Editors Policy

Computing and Information Systems offers an opportunity for the development of novel approaches, and the reinterpretation and further development of traditional methodologies taking into account the rate of change in computing technology, and its usage and impact in organisations.

Computing and Information Systems welcomes articles and short communications in a range of disciplines:

- Organisational Information Systems
- Computational Intelligence
- E-Business
- Knowledge and Information Management
- Interactive and Strategic Systems
- Engineering
- E-Learning
- Cloud Computing
- Computing Science

The website for Computing and Information Systems is http://cis.uws.ac.uk
Notes for Authors:

Articles for publication should be submitted on disk or via e-mail, in Microsoft Word for Windows format. A style template can be found at http://cis.uws.ac.uk/crow-ci0/articles/style.doc. Please observe the following guidelines:

1. The Normal paragraph style is Times New Roman 11 point, justified, with 6 pt space before. Avoid inserting extra empty paragraphs. Avoid the use of additional spaces between words or after a full stop. Use full stops after initials and precede subsequent initials by a space. Use the standard heading styles Heading 1, Heading 2 etc for top-level and subsidiary headings within the article.

2. Compose figures in an enclosing invisible frame or rectangle, with captions and any text on the figure provided in embedded textboxes, and all including the enclosing rectangle grouped into one object, defined to have "Top and Bottom" text wrapping. Try to avoid using anything based on the Normal paragraph style in embedded text boxes.

3. Equations should be embedded in the text, possibly as paragraphs in their own right. They should not be treated as figures.

4. All entries in the reference section should be cited in the text. List references at the end of the paper, alphabetically by the first letter of the first author’s last name. References should be identified in the text of the paper by typing the corresponding name and year in parentheses: letters a, b, etc can be added to the year if several references by the same author for the same year are cited. If a page number is set it should be set as (author, year, page number). DO NOT NUMBER REFERENCES; they must be alphabetical and unnumbered. References should have a half-inch hanging indent as shown below and there should be no additional lines between references. Book titles and names of journals should be printed in italics, not underlined. Examples:

Subscriptions:

Computing and Information Systems is issued normally three times per year, in February, May, and October. It is available at an annual subscription of £50, or £20 per issue. Contact the School of Engineering and Computing, University of the West of Scotland, Paisley PA1 2BE. Telephone: (+44) 141 848 3959; Fax (+44) 141 848 3542.
AN ANFIS-BASED FRAMEWORK FOR DETECTING DDOS ATTACK

Gabriel O Ogunleye  
Department of Computer Science,  
Federal University, Oye-Ekiti  
Ekiti State, Nigeria  
gabriel.ogunleye@fuoye.edu.ng

Stephen G Fashoto  
Department of Decision Sciences,  
University of South Africa  
Pretoria, South Africa, gbengafash@yahoo.com

Teresa O Ojewumi  
Department of Computer Science,  
Redeemer’s University  
Ede, Osun State, Nigeria, ojewumiti@run.edu.ng

ABSTRACT

Purpose: In this paper, a simulation environment is set-up to generate distributed denial of service data in Virtual Knowledge Community (VKC) in order to study the pattern of attack in multi-agent environment.

Design/Methodology/Approach: A threshold approach is used to construct the profile of the traffic for the agents in the network, and to identify anomalies whenever traffic goes out of profile. An Adaptive Neuro-Fuzzy Inference System (ANFIS) is used in classifying attack type and to determine its location. Four ANFIS are trained and tested in this research work to provide attack detection and classification.

Findings: An Adaptive Neuro-Fuzzy Inference System (ANFIS) is used to categorise the datasets into attack and normal traffic using some standardized threshold values.

Research limitations/Implications: The study is only limited to detecting distributed denial of service attack in multi-agent environment.

Practical Implications: The experiment and the simulation are carried out in a computer laboratory using Java Agent Development Framework and Adaptive Neuro-Fuzzy Inference System.

Originality: An adaptive security is implemented in this study for detecting distributed denial of service attack in an agent-based virtual knowledge community.

Keywords: ANFIS; DDoS; DoS; Attack

Paper Type: Research Paper

1. Introduction

A Denial of Service (DoS) attack is an harmful process by individual person or a collection of people to cause the target, site, or point to deny service to its authorize users. When this crack is encountered by a lone host of the network, it signifies a DoS attack. Furthermore, a lot of compromised hosts can organize to flood the victim with plenty of attack packets, so that the attack occurs concurrently from different points. This attack method is termed a Distributed DoS, or DDoS attack.

DDoS attack is known as a resource and bandwidth exhaustion attack (Specht and Lee, 2004). In most cases, attackers imitate Internet Protocol (IP) addresses to evade source identity with enormous packets thereby bombarding them at the network, targeting routers, servers, and firewalls and preventing legitimate users from operating their normal services. Packet-flooding is an act or process of bandwidth exhaustion attack where attacker’s intent is to flood the following protocols such as Transmission Control Protocol (TCP), User Datagram Protocol (UDP), Internet Control Messaging Protocol (ICMP), Domain Network System (DNS), and Hypertext Transfer Protocol (HTTP) packets to the target server (Mirkovic and Reiher, 2004). Past research works had it that DoS attack programs have been around
for decades. These earlier single source attacks can be modified easily by means of a protection schemes and where the attacks emanated from can be revolted or terminated with improved tracking capabilities. Nevertheless, as a result of the exponential growth of the internet during the previous years, a big number of susceptible systems are currently useable by the attackers to launch an attack. Attackers can now engage a huge numbers of these weak hosts to establish an attack instead of using a single server, an approach which is not very effective and easily detected. The most eminent DDoS attack took place in August 1999, when a DDoS attack tool known as Trinoo was launched out with 227 compromised systems to bombard a University of Minnesota single computer, which made the victim’s system to be put out of service for more than two days. In the same vein, the biggest scale of DDoS attack also took place on February 2000. It was also on record that similar DDoS attacks occurred in February, 2000 on Yahoo and some websites which made their internet portal and other services to be rendered useless and inaccessible for three hours. The attack disturbed their normal operations resulting into financial losses, service degradation and lack of availability (Arora et al., 2011).

Virtual Knowledge Community (VKC) is defined as an environment where agents share and exchange knowledge (Maret and Calmet, 2009). (VKC) consists of community of communities, agents and community. However, in the past years, the researches conducted in VKC were majorly targeted on knowledge sharing aspects instead of determining a way of securing the VKC itself against any act of attack. Considering the widespread threat direction, it is high time to establish a framework, algorithms, and protocols to realize a protected and steady VKC infrastructure. To maintain a more reliable VKC network territory is onerous because it is more liable to intrusion. Some researchers have proposed a Virtual Communities based on a Multi-Agent Oriented Programming (MAOP) approach (Muhammad et al, 2012). Using the MAOP JaCaMo platform, the authors proposed a multi-agent based Virtual Communities Management System, where agents can help their users in accessing, exchanging and managing information within the community of interest or group of people. To realize the goals of Virtual Communities Management System, the authors used the JaCaMo platform that combines three separate open source technologies, such as, Jason, CARtAgO and Moise for programming MAS. From the combination of these technologies, JaCaMo provides all levels of abstractions that are required for the development of sophisticated multi-agent system, such as, Virtual Community Management System. In addition, the approach offers many advantages, such as decentralization of the community management, personalized automatic management and discovery of communities, and flexibility so that any agent can create its own community. The first implementation of the research work for the gathering of data was carried out using the Java Development Kit, Netbeans IDE 6.5, VMware workstation, Mint Linux operating system, Ubuntu Linux operating system, Windows 8 and Windows 7 operating systems.

