which is generated by the task that owns the process.

$$t_i \rightarrow t_j \stackrel{s_j}{\rightarrow} p \in p \qquad \stackrel{s_j}{\rightarrow} d.....(4)$$

Any foreign process that needs to access data have to get a security token from the actual task attempting to access data and employed it. This is to prevent leakage of privileges.



### International Conference

#### 6.0 SECURITY LEVELS OF THE MODEL

- 1) In a secure environment and trusted processes: there is no need of using security tokens.
- 2) Semi secure environment or trusted processes: In this level, security tokens are used and have validation times.
- 3) Unsecured environment and untrusted processes: It uses the same principle as a semi secure environment; although a security token is issued for every access to data or assets so as toprevent any chance of reusing the security token or cheating by either tasks or processes.

### 7.0 ANALYSIS OF THE PROPOSED ARCHITECTURE

Although Task-Role based access control is used as a backbone to the proposed model because of its simplicity and flexibility in the configuration, it needs to be combined with an identity management system, developed and extended for cloud computing to support the delegation principle, dynamic and random behaviours of users, local and global access. The proposed system as shown in Figure 1 can deal with the previous apprehensions by employing the following concepts:

- In organizations and enterprises, roles are used to specify a user's responsibility and control access to resources. A user responsibility takes precedence over whom the useris. Hence roles are used since it gives organizations the ability to impose their constraints.
- 2. Each user that has been authenticated by the identity management system is assigned a role; subsequently, roles are assigned tasks that have permissions. Every role within the system will have a set of the most relevant and needed tasks for practicing this role. These permissions are assigned to roles with respect to their tasks and the given permissions

dynamically change according to the task at hand.

- 3. The current task or process activates or deactivates permissions. Permissions are dynamically changed according to users' behaviours.
- Resources and information flow are controlled using classifications and security labels. Sensitivity labels are used to watermark data internally according to their sensitivity and value, which are highly classified (HC), Classified (CL), Confidential (C), and Unclassified (U).
- 5. The model supports time and location constraints, least privilege principle, separation of duties either static or dynamic and delegation of capabilities.
- 6. Based on a user's previous behaviours, this model uses a risk engine to deal with some security concerns by controlling and crediting users based on how they comply with security and access policies.
- 7. A security token engine is used for generating security tags for tasks, which might be passed to employed processes. The security tags are utilized in untrusted environments to control access to a system's resources.

#### 8.0 CONCLUSION

In this paper, effort has been made to review other access control models and proposed a unique access control architectural framework for cloud computing. It is our believe that the proposed model can fulfill the access control requirements in cloud environment. Its backbone is based on the role and task principles which make assigning privileges dynamic and easy. In this model, each user that has been authenticated by the identity management system is assigned a role; subsequently, roles are assigned tasks that have



permissions. Every role within the system will have a set of the most relevant and needed tasks for practicing this role. These permissions are assigned to roles with respect to their tasks and the given permissions dynamically change according to the task in hand. The current task or process activate or deactivate permissions. Permissions are dynamically changed according to users' behaviours. The full scale implementation of this model with CloudSim will reveal the efficiency of the proposed architecture.

#### 9.0 REFERENCES

- [1 S Yu, C Wang, K Ren, and W Lou. (2010) IEEE INFOCOM. [Online]. http://dx.doi.org/10.1109/INFCOM.2010.54621 74
- [2 Dan W Hubbard and Michael Z Sutton. Top threats to cloud computing. [Online]. https://cloudsecurityalliance.org/topthreats/
- [3 Carnegie Mellon University. (2015, April) Cert Software Engineering Institute. [Online]. https://www.cert.org/insider-threat/
- [4 Symantec and VMware, Securing the Cloud for the Enterprise, 2011.
- [5 R Anderson, Security Engineering: a guide to building dependable distributed systems.: John Wiley & Sons, 2010.
- [6 Xuan Hung Le, Terry Doll, Monica Barbosu, Amneris Luque, and Dongwen Wang, "An enhancement of the Role-Based Access Control model to facilitate," elsevier, pp. 1084-1107, 2014.
- [7 Younis A Younis , Kashif Kifayat, and Madjid Merabti, "An access control model for cloud computing," Elsevier, pp. 45-60, 2014.
- [8 Junqi Duan, Deyun Gao , Chuan Heng Foh, and Hongke Zhang, "TC-BAC: A trust and centrality degree based access control," Elsevier, pp. 2675-2692, 2013.
- [9 Ryan Ausanka-Crues. http://www.cs.hmc.edu.

#### [Online].

http://www.cs.hmc.edu/~mike/public\_html/c ourses/security/so6/projects/ryan.pdf

- [1 Andras Belokosztolszki, "Role-based access control policy," 2004.
- [1 S Oh and S Park , "Task role-based access control model," 2003.
- [1 Zhu Tianyi, Liu Weidong, and Song Jiaxing, "An efficient Role Based Access Control System for Cloud Computing," in 11th IEEE International Conference on Computer and Information Technology, 2011.
- [1 Eric Freudenthal, Tracy Pesin, Lawrence Port, Edward Keenan, and Vijay Karamcheti. www.cs.nyu.edu. [Online]. www.cs.nyu.edu/vijayk/papers/drbacicdcso2.pdf
- [1 Vincent C. Hu et al., "Guide to Attribute Based Access Control (ABAC) Definition and Considerations,".
- [1 Bokefode D Jayant., Ubale A Swapnaja, Apte S Sulabha, and Modani G Dattatray, "Analysis of DAC MAC RBAC Access Control based Models for Security," International Journal of Computer Applications, pp. 6-13, 2014.
- [1 Marcos A. P. Leandro, Tiago J. Nascimento, Daniel R. dos Santos, Carla M. Westphall, and Carlos B. Westphall, "Multi-Tenancy Authorization System with Federated Identity for Cloud-Based Environments Using Shibboleth," in ICN 2012 : The Eleventh International Conference on Networks.
- [1 P Cheng and P Rohatgi , "Fuzzy multi-level security: an experiment on quantified riskadaptive access control," in IEEE Symposium on Security and Privacy, 2007.
- [1 V Suhendra. (2011) Secure Technology. [Online]. http://www.springerlink.com/index/J3101024 2555W867.pdf
- [1 Y Demchenko, L Gommans, C De Laat, and A



Van Buuren, "Domain Based Access Control Model for Distributed Collaborative Applications," in The 2nd IEEE International Conference on e-Science and Grid Computing.

- [2 Anu Gopalakrishnan, Cloud Computing Identity Management, 2009.
- [2 Eghbal Ghazizadeh, Jamalul-lail Ab Manan, Mazdak Zamani, and Abolghasem Pashang, "A Survey on Security Issues of Federated Identity in the Cloud Computing," in IEEE 4th International Conference on Cloud Computing Technology and Science, 2012, pp. 562-565.
- [2 (2012) Fred B Schneider publication. [Online]. www.cs.cornell.edu/fbs/publication/chptr.D AC.pdf
- [2 Elliott D Bell and Leonard J LaPadula, "Secure Computer Systems: Mathematical Foundations," MITRE Technical Report , Research 1996.
- [2 Keleta Yared, J H Eloff, and H S Venter, "Proposing a Secure XACML architecture ensuring privacy and trust," Pretoria,.



International Conference

#### PREDICTION OF PEADIATRIC HIV/AIDS SURVIVAL IN NIGERIA USING NAIVE BAYE'S APPROACH

#### O. Agbelusi<sup>1</sup>, T. A. Aladekomo<sup>2</sup>, and \*P. A Idowu<sup>3</sup>

<sup>1</sup>Department of Computer Science, Rufus Giwa Polytechnic, Owo, Ondo State <sup>2</sup>Dept. of Pediatric and Child Health, Obafemi Awolowo University, Ile-Ife, Nigeria <sup>3</sup>Dept. of Computer Science & Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria

<sup>1</sup>tola52001@yahoo.com, <sup>2</sup>aladekomotheo@yahoo.com, <sup>3</sup>paidowu1@yahoo.com