The second implementation of the research work for the simulation of the data was carried out using the Matlab version R2012b. In this research, dataset was collected, analyzed and pre-processed to the required format of ANFIS. About 65% of the dataset collected was used as a training data while the remaining 35% was used as a testing data. The rest of the paper is organized as follows: section 2 contains the review of the existing system; section 3 presents the methodology; section 4 presents the system implementation, and section 5 presents conclusion and future works.

2. Literature Review

2.1 Mobile Agent

A mobile agent is regarded as a computer program that has the ability to move or transit from one host to another in a different network (Kuo-h and Yu-F. 2009). They are also categorized as moving agents, in such a way that these programs can travel back and forth with their code and state among resources. They are known as network roamers that act as personal representative that is they can work independently through networks. They have potential of visiting network nodes directly using the obtainable computing power and they are not in any way restricted by platform. The technology has become an alternative approach for the design
and implementation of distributed systems to the traditional Client/Server architecture.

Mobile agents are able to migrate from one system to another during their execution and communicate amongst one another, clone, merge and co-ordinate their computations. Mobile agents are autonomous agents in such a way as they control their relocation behavior in pursuit of the goals with which they are tasked (MengX et al., 2003). Main fields of application for mobile agents are information retrieval on the www, distributed database access, parallel processing, automation of electronic marketplaces and others. Mobile agent frameworks are currently rare, due to the high level of trust required to accept a foreign agent into one’s data server. However, with the advances in technologies for accountability and immunity, mobile agent systems are already becoming popular.

2.2 Issues of Security in Mobile Agents

2.2.1 Attacks on Mobile Agents by Mobile Agent Platforms

For a strong mobility of mobile agent, all its code, data and state are exposed to the mobile agent platform in which it migrates for execution of operation. Because of this, mobile agent faces more severe security risks. The following are possible attacks by malicious platforms (Mousa and Lijiljana 2004):

(i) Leak out/ modify mobile agent’s code
Due to the fact that mobile agent’s code has to be processed by a guest platform, so this malicious platform can read and remember instructions going to be executed to infer rest of the program based on that knowledge. By this process, platform acknowledges the strategy and purpose of mobile agents (Schelderup, 1999). If mobile agents are generated from standard building libraries, the malicious platform knows a complete picture of mobile agent’s behavior and it finds out the physical address and can access its code memory to modify its code either directly or by insertion of virus. It can even modify code temporarily, execute it and finally resuming original code before the mobile agent leaves.

(ii) Leak out/ modify mobile agent’s data
There are many data which are very security sensitive like security keys, electronic cash, social security number that cause leak of privacy or loss of money. If the malicious platform gets to know the original location of data it can modify the data in accordance with the semantics of data (Meadows, 1997). Above tasks can lead to severe consequences. Even if data is not sensitive, malicious platform can attack on normal data like traveling data of person and leaking it to somebody.

(iii) Leak out/ modify mobile agent’s execution flow
By knowing the mobile agents physical location of program counter, mobile agent’s code and data the malicious platform can predict what will be set of instructions to be executed next and deduce the state of that mobile agent. By help of this process, it can change the execution flow according to its will to achieve its goal (Jansen and Karygiannis, 2000). It can even modify mobile agent’s execution to deliberately execute agent’s code in a wrong way.

2.3 Virtual Knowledge Communities

Maret and Calmet, (2009) came up with an approach or initiated the idea of agent-based communities which enhance the reflection of a mobile agent, in such a way that it acts within the system, inquiring about knowledge among other agents through multi-agent communities. The model developed enables the agents to make a decision to join, leave, create and destroy a community. They can seek for information and send information to the community, and they can be a member of several communities synchronously. Agents can meet, interact and communicate among themselves which is termed “Virtual Knowledge Communities”. This is because a VKC is focused on a topic, similar to a
domain of interest for which the interested agents have joined this community. This notion allows an increased availability of data and knowledge within the various communities.

Muhammad (2012) proposed a multi-agent based Virtual Communities Management System (VCMS) where agents can assist their users in accessing, exchanging and managing information within the community of interest or group of people. To actualize the goals of Virtual Communities Management System, the authors used the JaCaMo platform that combines three different open source technologies, that is, Jason, CArtAgO and Moise for programming multi agent system. From the blending of these technologies, JaCaMo covers all levels of abstractions that are considered for the development of sophisticated multi-agent system, such as, VCMS. In addition, the approach offers many advantages, such as decentralization of the community management, personalized automatic management and discovery of communities, and flexibility so that any agent can create its own community. The VCMS that was created, implements a generic approach for creating "mass" of places of local exchanges and local knowledge bases in the context of smart places. VCMS is focused on a multi-agent architecture that enables community assistants (participants) to meet, share and gain quick and efficient access to the information of their interest. In VCMS, virtual communities are realized by means of a set of Jason agents encapsulating the user profile and the logic and control of the specific operations involved in the community pack: community management (for example, joining/leaving, creating/deleting a community), information sharing (for example, publishing posts, notifying members, etc.).

![Figure 1: A VCMS Architecture](source: Muhammad et al. (2012))
2.4 Existing Methods for DDoS Attack Detection

In this section, a summary of existing literature on DDoS attack detection methods are presented. These processes are based on the architectures discussed above namely, victim-end, source-end and in-network.

Past research works have shown that soft computing techniques have been employed largely for DDoS attack detection without applying the scheme towards VKC environment. It was discovered that a group of classifiers for DDoS attacks have also performed satisfactorily with high detection rates. DDoS attack detection can be classified into four major categories as follows:

(a) Statistical Methods
Statistical properties can be applied for detection of DDoS attacks using normal and attack patterns. Typically, a statistical model with normal traffic method is adapted and then a statistical inference test is utilized to ascertain if a new instance conforms to this model. Processes that do not adapt to the learnt model, based on the applied test statistics, are classified as anomalies. Some authors have also proposed a method of Distributed Change Point (DCP) detection framework using change aggregation trees (CATs).

(b) Soft Computing Methods
Learning paradigms, for example neural networks, fuzzy logic, neuro-fuzzy algorithm, radial basis functions and genetic algorithms are extensively applied in DDoS attack detection as a result of their potency to classify intelligently and automatically. Soft computing is a general term for identifying a collection of optimization and processing styles that are permissive of imprecision and uncertainty. Jalili (2005) introduced a DDoS attack detection system known as SPUNNID which is based on statistical pre-processor and unsupervised artificial neural nets. They used a statistical pre-processing to extract features from the traffic, and an unsupervised neural net to analyze and classify traffic patterns as either a DDoS attack or normal.
Karimazad (2011) proposed an anomaly-based DDoS detection method based on characteristics of attack packets, analyzing them by using Radial Basis Function (RBF) neural networks. The method can be used for edge routers of victim networks.