#### ABSTRACT

In this paper we developed predictive model for survival of pediatric HIV/AIDS patients who were receiving antiretroviral drug in Nigeria based on identified variables. Interviews were conducted with the virologists and pediatricians. Pediatric HIV/AIDS patients' data (216) were also collected from two tertiary health institutions in Nigeria, preprocessed and the 10-fold cross validation technique was used to partition the datasets into training and testing data. The model was developed using supervised learning techniques (Naïve Bayes' classifiers) and the Waikato Environment for Knowledge Analysis (WEKA) was used to simulate the model in which CD4 count, viral load, opportunistic infections and nutritional status were used as the independent variables for the prediction. The result shows that Naïve Bayes' classifier was suitable in carrying out the task of forecasting the survival of paediatric HIV/AIDS patients with an accuracy of 81.02% (175 correct classifications out of 216), the mean absolute error rate was 0.2025, 0.2920 for the root mean square error and 40.92% for the relative absolute error. The ROC area for the model was also 0.993 showing that the level of bias was very low (0.007) 81.0185. The validation was done by comparing the developed model with the historical data from the two selected health tertiary institutions in Nigeria.

**Keywords:** Paediatric, Acquired Immunodeficiency Disease Syndrome, Human Immunodeficiency Virus, Antiretroviral Drug, Predictive Model.

#### 1. INTRODUCTION

Epidemic diseases have highly destructive effects around the world and these diseases have affected both developed and developing nations. Disease epidemics are common in developing nations especially in Sub Saharan Africa in which Human Immunodeficiency Virus /Aquired Imunodeficiency Disease Syndrone (HIV/AIDS) is the most serious of all (Idowu *et al.*, 2009). HIV is one of the world's most serious health and development challenges (Henry, 2013). It is a type of virus called a retrovirus which infects humans when it comes in contact with tissues such as those that line the vagina, anal area, mouth, eyes or through a break in the skin (Eric and Daria, 2012), while Acquired Immunodeficiency Syndrome (AIDS) is the advanced stage of the retroviral infection that swept through sub-Saharan Africa with venom (Hoe, 2011; Idowu, 2012).

Globally, HIV continues to be a very serious health issue facing the world (Ojunga *et al.*, 2014). About



34 million (31.4 million–35.9 million) people were living with HIV at the end of 2011 and an estimated 0.8% of adults aged 15-49 years worldwide are living with the virus, although the burden of the epidemic continues to vary considerably between countries and regions (UNAIDS, 2012).

In Sub-Saharan Africa, roughly 25 million people were living with HIV in 2012, accounting for nearly 70 percent of the global total. The epidemic has had widespread social and economic consequences, not only in the health sector but also in education, industry and the wider economy (WHO, 2012; Shearer, 2000). The epidemic has heavy impact on education, school attendance drops as children become sick or return home to look after affected family members (Picat et al., 2013). Moreover, Sub-Saharan Africa remains most severely affected, with nearly 1 in every 20 adults (4.9%) living with HIV which accounts for 69% of the people living with HIV worldwide. Although, the regional prevalence of HIV infection is nearly 25 times higher in sub-Saharan Africa than in Asia, almost 5 million people are living with HIV in South, South-East and East Asia combined and sub-Saharan Africa region is the most heavily affected region follow by the Caribbean, Eastern Europe and Central Asia, where 1.0% of adults were living with HIV as at 2011 (Henry, 2013).

In addition, Nigeria is the most populous nation in Africa with an estimated population of over 160 million people. Government reports claim that over 300,000 Nigerians die yearly of complications arising from AIDS. Nigeria has the highest HIV populations in Africa with 5.7 million infected people. It is estimated that over 200,000 people die yearly in Nigeria as a result of HIV/AIDS (Nigerian Bulletin, 2014).

At present, there is no cure for HIV but it is being managed with antiretroviral drugs (ARV). There is optimal combination of ARV which is known as Highly Active Antiretroviral drug (HAART) (Rosma *et al.*, 2012; Kama and Prem, 2013). Antiretroviral therapy is the mechanism of treating retroviral infections with drugs. The drugs do not kill the virus but they slow down the growth of the virus (Ojunga *et al.*, 2014). HAART refers to the use of combinations of various antiretroviral drugs with different mechanisms of action to treat HIV.

The epidemic of HIV/AIDS affects two classes of people: the pediatric and the non-pediatric individual. The non-pediatric patients are patients above 15 years of age while the pediatric patients who form the main target of this research are patients whose age is less than 15 years (Sanjel, 2009).