(c) Knowledge-based methods
From this method, network events are analyzed against predefined rules or patterns of attack. The approach gives general representations of known attacks which are formulated to identify actual occurrences of attacks. Examples of knowledge-based approaches include signature analysis, self-organizing maps, and state transition analysis, expert system.

(d) Other data mining and machine learning methods
An effective defense system to protect network servers, network routers, and client hosts from becoming handlers, zombies, and victims of DDoS flood attacks is presented in Hwang et al, (2003). The NetShield system protects any IP-based public network on the internet. It uses preventive and deterrent controls to remove system vulnerabilities on target machines. Adaptation techniques are used to launch protocol anomaly detection and provide corrective intrusion responses. The NetShield system enforces dynamic security policies. NetShield is especially tailored for protecting network resources against DDoS flood attacks. Chen et al, (2007) presents a comprehensive framework for DDoS attack detection known as DDoS Container. It uses a network based detection method to overcome complex and evasive types of DDoS attacks. It works in inline mode to inspect and manipulate ongoing traffic in real time. By continuous monitoring of both DDoS attacks and legitimate applications, DDoS Container covers stateful inspection on data streams and correlates events among different sessions. It proactively terminates the session when it detects an attack.

Zahra and Mohammad (2011) proposed different approaches to implement and improve the security level of Mobile Ad hoc Networks. They designed a mechanism for intrusion detection and security framework to detect a security attack. In a type of attack considered in this research, an intruder node injects a large amount of junk packets into the network and causes a denial in the services of the attacked node to the network. The model was developed using 2 method of detection – ANFIS and ANNs – in a simulated environment. It was showed that almost all of models could detect Dos attack effectively.

Ogunleye et al., (2011) proposed a framework for securing knowledge agents in virtual knowledge communities against malicious attack by sharing knowledge in an encrypted format. Their approach was to enhance privacy so that malicious agents were not able to read the contents of important knowledge. Also, it will also help to check the case of eavesdropping. The last goal of their work is that no one except the knowledge receiver must be able to decrypt the knowledge and the result.

3 System Design
3.1 Agent Communication in VKC
Agent Communication can be defined as a form of interaction in which the dynamic relationship between agents is expressed through the intermediary of signals, which once interpreted will affect these agents. A large number of agent communication forms exist. In this research paper, an act of communication was used by sending some information from a sender to a set of (intended) receivers is used. This information is encoded with the help of languages and decoded upon arrival by the receivers. We used an abstract communication model of FIPA that derives from speech act theory. In this model, communication occurs through the exchange of asynchronous messages corresponding to communicative acts. The Agent Communication Language (ACL) format defines the format of these messages. The
ACL messages used in this research paper is characterized by the following:

- **Intention** e.g., REQUEST, INFORM, QUERY_REF
- **Attendees** i.e. the sender and a set of receivers
- **A content** i.e. the actual information that is exchanged.
- **Content Description**, i.e: an indication of (i) the content language used to express the content and (ii) the ontology by means of which both the sender and the receiver ascribe a proper meaning to the terms used in the content.
- **Conversation control**: interaction protocol and conversation identities.

ACLMessage: Example of the message expressed in XML with intention INFORM;

```
<agent-identifier>
<name id = "Ope@host1:8888/JADE" />
</agent-identifier>
<brider>
<content>
(Greeting Message )
</content>
</fipa-message>
```

**3.2 Membership Function used for each of the Input Variable**

In this section, the linguistic values are mapped to their respective membership functions using the triangular membership formula as described in the equations 1 to 12.

- The membership used for each of the input variable is as follows:
  - The median point (n) =0.55, the lower boundary (l) =0.2 and the upper boundary (k) =0.95 as shown in equation 1, 2 and 3.
  
  for %Synchronous flag,
  
  \[ \mu_{\text{Low}}(x) = \begin{cases} 0.1 & \text{if } 0.2 \leq x < 0.2 \\ 0 & \text{otherwise} \end{cases} \]  
  
  \[ \mu_{\text{medium}}(x) = \begin{cases} 0.2 & \text{if } 0.2 \leq x < 0.55 \\ 0 & \text{otherwise} \end{cases} \]  
  
  \[ \mu_{\text{high}}(x) = \begin{cases} 0.65 & \text{if } 0.65 \leq x < 0.95 \\ 0 & \text{otherwise} \end{cases} \]  

  for probability of distinct source ports,

  \[ \mu_{\text{Low}}(x) = \begin{cases} 0.15 & \text{if } 0.15 \leq x < 0.25 \\ 0 & \text{otherwise} \end{cases} \]  
  
  \[ \mu_{\text{medium}}(x) = \begin{cases} 0.3 & \text{if } 0.3 \leq x < 0.65 \\ 0 & \text{otherwise} \end{cases} \]  
  
  \[ \mu_{\text{high}}(x) = \begin{cases} 0.7 & \text{if } 0.7 \leq x < 0.98 \\ 0 & \text{otherwise} \end{cases} \]  

  for bit rate,

  \[ \mu_{\text{Low}}(x) = \exp \left\{ 0.5 \leq x < 2.5 \right\} \]  
  
  \[ \mu_{\text{medium}}(x) = \exp \left\{ 3.0 \leq x < 4.0 \right\} \]  
  
  \[ \mu_{\text{high}}(x) = \exp \left\{ 6.0 \leq x < 8.5 \right\} \]  

The median point (n) =4.0 * 10^5, the lower boundary (l) = 2.5 * 10^5, and the upper boundary (k) =8.5 * 10^5, as shown in equation 1,2 and 3 as represented in equation 7, 8 and 9 as seen in equation 7, 8 and 9.
for rate of change of bit rate,\\
\[ \mu_{Low}(x) = \exp\{0.5 \leq x < 0.9\}^{5} \] (1)\\
\[ \mu_{Medium}(x) = \exp\{1.5 \leq x < 2.5\}^{5} \] (1)\\
\[ \mu_{High}(x) = \exp\{3.0 \leq x < 5.5\}^{5} \] (12)

The median point (n) = 2.5 \times 10^5, the lower boundary (l) = 0.9 \times 10^5, and the upper boundary (k) = 5.5 \times 10^5, as shown in equation 1, 2 and 3 as represented in equation 4, 5 and 6 as seen in equation 10, 11 and 12.

3.3 Output Interpreter or Output Variable

This section refers to the presence of attack in the VKC network. It comprised of integer value from 0 (no presence) to 1. By increasing of integer value, attack risk increases in the network. We considered a different output variable, which divides to 3 fuzzy sets (normal traffic(n1), possible attack traffic(n2) and attack traffic(n3). Membership functions of “normal” traffic (n1), possible attack traffic(n2) and attack traffic(n3) were considered using a bell membership function. The thresholds that were applied in this technique were adapted from the work of [23] which states that:

- when the Output < 0.25 – for normal
- when the range is 0.25 < output < 0.65 – for a possible attack
- when Output > 0.65 – for attack

As a result of this, ranges (0.25, 0.65) were employed to notify administrators or concerned personnel of the presence of suspicious activity in the network.

4.0 System Implementation

To generate attack traffics, Figure 3 shows the interface generation window for the data collection for the experiment. Tcp-syn flood attack command was used as the reference point in this research work.