There are four distinct stages of HIV infection which includes: the primary HIV infection stage or clinical stage 1 which involves asymptomatic and acute retroviral syndrome; clinically asymptomatic stage or clinical stage 2 which involves moderate and unexplained weight loss (<10% of presumed or measured body weight; symptomatic stage or clinical stage 3 is also a conditions where a presumptive diagnosis could be made on the basis of clinical signs or simple investigations like unexplained chronic diarrhea for longer than one month, unexplained persistent fever (intermittent or constant for longer than one month) and progression or Clinical stage 4 that is a condition where a presumptive diagnosis can be made on the basis of clinical signs or simple investigations like HIV wasting syndrome, pneumocystis pneumonia, recurrent severe or radiological bacterial pneumonia etc. Patient at this stage is in a condition where confirmatory diagnostic testing is necessary (Avert, 2014; Centre for Disease Control, 2011).

There are different modes of transmission of this virus one of which is mother-to-child transmission. Here, about nine out of ten children exposed are infected with HIV during pregnancy, labour, delivery or while breastfeeding (UNAIDS, 2010). Without treatment, 15-30 percent of babies born to HIV positive women are infected with the virus during pregnancy and delivery and a further 5-20 percent are also infected through breastfeeding (WHO, 2006). In high-income countries, preventive measures are undertaken to ensure that the transmission of HIV from mother-to-child is relatively rare and in cases where it occurs, a range of treatment options are undertaken so that the child can survive into adulthood. Blood transfusion



is another route in which HIV infection can occur in medical setting (Mira and Denis, 2010).

HIV epidemic in Nigeria is one of the incurable deadly diseases and it varies widely by region (Petros *et al.*, 2010). The impact of HIV/AIDS is pervasive and far-reaching, affecting individuals and communities not only psychologically but also economically and socially. Families lose their most productive members to this disease, leaving children and elderly people without means of support (UNAIDS, 2013).

Despite the state of this deadly disease in Nigeria and most especially among children, there is no existing model in which survival of infected patient can be predicted. Therefore, this paper present the development of survival model among pediatric HIV/AIDS patients in South Western Nigeria and the objectives were to identify survival variables for HIV/AIDS pediatric patients in the South Western Nigeria and formulate survival predictive models based on variables identified (CD4 count, viral load, nutritional status and opportunistic infection) from the interview conducted with the virologist and pediatrician at the study area.

#### 2. RELATED WORKS

Researchers had worked on the prediction of HIV/AID prediction using different types of variables like CD4 count, CD8. Some of the researcher and the result of their finding are:

In 1996 Lucia developed a model to predict survival of HIV/AIDS using sequential and standard neural networks. The aim of the study was to produce a model of disease progression in AIDS using sequential neural network and compare the model's accuracy with that of a model constructed using only standard neural networks based on demographic and socioeconomic variables. The strength of the study was that sequential neural networks could discriminate patients who die and patients who survive more accurately than the standard neural networks. The weakness of the study is that only demographic and socioeconomic variables were used, which is not sufficient to predict survival. CD4 count, viral load, opportunistic infections and nutritional status would be enough to predict survival accurately in this paper.

Brain *et al* (2006) applied neural network in the prediction of HIV status of an individual based on demographic and socioeconomic characteristics. The aim of the study was to use supervised learning to train neural networks, to classify the HIV status of an individual given certain demographic factors. The strength of the study is that the neural networks used for prediction has high predictive capability. The weakness is that demographic and socioeconomic characteristics are not enough to accurately predict survival status.

Torsak et al (2009) used highly active antiretroviral therapy to predict survival among HIV-infected children in Asian countries. The aim of this research is to conduct a general review of Paediatric ART effectiveness in Asian countries using Kaplan-Meier survival analysis to estimate survival time probability, after the introduction of ART and Cox proportional hazard model was used for multivariate analysis. The strength of this research is that there were beneficial outcomes of first-line antiretroviral therapy for HIV infected children in Asian countries. The weaknesses were (1) limited information about the management of children who failed first line NHRT I regimen and (2) There was need to improve access to early diagnostic testing and treatment in infancy. CD4 count was the only predictive variable used in this study which is not enough to determine the survival of infected HIV patients CD4, viral load, opportunistic infection and nutritional status were added to the existing predictive factor in this thesis.

Sameem et al (2010) used Classification and Regression Tree (CART) for the Prediction of Survival of Aids Patients receiving antiretroviral therapy in Malaysia. The aim is to investigate the use of CART as a tool for prediction of AIDS survival using CD4, CD8, Viral Load and Weight as predictor variables. The strength of the research is that the potential treatment methods and monitoring the progress of treatment of AIDS patients could be determined with the approach experimented and the results obtained. Fewer variables were considered for the prediction of survival of AIDS



patients and data limitation is also a constraint in the study. Also opportunistic infections and nutritional status that are very important for HIV/AIDS prediction were not used in the research. CD4, CD8, viral load and weight are not enough factors to predict survival.

Dalton *et al* (2010) carried out a prospective cohort on the Predictors of mortality in HIV-1 infected children on antiretroviral therapy in Kenya. The aim of this work was to carry out a study on early mortality following highly active antiretroviral therapy (HAART) using Cox proportional hazard model to determine the baseline characteristics associated with mortality and Kaplan-Meier method to estimate the probability of survival. The study shows that Low baseline hemoglobin was an independent risk factor for death. The weakness is that only hemoglobin was used as predictor variable in this research which is not enough to determine survival rate.

Heamoglobin is not enough to predict the survival of pediatric infected HIV patients but in this thesis, CD4, CD8 and viral load, opportunistic infection and nutritional status were used as predictive factors.

Rosma et al (2012) developed a predictive model for AIDS Survival using Data Mining Approach. The aim of the research was to describe the feasibility of applying data mining technique to predict the survival of HIV/AIDS. An adaptive fuzzy regression technique, FuReA, was used to predict the length of survival of AIDS patients based on their CD4, CD8 and viral load counts. The strength of the research is that CD4, CD8 and viral load counts were used because the authors believed that predictors / markers are appropriate for predicting AIDS survival due to the high accuracies demonstrated by Fuzzy regression analysis (FUREA). The weakness is that fuzzy neural network prediction results on AIDS survival could not be made possible because of data limitation. This is because Fuzzy neural network requires the use of large volume of data for prediction. Opportunistic infections and Nutritional status are important predictor variables together with CD4, CD8 and viral load counts that can be used to predict the survival of HIV/AIDS patients.

Nurilign *et al* (2014) carried out a retrospective cohort study on the survival status of HIV positive adult on antiretroviral treatment in Debre Markos Referral Hospital, Northwest Ethiopia. The aim of this study was to determine survival status and associated factors among HIV positive adult on antiretroviral treatment using Kaplan Meier to estimate survival and Cox regression for the analysis. Lost, drop out, transfer out and transfer in patients with high risk of death were excluded so as not to under estimate mortality of infected patients. Secondary data in which some important variables were not documented well were used and many opportunistic infections were presumed diagnosis. Pre-processing exercise was carried out in order to remove incomplete data before classification and prediction of survival took place and more independent variables were also used in this study. CD4, CD8, viral load, opportunistic infection and nutritional status were in this thesis as predictive factors.

Ojunga et al (2014) applied logistic regression in modelling of survival chances of HIV-positive patients under highly active antiretroviral therapy (HAART) in Nyakach District, Kenya. The aim of this study was to outline the various social and economic factors affecting survival of HIV patients under highly active antiretroviral therapy (HAART). The study was expected to provide suitable model for predicting the chances of survival among the HIV positives attending ART clinic in Nyakachi District and also provide information for policy makers on the factors affecting survival of HIV positive ARVs. The strength shows that the survival of infected patient under study can be improved if access to socio-economic factors their is considered. The outcome may only be obtained in services that have smaller numbers of patients. Socioeconomic factors are not enough to predict survival as CD4, CD8, viral load, opportunistic infections and nutritional status were added to the existing study in this paper as predictive factors.

#### 3. METHODS

Extensive review of literature on related areas in HIV/AIDS survival prediction was carried out and interview was conducted with virologist and



Paediatrician in order to identify required survival variables for HIV/AIDS. Predictive model was developed using supervised learning techniques (Naïve Bayes classifier) and the Waikato Environment for Knowledge Analysis (WEKA) was used to simulate the model. The performance of this model was tested and we found out that it can predict HIV/AIDS survival with an accuracy of 81.015% based on selected dependent variables. The validation of the models was done by comparing the historical data collected from The Federal Medical Centre, Ondo State and Obafemi Awolowo University Teaching Hospital Complex, Ile- Ife with predicted value.