The experiments were performed and the sizes of the training and testing sets were determined by taking into consideration the classification accuracies. The dataset was divided into two separate data sets – the training data set and the testing data set. The training data set was used to train the ANFIS, whereas the testing data set was used to verify the accuracy and the effectiveness of the trained ANFIS model for the detection of DDoS attack in VKC.

Figure 3: Attack/Normal Traffic Generation Window
For normal data, from figures 4(a) and (b), we observed that when the percentage of tcp-syn packets was very small there was no suspicious attack. Furthermore, when the data rate is small and the tcp-syn packet is large, there is likewise no suspicious or traces of attack.

**Figure 4(a): Percentage of TCP SYN packets and probability of distinct source ports**

![Graph](image1)

For attack data, Figures 5(a) and (b) illustrate the behavior of the network attack traffic. When the tcp-syn pack is high, it shows that the distinct source port is likewise high making the data rate to be also large. This shows that there is attack traffic arriving at the victim’s end.

**Figure 4(b): Data rate and rate of change of data rate for Normal network traffic**

![Graph](image2)
The second implementation of the proposed hybrid intelligent system for the simulation of the data was done using the Matlab version R2008b. In this paper, dataset was collected, analyzed and preprocessed to the exact format of ANFIS. About 65% of the dataset collected was used as a training data while the remaining 35% was used as a testing data. The dataset comprised of both attack and normal traffic.

Back-propagation method was used and the number epochs employed in this research work is 30. Epochs provide the different iterations for the system. The ANFIS employed is the Sugeno inference system as analyzed in the literature above.
The threshold set for the output variable is shown below:

- Output <0.25 – for normal 
- 0.25< output < 0.65 – for a possible attack 
- Output > 0.65– for attack 

Hence, from Figure 6, output less than 0.25 is taken to be a normal traffic and output that falls between the range 0.25 and 0.65 is taken to be a suspicious attack, when the output is greater than 0.65 is taking to be an attack. The hybrid intelligent system (ANFIS) had four inputs. Each had three membership functions corresponding to low level, medium level and high level of the quantity, respectively as it is shown in figure 6.

Model validation is the process by which the input vectors from input/output data sets on which the FIS was not trained, are presented to the trained FIS model, to see how well the FIS model predicts the corresponding data set output values. This is accomplished with the ANFIS Editor GUI using the so-called testing data set, and it is shown in figure 7.
Since attack data and normal data were assumed an output value of 1 and 0 respectively, then figure 7 shows a good fit. The dots at point 1 and 0 showed that the output is a good fit comprising of purely an attack data and normal data respectively.

5 Conclusion and Future Works

In this research work, DDoS detection in virtual knowledge communities has been studied. Thus, DDoS attack detection is a serious issue in cyber security. The agents used in this work as highlighted in this research were created under the Java Agent Development Framework (JADE) system. The implementation as well as the simulation addressed the problem of DDoS problem in knowledge sharing activities of the agents. Datasets were generated through the data collection agent which records all the traffics in the VKC environment. About 65% of the data generated were pre-processed and fed into the ANFIS as training data while the remaining 35% were used as a testing data. The paper has indicated that there is a need to develop efficient systems capable of resolving DDoS problems during the sharing of knowledge in VKC.

5.1 Future Works

In order to provide a comprehensive network security, the ability to prevent, detect as well as respond to threats is critical. There are many possible directions for future work, which include the following:

(a) In this research work, the number of systems used to set-up the experiment is minimal; more systems should be gathered in the future to get more training and testing data.

(b) Furthermore, only tcp-syn attack was studied, future works should investigate other types of DDoS attacks in the VKC network.

(c) In addition, future work should concentrate more on developing a more intelligent and robust system that will be able to detect and resolve DDoS in heterogeneous network communities.

References


About the Authors

FIRST AUTHOR: Dr. Ogunleye Gabriel Opeyemi is a lecturer in the Department of Computer Science, Federal University, Oye-Ekiti, Ekiti State, Nigeria. He holds a Ph.D degree in Computer Science from The Federal University of Technology, Akure, Ondo State. His research interests include Computer Security, Mobile Agent Systems, and Artificial Intelligence among others.

SECOND AUTHOR: Dr. FASHOTO, Stephen Gbenga had his Ph.D from University of Ilorin, Nigeria in 2014. He is currently a post-doctoral research fellow at Department of Decision Sciences, University of South Africa Pretoria, South Africa. He has published a number of articles in reputable journals and learned conferences.
THIRD AUTHORITY: Mrs. Ojewumi. Teresa O. is a programmer I in the Department of Computer Science, Redeemer’s University, Ede, Osun State, Nigeria. She holds a masters degree in Computer Science from University of Ibadan, Oyo State. She is currently pursuing her Ph.D degree at Ladoke Akintola University of Technology, Ogbomoso. Her research interests include E-Learning, Fault Tolerant Systems among others.
ABSTRACT

Purpose: This paper proposes High Availability (HA) Application Models for deploying eCommerce solutions leveraging open source and cloud computing technologies based on global best practices in production environment architecture for mission critical applications.

Design/Methodology/Approach: The key functionality requirements of the production environment were identified and incorporated into the design of the HA models based on existing designs for Database replication, load balancing, monitoring and backups. The designs of Anicas (2015a) on production server setups were adopted and modified to design three (3) High Availability eCommerce Production Environment models.

Findings: Recovery plan, load balancing and scaling were identified as the most critical components of HA systems. At least, two of these components were incorporated in each of the three (3) HA models with varied degree of feature implementation.

Research limitations/implications: The HA models outline production environment features and recommend open source technologies that could be used to implement their functionality. Technical details on implementation or site security were not discussed.

Practical implications: These models shall be very useful to system engineers working on eCommerce solution as they could adopt or modify the designs for implementation on their infrastructure.

Originality/value: The HA models were modified to suit eCommerce solution requirements and outline essential components of Highly Available eCommerce solutions which could be valuable information for the system engineer or eCommerce solution expert.

Keywords: eCommerce in Nigeria, high availability systems, Cloud Server architecture, Private networks, Production Environment

Paper type: Conceptual Paper

1. Introduction

eCommerce has been a little more active in Nigeria in the past couple of years due to the significant evolution of the Nigerian retail industry from traditional to more modern systems (Philips Consulting, 2014). Nigerians are gradually adopting online shopping, the primary channel of eCommerce, and the convenience of using POS (Point-of-sale) channels to pay for goods and services. Online shopping is gradually gaining popularity due to the fact that it is relatively convenient to pay for goods and services through this channel. Besides, the prices of goods and services on eCommerce platforms are reasonable and affordable (Philips Consulting, 2014). The internet penetration level has greatly influenced this evolution. As at the last quarter of 2015, over 45% of the Nigerian population could access the internet through their computers, tablets, smart phones etc. This number has grown from 1.3% in 2004 to 24% in 2010 and a little over 46% by the first quarter 2016 (InternetLiveStats, 2016).