#### 4. RESULT AND DISCUSSION

#### 4.1 Result

Mathematically, the Naïve Bayes' Classification is expressed as follows:

$$P(X|Ci) = \prod_{k=1}^{n} P(X_k|Ci)$$

- **Prior distribution**  $P(X|C_i) = P(x_1|Ci) * P(x_2|Ci) * P(x_3|Ci) * \dots * P(x_k|Ci)$ 

The classification of each training data was performed via the implementation of the Naïve Bayes' classification algorithm which calculates the probability and manipulates them into the necessary results. A typical demonstration of how this is achieved is shown as follows:

For HIV survival model, the variables used are: CD4 count, viral load, Opportunistic Infections and Nutritional Status and they are represented as X (input value) and the output class by C.

 $X = (x_1 = "CD4 \text{ status} = value", x_2 = "viral load = value", x_3 = "nutritional status = value", x_4 = "opportunistic infection = value")$ 

The probability that the output of the survival is **YES** =  $P(X|YES) = P(x_1|YES)^* P(x_2|YES)^* P(x_3|YES*P(x_4|YES))^*$ 

P(x<sub>1</sub>|YES) = probability of the value of CD4 status (High/Low) when survival is YES

 $P(x_2|YES) = probability of the value of Viral Load (High/Low) when survival is YES$ 

 $P(x_3|YES) =$  probability of the value of nutritional status (High/Low) when survival is YES

 $P(x_4|YES) = probability of the value of opportunistic infection (Yes/No) when survival is YES$ 

The probability that the output of the survival is **NO**  $P(X|NO) = P(x_1|NO)*P(x_2|NO)*P(x_3|NO)*P(x_4|NO)$ 

 $P(x_1|NO) = probability of the value of CD4 status (High/Low) when survival is NO$ 

 $P(x_2|NO) = probability of the value of Viral Load (High/Low) when survival is NO$ 

 $P(x_3|NO) = probability$  of the value of nutritional status(High/Low) when survival is NO

 $P(x_4|NO) = probability of the value of opportunistic infection (Yes/No) when survival is NO$ 

Therefore, the output value of survival is the probability **max** [P(x|YES),  $P(x_1|NO)$ ].

Hence,  $C_{i}$  is the required Output/Survival (YES or NO)



Figure 1: Naïve Bayes' Prediction of all 216 Data





Figure 2: Confusion Matrix of Naïve Bayes' Classification for all 216 data.

#### 4.2 Discussion

From the results of the analysis made on the dataset using Naïve Bayes' classification in developing the predictive model of HIV/AIDS survival for Paediatric patients; the following were observed from Figures 1 and 2.

In Figure 2, out of the 216 data collected from FMC. Owo and OAUTHC, Ile-Ife, after using the Naïve Bayes' Classifier to train the data and validate the model developed using 10-fold cross validation, it was discovered that the naïve bayes' prediction model made 95 correct classification and 80 incorrect classifications of the output of the survival. Figure 1 above shows the graph of the results of the classification made by the Naive Bayes' prediction model; blue crosses identify YES and red crosses identify NO while the boxes show misclassifications. From the results of the analysis made on the dataset using Naïve Bayes' classification in developing the predictive model of AIDS survival for Paediatric patients; the following were discovered: that out of 97 dataset which had survival of YES, 95 were classified as YES and 2 were classified as no and of the 119 dataset which had a survival of NO, 39 were classified as YES and 80 were classified as NO. The model also produced an accuracy of 81.02% in prediction. The above was also represents with a confusion matrix in figure 2 above.

From the Naïve Bayes' prediction model, the True Positive (TP) rate values were 1.00, 0.659 and 0.979; False Positive (FP) rate values were 0.359, 0.026 and 0.328 and precision rate were 0.678, 0.964 and 0.709 for the FMC, OAUTHC and both data sources respectively. The area under the ROC curve were 0.995, 0.985 and 0.993 for FMC, OAUTHC and both data sources respectively showing that the Naïve Bayes' Model has a low rate of bias; an average of 0.01 (1% of results).

#### 5. VALIDATION OF THE MODEL

Validation is the task of demonstrating that the model is a reasonable representation of the actual

system. It reproduces system behaviour with enough fidelity to satisfy analysis objectives. In table 1 below, the validation of the model was done by comparing the values of the developed survival predictive model with the actual values collected from the two health institutions (historical data) and it was discovered that the values from the decision tree which is the most efficient of the three model were almost the same as the actual values using all the input variables (CD4 count, Viral load, Opportunistic infections and Nutritional status).

Table 1: Validation of the Model with the Actual Data

Status	(nan	Multi Gonul Status	Opportunistic Intertion	Survival	Name:
High	1.64502	High	We Car	¥1.44	
Lanse-	High	Hards	Y		****
Lilgh	Low	Low	Y2.8	ne	Yes
High	LOW	High	Y	yes.	Yes
110.21	1.000	3 a 1 sec	WH N	100.0	¥mo.
1 H Th	1.0442	tuch	Vec	yes	240
Heght	Liow	High	Y	8425	¥-cos
High	Lines	Low	944	110	Yes
Lilen.	Casso	40mb	Vec	Ves.	. yes
Low	Ligh	1.039	And.	no	no
Law	Light	1 OW	00	no	80
Hegh	Liney	High	Y	1000	yes
Hugh	LOW	High	Y-a	3425	yes
Lillets	1000	1 mile	Vec	ves	yes.
LII.zh	0447	1.799	Vec	no	Ves
High	LOW	Hisis	W-Carl	V425	¥-cs
High	#188.00 F	High	W-1-+	648 MA	Wars
(IIzh	Down	1.539	Vec	no	Ves
111 ah	Low	Ligh	Vec	yes	Ver
Hat the	Paase	1 500	Are in	() () () () () () () () () () () () () (	Water
Laiss	10 activ	The la	Y	100.0	Film
Lifets	LOW	tilgh	YES	Yes	Yes
Low	High	LUW	8	110	No
ru,th	0.00	1.5W	VPC	no	1000

#### 6. CONCLUSION

Child mortality is a factor that can be associated with the well-being of a population and taken as one of the development indicators of health and socioeconomic status in any country. HIV/AIDS epidemic has devastated many individuals, families and communities. Therefore, in order to reduce child mortality which is one of the important millennium goals, there is need to have effective and efficient model that can be used to forecast the survival of Paediatric HIV/AIDS patients in South Western Nigeria. It will also help individuals, NGO and the government to make adjustments in areas where these affected children are suffering. The naive baye's predictive model serves as an effective model from the analysis above and will be recommended for use to predict Pediatric HIV/AIDS patients survival in order to justify results collected



from the two health institutions (FMC, Qwo and OAUTHC, Ile-Ife). Pediatric HIV/AIDS patients survival prediction depends on some major factors as shown in the result of this paper and the factors serves as contributing factors to the patient's status

#### 7. ACKNOWLEDGEMENT

We acknowledge all the referees, and those that contributed to the success of this work.

#### 8. REFERENCES

Avert, 2014. Stages of HIV Infection. Available from <u>www.averts.org/stages-hiv-infection.htm</u> [Accessed 28 July, 2014]

Brain, et al., 2006. Prediction of HIV status from Demographic Data using Neural Networks. IEEE conference on Systems Man and Cybernetics. Taipei, Taiwan.

Centre for Disease Control, 2011. HIV classification: CDC and WHO staging system. Available from www.hab.hrsa.gov/deliverhivaidscare/clinicalguide1 1/cg- 205\_hiv\_classification.html. [Accessed 4th August, 2014].

Dalton, C., et al., 2010. Predictors of Mortality in HIV-1Infected Children on Antiretroviral Therapy in Kenya: A Prospective Cohort. *Journal of Pediatrics*. 10(33). pp 1471-2431.

Eric, J. and Daria, J., 2012. HIV-1 Antiretroviral Drug Therapy. Journal of Cold Spring Harb Perspect Medical. Vol. 2, No 4. pp 23-45

Henry, K., 2013. The Global HIV/AIDS Epidemic. Available from <u>www.kff.org/global-health-</u>policy/fact.../the-global-hivaids-epidemic/. [Accessed 12 January, 2014].

Hoa, M., 2011. ART Adherence among People Living With HIV/Aids in North Vietnam, Queensland University of Technology, Brisbane Australia

Idowu, P., et al., 2009. Spatial Predictive Model for Malaria in Nigeria, Journal of Health Informatics in Developing Countries, Vol. 3, No 2. pp 31-36.