There has been significant growth in eCommerce adoption in Nigeria partly influenced by the Central Bank of Nigeria’s “Cashless Policy”. This adoption benefits from economic advantages such as market expansion, reduction of product source prices, promotion of productivity, reduction of transaction costs and inflation, lowering uncertainty, sharing market information, and aiding in distribution channel efficiency (Mohammed & Abdulkadir, 2012). The online shopping sector in Nigeria recorded about N49.9 billion worth of transactions in 2010, N62.4billion in 2011 and N78billion in 2012. This figures account for over 25% growth between 2010-2011 and 2011-2012 respectively (Philips Consulting, 2014). Business owners are constantly exploring alternative channels for enhanced productivity and profitability. With the current trends, business owners are gradually adjusting towards fitting their service delivery options to the eCommerce ecosystem.
There are numerous benefits of eCommerce to the entrepreneur, business owner and the consumer. Some of the benefits for the entrepreneur and business owner include online presence, improved sales conversion rate, reduced operational cost, brand boost and better understanding of customers. The consumers on the other hand enjoy access to different pricing for a product, price comparison, convenience of office or home delivery, access to product specifications, user reviews etc. It is believed that the awareness of these benefits drive the adoption of eCommerce in Nigeria (Philips Consulting, 2014). eCommerce opens up a single business to the world and provides an international marketplace for businesses to advertise goods and services. Some of the marketplaces that have contributed to the development and evolution of eCommerce in Nigeria include Konga, Kaymu, Dealdey, Olx and Cheki.

Jumia, Konga and Kaymu are ranked as the top three (3) eCommerce websites in Nigeria (Egbonwon, 2015). Kaymu.com was launched as an online marketplace for buyers and sellers. Business owners do not have to go through the hassles of setting up an eCommerce website; they just sign up as a merchant, upload and configure their products on the platform and Kaymu handles the rest. Konga on the other hand, offers her eCommerce platform as a service to merchants and also sells her products. The Konga marketplace was launched in 2014 and is currently serving thousands of merchants across Nigerian cities. This top three eCommerce websites were all launched in 2012 and have been very active till date with most of them expanding beyond Nigeria to other countries in Africa (Egbonwon, 2015). Konga registered “Nigeria’s record breaking online sales event” on her 2014 edition of black Friday tagged “YAKATA”. The event was successful, breaking all the records for online sales in a day for the first time and generated over N200 million naira for Konga merchants in just one day (Eluwande, 2014). With success stories of existing eCommerce solution providers and merchants, Nigerian entrepreneurs and business owners are interested in being part of the ecosystem. Unfortunately, majority of new or prospective entrants into the Nigerian eCommerce ecosystem only have a firm understanding of the business part of eCommerce and have little or no knowledge of the technology requirements. There is no gain-saying the fact that technology is an integral part of eCommerce and goes beyond just one software for selling products online. The platform on which a robust eCommerce solution sits on comprises server hardware and software with other local or remote services for monitoring, logging, analytics, security, automated scaling and security. This is a basic requirement which most new or prospective entrants are either not aware of or choose to defer integration of the services because of the cost. There is need to emphasize how essential it is to deploy high availability eCommerce solutions and recommend low cost models for its implementation. This paper discusses the basic requirements for deploying high availability eCommerce solutions and focuses on adopting open source with cloud technologies for implementation.

2. Open Source eCommerce

Open Source, a term commonly used by programmers and application developers, refers to software or an application released under a family of “free to use, modify and change” licenses. A lot of people from the large enterprises tend to be very sceptical about open source solutions yet most of the fortune 500 hundred companies in the world including Apple, Google, Facebook, Microsoft and Oracle have contributed in various ways to the development of the open source ecosystem.

eCommerce technologies emerged a few decades ago but these solutions were mainly driven by proprietary applications developed in Java. In recent times, open source eCommerce solutions have been gaining prominence and boldly challenging the proprietary eCommerce solutions in the overall market share. Open source solutions (mainly PHP based) lead in market share for emerging to medium sites while proprietary eCommerce solutions (mainly JAVA based) lead in the market share of big to very big sites (Humeau & Jung, 2013). Whenever the term “open source” pops up in eCommerce solution deployment discussions, some people tend to ask questions like: what does open source mean? (Wood, 2015).

2.1 Important Open Source Facts

2.1.1 What does Open Source mean?

Open source software is a software whose source code has been made available for anyone to modify or update. In a more technical perspective, the software is distributed to the end user with full access to the source code under any of the open source licenses which outline terms of software use (Wikipedia). Most often, an open source software is released as a result of the collaborative effort of developers from different companies, countries or continent. Some of the notable open sourced software solutions include:
• **Android** – Operating System for Mobile phones
• **Linux** Based OS Flavours – Operating System for Desktop, Laptops and Servers similar to Windows and Macintosh
• **PHP** – Server side scripting language used by Google, Yahoo, Facebook etc.
• **Moodle** – Learning Management System
• **Magento** – Most popular open sourced eCommerce framework

### 2.1.2 Is Open Source Really Free?

Open source means free technically. It is free in the sense that, the source code has been released for public consumption but programing expertise is required in order to modify or change the program codes. The ability to implement a particular open source solution is dependent on your proficiency in the programming the language used for its development and the level of difficulty associated with it. If you are not proficient in the programming language, you may hire an experience developer to customize and maintain the solution for you. Otherwise, you can opt for a hosted eCommerce service. For continuity, you may need to run estimates to see if the cost of a developer is cheaper over time than the cost of a hosted eCommerce solution (Wood, 2015). Some of the notable open source eCommerce frameworks are: Magento, Open Cart, Prestashop, Simple Cart, ZenCart, OsCommerce and WooCommerce.

### 2.1.3 Is it a good choice for my small store?

Setting up an open source store will likely take a lot of time and/or money, depending on what and who you know. Generally, if you are not a developer or do not have an employed developer that is proficient in the technology used for the eCommerce solution, you should probably avoid open source. If you need a small online store, you may sign up for merchant account on Kaymu or Konga.

### 2.1.4 Is Open Source eCommerce powerful enough to create a store for my very large chain?

Yes. Certain solutions perform better than others but you need a creative technical team. Most of the time, you can do more with open source than you can with a hosted or proprietary solutions. Open source eCommerce solutions are very flexible, robust and extensible. In fact, the top three eCommerce websites in Nigeria use eCommerce solutions written with open source server side language – PHP (e.g Konga uses Magento eCommerce framework).

### 2.2 eCommerce and Cloud Computing

In recent times, eCommerce solutions are preferably hosted in the cloud. This is attributed to the flexible subscription model of Cloud technologies (pay-as-you-use) and service delivery model were services are served based on demand (Wang, 2013). For large eCommerce sites, during peak periods, the services of additional ten to hundred servers may be needed to handle high traffic while fewer servers may be needed during off-peak; cloud technologies comfortably fit into this mould.

The influences of cloud computing on eCommerce technical architecture is demonstrated by building, implementing and maintaining of the technical layer of an eCommerce solution. This separation enables enterprises rent, rather than purchase hardware and software for server infrastructure (Wang, 2013). Employing cloud technologies to handle the technical layer is like outsourcing. Outsourcing to Cloud computing sets the eCommerce enterprises free from the complicated technical architecture planning, designing and maintaining of servers. It enables them to focus on the core businesses and service delivery to clients (Wang, 2013).

### 2.3 Considerations for Deploying eCommerce Application

eCommerce applications are mission critical applications which are tailored to satisfy the needs of the customers round the clock. A lot of entrepreneurs in Nigeria have gained interest in participating in the eCommerce ecosystem. Though some big players in the Nigerian ecosystem like Konga.com provide merchant services on their marketplace platform for business owners to sell online, a few entrepreneurs prefer to own and be in control of their eCommerce solution. Whether you are planning to launch a start-up eCommerce solution or move your existing business online, here are a few considerations you need to factor in before you deploy.

#### 2.3.1 Budget

Your budget is key in your decision making regarding deploying an eCommerce application. If you are spinning off a start-up, you need to start small but if you have already been in business for a couple of years and you need a robust online presence, then you can explore the best options or go for enterprise solutions.
2.3.2 Technology

Computing is basically about software and hardware technologies. The robustness of an eCommerce solution is highly dependent on software and hardware technologies. Hence, the decision on what technology one adopts for deploying eCommerce application should come at the early stage of planning and feasibility study. The choice of technology could be limited by budget and technical requirements but the good news is that there is a solution for everyone. Below is a hierarchy diagram which I have drafted to highlight technology relationships for an eCommerce solution.

![Technology Relationships Diagram](image)

**Fig.1 – Technology Relationships**

2.3.3 Scalability

Scalability is the capability of a system, network, or process to handle a growing amount of work, or its potential to be enlarged in order to accommodate that growth (Bondi, 2000). For eCommerce applications, scalability is very essential. Chosen software and hardware technologies must be able to accommodate expansion. Most often, when an eCommerce application is deployed, the number of users accessing and performing transactions increase gradually. The number of daily unique visitors increases from tens to hundreds and then thousands etc. What usually causes problem is the number of concurrent users. When hundreds, thousands or millions of users try to login or perform transaction at the same time, the server load increases and this affects both software and hardware components. At first, your system hardware may have the capacity to handle a few hundred concurrent users, you should consider accommodating more concurrent users as the user-base grows. With cloud computing, it is a lot easier to scale hardware horizontally or vertically depending on software limitations.

Scaling in cloud computing involves vertical scaling (increasing computing power by adding extra RAM/CPU Cores) and horizontal scaling (expanding server cluster by adding more machines to the pool). There are options for automated scaling and manual scaling. Whatever the case may be, it is more efficient when the scalability factor is tailored into the architecture of the eCommerce solution from the beginning.

2.3.5 Technical Capacity

It is basic to have employees with sound knowledge of eCommerce technologies and cloud computing. The various options available for deploying eCommerce application still require a certain level of
technical expertise in the subject. A web developer with basic knowledge of eCommerce or open source eCommerce solution is essential as an in-house technical contact. If you do not have employees with sound knowledge of eCommerce solutions or cloud computing, it is advisable to go with managed services (hosted eCommerce & manage cloud). But, employees with in-depth knowledge of cloud computing and or eCommerce technologies, can help in starting-off with an open source eCommerce framework and unmanaged cloud service at a lower cost.

### 2.3.6 Production Environment

As earlier mentioned, a robust eCommerce solution comprises software and hardware technologies integrated into a single platform to enable users perform electronic transactions through the World Wide Web. This single platform handles all eCommerce related activities and is otherwise referred to as the “Production Environment”. In cloud computing, servers are more common terms used to describe the hardware/software services. A production environment typically refers to a server environment that is designed specifically for an application or a set of related applications (Anicas, 2015a). In this case, we are referring to a dedicated production environment. A simple production environment could be seen in the figure below:

![Fig.2 – Simple Production Environment](http://example.com/)

The figure above shows a production environment with four (4) cloud servers namely: load balancer, application servers (app-2 & app-2) and database server. The eCommerce application resides in the app-backend while the database is hosted in a dedicated database server. This simple example was designed with scaling in mind, as opposed to the traditional or conventional server setup where everything is in one box like the figure below.

![Fig. 3 – Single Server Setup](http://example.com/)

### 3. Design Considerations for Production Environment

There are three basic factors to consider when you are designing Production Environment architecture for mission critical applications. This factors are availability, recoverability and performance.

- **Availability**: The ability of an eCommerce application to be usable by intended users during advertised hours without service
disruption. Failure of any critical component can cause service disruption (e.g. the application crashes due to a bug, the database storage device fails, or the system administrator accidentally powers off the application server). To promote availability, the number of single points of failure must be decreased in the production environment (Anicas, 2015a).

- **Recoverability**: The ability of an application environment to recover in the event of system failure or data loss. If a critical component fails, and is not recoverable, the application will become unavailable. Designing production environment architecture with a recovery plan ensures recoverability in case of system failure.

- **Performance**: When an eCommerce application performs as expected under average or peak load, its performance could be rated as satisfactory. The performance of an application is an important factor for users but only when the application is available (Anicas, 2015a). Hence, the eCommerce solution must be available (online) for users to enjoy the application performance benefits.

### 3.1 Production Environment Architecture

The production environment comprises three main modules namely, User Module, Application Module and Recovery Module. Most often, only the first two modules are considered in production setup. In a production environment, the recovery module is very essential. A simple production environment architecture is shown in the figure below.

![Production Server Setup Diagram](image)

**Fig. 4 – Standard Production Environment Setup**

*Culled from Anicas, M. (2015a).*

**The User Module**: The user module represents users who may access the application with their computers, laptops and mobile phones. It also represent third party applications that may wish to communicate with the application through a Web API.

**The Application Module**: This module comprises the load balancer, application backend and database server. The components of this module are usually servers with specific functions. Other components may include DNS and cache servers.

- **Load balancer**: This is a server that improves performance and availability by distributing workload to multiple servers (app backend) with horizontal scaling capability.

- **Application Backend**: This is a server or group of servers where the eCommerce application is hosted. If more than one server is used, a replication software may be used to ensure that application files, server configuration and application assets are the same across application servers. This setup allows for horizontal scaling as more servers can be added as the need arises.

- **Database Server**: This is a single server or a group of servers (one master server, many slave servers) hosting the database application. A replication software is often used to enhance the functions of the Master and Slave databases and to ensure data management.
integrity across database server within the layer.

**Master-Slave Database Replication**

![Diagram of Master-Slave Database Replication]

*Fig. 5 – Production Environment with Master-Slave Database Replication Culled from Anicas, M. (2015a).*

**The Recovery Module:** Recoverability is very essential in a production environment. Though there are a few other components in the recovery module, it basically comprises a recovery plan and a backup system. The backup system comprises local and remote backup systems while the recovery plan is a set of actions to be performed manually or automatically if any critical system component fails. Other components of the recovery module are monitoring and central logging servers.

![Diagram of Recovery Plan: Database Server Failure]

*APPLICATION*

*Fig. 6 – Backup and Recovery Plan Culled from Anicas, M. (2015a).*

### 3.2 High Availability System

A High Availability (HA) system is one whose services automatically switch to a backup/passive system on event of failure and switches back to the primary system once its health status is restored to normal. HA systems are designed such that there is no single point of failure in the system. In a typical HA system, at least two load balancers are required for a simple setup to protect against downtime by ensuring that one load balancer is always passively available to accept traffic if the active load balancer is unavailable (Anicas, 2015b). The figure below demonstrates the HA system setup.
The user, accesses the website through the domain (e.g. www.example.com), the domain forwards to the Floating IP. The Floating IP is the gateway to other layers of the production environment. The request is then forwarded to the Primary Load Balancer which is active. If the Primary Load balancer is not healthy, the request is forwarded to the secondary load balancer instead. Once the primary Load Balancer has been fixed, the system automatically reverts to it.