Idowu, P., 2012. Development of A Web-Based Geo-Spatial Environmental HealthTracking System for South Western Nigeria. Unpublished PhD Thesis Submitted To Department Of Computer Science and Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria.

Kama, K. and Prem, S., 2013. Utilization of Data Mining Techniques for Prediction and Diagnosis of Major Life Threatening Diseases Survivability-Review. International Journal of Scientific & Engineering Research. Vol 4, No 6. pp 923-932.

Lucia, O. , 1996. Sequential Use of Neural Network for Survival Prediction in AIDS. Available from <u>www.ncbi.nlm.nih.gov/pmc/articles/PMC2233186/</u>. [Accessed 4<sup>th</sup> August, 2014].

Mira, J. and Denis, A., 2010. The Cost-Effectiveness of Preventing Mother-to-Child Transmission of HIV in Low- and Middle-Income Countries: Systematic Review. *Journal of Bio Med Central*. Vol. 9, No 3. pp 12-19.

Nigerian Bulletin, 2014. World HIV/AIDS Day: 10 Facts about HIV/AIDS in Nigeria You Probably Didn't Know. Available from www.nigeriabulletin.com/threads/world- hiv- aidsday-10-fact-about-hiv-aids-in- nigeria-youprobably-didn't-know.24303/. [Accessed 28 July, 2014].

Nurilign, A., et al., (2014) Survival Status of HIV Positive Adults on Antiretroviral Treatment in Debre Markos Referral Hospital, Northwest Ethiopia: Retrospective Cohort Study. *The Pan African Medical Journal*. Vol. 17, No. 88 pp. 19-24.

Ojunga, N., et al., 2014. The Application of Logistic Regression in Modeling of Survival Chances of HIV-Positive Patients under Highly Active Antiretroviral Therapy (HAART): A Case of Nyakach District, Kenya. Journal of Medicine and Clinical Sciences. Vol. 3, No 3. pp. 14-20.

Petros, I., et al., 2010. High Survival and Treatment Success Sustained After Two and Three Years of First- Line ART for Children in Cambodia. *Journal of International AIDS Society*. Vol.13, No 11, pp 21-29.

Picat, M., et al ., 2013. Predicting Patterns of Long-Term CD4 Reconstitution in HIV-Infected Children Starting Antiretroviral Therapy in Sub-Saharan Africa: A Cohort-Based Modelling Study. *Journal of Pediatric Medicine*. Vol. 10, No 10. pp 45-49



Rosma, M., et al., 2012. The Prediction of AIDS Survival: A Data Mining Approach. 2nd WSEAS International Conference. Malaysia.

Sanjel, S., 2009. An Application of Cox Model for Lifetimes of HIV Patients. Unpublished M.Sc thesis submitted to the department of Power Analysis. USA. MC Master University.

Sameem, A., et al., 2010. Classification and Regression Tree in Prediction of Survival of AIDS Patients, *Malaysian Journal of Computer Science*. Vol. 23. No 3. pp 153-165

Shearer, W., 2000. Evaluation of Immune Survival Factors in Pediatric HIV-1 Infection. Annual National Academic Journal. Vol. 91, No. 8. pp 298-312.

Torsak ., A. et al., 2009. The Effectiveness of Highly Active Antiretroviral Therapy among HIV-Infected Children in Asian Countries. *Asian Biomedicine Journal.* Vol. 3, No 1, pp 89-100.

UNAIDS, 2010. Report on the Global AIDS Epidemic. Available From www.unaids.org/globalreport/global\_report.htm. [Accessed 12 June, 2013] UNAIDS, 2012. Together We Will End AIDS. Available from www.unaids.org . [Accessed 12 June, 2013]

UNAIDS, 2013. Global Report on HIV. Available from http://www.unaids.org/en/resources/documents/20 13/name,85053,en.asp. [Accessed 2<sup>nd</sup> August, 2014]

WHO, 2006. Preventing HIV/AIDS in Young People, A Systematic Review of the Evidence From Developing Countries, UNAIDS Inter-Agency Task Team on Young People. World Health Organization, Geneva. Available From: www.who.org . [Accessed 12 January, 2013].

WHO, 2012. Towards Universal access: Scaling up Priority HIV/AIDS interventions in the Health Sector. Progress

Report 2010. Available from

http://whqlibdoc.who.int [Accessed January 15, 2014]