E Commerce applications are usually designed such that its services are delivered round the clock (24/7) with reliable live support. To achieve this primary goal, the production environment must be Always Available, Scalable and Recoverable with satisfactory Performance during peak and off-peak periods. The production environment should be designed to handle all possible scenarios and events involving high traffic, DDOS (Distributed Denial of Service) attacks, application bottlenecks, crash of application due to bugs, system maintenance etc. Fail-over solutions should be tailored into the architecture to ensure that core application services are “Always Available”.

3.3 High Availability Application Models

It is common practice nowadays that cloud hosted applications especially eCommerce applications are “built to scale". The application infrastructure is often designed such that, as the website traffic increases, the cloud infrastructure can also be scaled to accommodate the difference. When deploying an eCommerce website, you may be confronted with the fact that it is a basic requirement for the cloud infrastructure to ensure that your application is “always available” since online stores are expected to be open 24/7. The cost of the cloud hosting infrastructure is an important factor that should be considered. Some entrepreneurs or business owners prefer hosted cloud services or managed eCommerce solutions while others prefer to setup, deploy and manage the application infrastructure by themselves. Managing your cloud infrastructure definitely gives you more control. For starters, you may not have to setup a standard highly available system. Since global best practice in cloud computing for high availability setup requires a minimum of sixteen (16) cloud servers (15 local and 1 remote), you may have to adopt other HA models that will be cost effective and efficient depending on your traffic or concurrent user demands.

Having considered the basic requirements for a production environment, and other requirements for High Availability systems, the following High Availability Application Models have been designed based on Anicas (2015b) production environment setup designs, to suit three (3) levels of traffic demands on an eCommerce website. The models are:

1. ePE-Standard Performance (ePE-SP)
2. ePE-High Performance (ePE-HP)
3. ePE-Turbo Performance (ePE-TP)
3.3.1 ePE-Standard (ePE-SP)

The ePE-Standard Performance model is the simplest production environment setup for eCommerce applications with minimal cloud server infrastructure requirements. With a minimum of three (3) cloud servers and a few remote services, the eCommerce application can be deployed. The basic requirement includes: one (1) Application Server, one (1) Database Server and one (1) Backup server. For an eCommerce website handling a few hundreds of requests and a few hundreds of concurrent users, this setup is ok. RAM size, HDD/SSD space, CPU cores may be scaled manually for higher system performance. A remote CDN & Web application firewall (WAF) service may be integrated in the architecture to enhance performance and protect against DDOS attacks. A cloud service for monitoring application health is also an essential component of this setup, as it sends notifications via email on event of application module failure.

Fig. 8 - ePE-Standard Performance Model

3.3.2 ePE-High Performance (ePE-HP)

The ePE-High Performance model is a little more advanced than the ePE-SP model. It allows for more robust scaling of cloud infrastructure (manually and automatically) and can handle very large volumes of user requests. With automated replication, the ePE-HP can atomically add more servers to the application backend or database backend to cater for heavy traffic on the eCommerce website. The setup requires a minimum of ten (10) local cloud servers with a common private network and one (1) remote backup server.

The Web Application Firewall (WAF) service filters and forwards user requests to the Load Balancer which in turn evenly distributes it across the application backend servers. The database backend could be scaled to add more Database Slave servers to the pool while the DNS servers ensure that the hostnames of local servers within the private network is propagated efficiently. This setup could be scaled to a few tens to hundreds of cloud servers for improved performance.
3.3.3 ePE-Turbo Performance Model

The ePE-Turbo Performance model implements advanced features of the High Availability system. It ensures that there is no single point of failure in the application production environment. EPE-TP requires a minimum of fifteen (15) local cloud servers and one (1) remote server. This setup is very flexible and highly scalable. With primary (active) and secondary (passive) load balancers handling user requests, a background application checks the health of the primary load balancer and switches to the secondary load balancer on event of system failure. It is pertinent to note that very large eCommerce sites like Ebay employ eCommerce Production Environment model similar to the ePE-TP. Automated replication programs are used to spin-off new servers and add to the pool during peak periods and destroy some server instances when the traffic reduces. The ePE-TP could be scaled to several hundreds or thousands of cloud servers depending on traffic demands.

Fig. 9 - ePE-High Performance Model

Fig. 10a - ePE-Turbo Performance Model
4. Recommendations

The following Open Source Technologies may be chose of implementation of a robust eCommerce production environment:

- Linux Operating System (Ubuntu, Debian, Centos) - Operating System
- Apache/Nginx - Web Server
- Magento/Prestashop/OpenCart - eCommerce Framework
- Github - Revision Control
- Bacula - System Backup
- Percona XtraBackup - Database Backup
- Nagios 4 - Monitoring
- MySQL/MariaDb - Database
- BIND - DNS
- HAProxy - Load Balancing
- Puppet/Chef - Automated Scaling
- ELK Stack (Elasticsearch, Logstash,Kibana) - Centralized Logging

Other remote Service include

- Cloudflare - Web Application Firewall, CDN, SSL (free & Paid subscription)
- New Relic - Software analytics (paid subscription only)
- Anturis - Remote Monitoring (free & Paid subscription)

5. Conclusion

The implementation of a High Availability model in the production environment setup of an eCommerce application is very essential. Apart from ensuring data integrity, it builds client trust on the reliability of the service. Though high availability system may seem expensive, on the long run, it is cheaper as you may never be exposed to the crisis of data loss due to service disruption and other related hazards. The models highlighted above could be implemented according to the business needs and gradually scale to higher or better alternatives.

References


About the authors

David Alphonsus Inyang-Etoh holds a B.Eng degree in Mechanical Engineering and awaiting the award of M.Eng in Production Engineering. He is a seasoned programmer and web developer with experience in web application development and implementation of open source technologies. He has been researching on emerging open source and cloud technologies for development of ecommerce solutions, eLearning solutions and artificial intelligence for task automation in production and service delivery. He is currently serving as a Programmer at the Federal Polytechnic, Ukana, Akwa Ibom State.
Email: dinyangetoh@gmail.com
Improving the recording and reporting of facility-based mortality using open source mobile technology: Lessons from Cross River HDSS, Nigeria

Iwara Arikpo$^{1,2,*}$, Idongesit Eteng$^1$, Anthony Okoro$^2$, Uchenna Nnabuko$^1$

$^1$Department of Computer Science
$^2$Cross River HDSS, Directorate of Research & Quality Assurance
University of Calabar, Nigeria

*Corresponding Author: iiarikpo@gmail.com, iwara.arikpo@unical.edu.ng

ABSTRACT

Purpose: To improve the recording and reporting of facility-based deaths, using open source mobile information technology, in resource-constrained settings like Nigeria, where there is near-absence of efficient, robust and cost-effective electronic systems for the collection and timely reporting of mortality data.

Methodology: The study was carried out at the Cross River HDSS, a health and demographic surveillance research platform in Cross River State, southern Nigeria, between October 2014 and August 2015. The frontend mobile client for data collection, in XML data format, was implemented with ODK Collect and installed on Android Tablet PCs and smartphones, while the backend data management was implemented in ODK Aggregate server hosted on Google App Engine Cloud.

Findings/Results: The study found that, there were no robust, affordable electronic systems for tracking and reporting facility-based mortalities in ideal hospital settings in Nigeria. The system was piloted with 151 retrospective mortality records from the University of Calabar Teaching Hospital (UCTH) Calabar, Nigeria, one of the tertiary healthcare facilities within the coverage of the Cross River HDSS. Data was collected and uploaded to the web server; downloaded in CSV format, cleaned and analyzed using R Statistical Computing Software. The data collection and validation process was simplified with the use of dropdown and other select fields in the electronic formats.

Research limitations: This study is limited by the difficulty with which text can be typed using smartphones, especially on mortality issues, where sometimes a narrative on symptoms during illness may be informative.

Research implications: This study has demonstrated once again, the value of the free and open source software revolution to the deployment of information technology in organizations in resource-poor settings. The simplification of the recording and reporting of facility-based deaths can go a long way in informing the direction of necessary interventions in healthcare service delivery, through the availability of information on cause of death from health facilities.

Keywords: Open source software, mobile, facility-based mortality, tools, information technology, paper-based system.

Paper Type: Research Paper

1. Introduction

Mortality, migration and fertility are three essential indicators that give insight into the population status, distribution and structure of any nation. Mortality (death event) is inevitable and a one-time occurrence that often creates deep vacuum, difficult to fill in a family, community, organization and the society in general. The impact of this unfortunate event can be too severe and irreparable. Any nation desiring to sustain or improve the growth of her economy therefore, makes effort at ensuring minimal occurrence of this event, especially preventable deaths. One of the common ways of reducing death rates is by tracking all deaths and identifying the root causes in order to help plan interventions and prevent future occurrence from the same causes.

However, the difficulty of tracking all deaths arises from the fact that death happens anywhere. Usually, deaths that occurred in the community are termed community-based deaths while all deaths in formal health facilities are known as facility-based
deaths. Over the years, governments have developed measures for tracking death events, wherever and whenever they occur. For instance, the National vital registration system was solely designed among other objectives, to register all deaths, be it in the facility or community. Health and demographic surveillance systems are also platforms setup especially in low and medium-income countries (LMICs) to track deaths in communities.

Countries are often at crossroad in setting up these data collection platforms due to cost. Fully functional vital registration systems are systems that would have guaranteed efficient and complete collection of death records. However, high implementation costs often prevent countries from adopting its use. Salawu (2009) asserts that standard vital registration systems are only functional in few countries (wealthy countries) where there are resources and better infrastructure. For poorer countries, implementation takes close to a decade or more; a process which has forced hospital managements to continue the use of paper-based approaches in spite of the numerous limitations in paper-based data collection systems. Another approach for recording death events is through the implementation of some form of facility-based surveillance systems and programmes in health facilities. A typical example is the tracking of deaths in South Africa health institutions to assess the impacts of disease burden using programmes such as Confidential Enquiry into Maternal Deaths (CEMD), Peri-natal Problem identification Programme (PPIP), Child Healthcare Problem Identification Programme (Child PIP) and National Injury Mortality Surveillance System [NIMSS] (Joubert, et al., 2012). Some routine health management information systems have features for capturing facility-based deaths also. The DHIS2 tracker is one of such tools used for sharing basic clinical data across health facilities. It permits the collection, management and analysis of disaggregated data such as immunization, neonatal and maternal deaths audits (DHIS2, 2016).

The importance of reliable data on facility-based deaths need not be over-emphasized. Accurate recording and timely reporting of facility deaths are at first instance, legal proof of the death event. It provides formal backing and reference on the issues (place, time, cause of death, ailment leading to death, etc.) surrounding the death. More so, the information needed in understanding the etiology (study of the causes and origins of diseases) and its economic burden on the nation as provided through timely and accurate collection of this one-time event is used by governments and researchers in measuring mortality rates and trends within a fixed-time periods. Policies that can strengthen socioeconomic development and reduce the mortality rate of nations can be formulated based on the outcome of mortality data that have been reviewed and analyzed.

Fatality rates of life-threatening health issues and emergencies can be reduced or even prevented, if there is timely availability of data on such a condition. Examples can be drawn from the experiences in tackling the Ebola and Lassa fever epidemics which Nigeria, Liberia, Guinea, and few other countries suffered great loss. The timely detection and recording of the Ebola hemorrhage fever led to Nigeria’s success story in curtailing the disease from spreading.

Mortality information is crucial and needed by nations for effective national planning (population forecasting, social description), health sector planning (planning and development, evaluation of health services and programmes) and epidemiological and medical research interventions [epidemiological studies, and evaluation of medical procedures] (Hill, 1984). Accuracy in recording and reporting can strengthen the integrity of facility-based mortality data by eliminating errors arising from mortality bias and under-reporting. Besides, the accuracy of this data can assist health policy planners to implement policies and programmes that will reduce premature death and improve the quality of life. An implication of not having accurate recording of mortalities which is often as a result of omission of death events and errors in the coding of key variables like age and dates is seen in the inability of health policy planners to measure death rates, cause of deaths, sex- and age-specific mortality patterns and differentials in countries (Mathers & Boerma, 2010). Reliable data sources are therefore required for comparing death rates and trends among countries. Without accurate or timely collection of these data, results from such comparisons will be skewed or inconclusive.
1.1 Problem with paper-based approaches

Notwithstanding the numerous benefits of collecting and reporting facility-based deaths, the collection method matters, as this affects the quality, access, and completeness of such data. There are many problems inherent in paper-based approaches to collecting and reporting of facility-based mortality data. Some of these problems are difficulty in tracing and correcting probable mistakes once written and archived, inconsistent and/or misleading terminologies in the classification and reporting of symptoms and diagnosed cause of death among health facilities or individual physicians. For instance, a physician may classify cardiopulmonary arrest and chronic heart failure as symptom and cause of death, while another records renal disease and cardiovascular disease. This disparity in documenting the cause of death can mislead data analysts, thereby yielding erroneous statistics and decision-making on deaths (Mirabootalebi, Mahboobi, & Khorgoei, 2011).

A well-designed electronic mortality data collection tool can prevent these problems, by coding likely symptoms and causes as dropdown options, radio buttons and check boxes from which users can choose. Another solution would be enforcing data integrity through complex skip pattern and restriction of some sets of possible responses in the electronic form design (Tomlinson, et al., 2009). The inclusion of these form design features into clinical data collection tools has the potential to: reduce entry errors, reduce data collection time and improve the quality of mortality data collected. This is justified in the study by Paudel, Ahmed, and Pradhan, (2013) which assessed the benefits and challenges of using tablet personal computers (PCs) and wireless technologies in administering demographic and health survey in Nepal, South Asia.

Other problems associated with paper-based methods include: huge budgets on paper, storage space and filing infrastructures. There is also possibility of losing collected death records to natural disasters and accidents such as flood and fire outbreaks. Quick access to and retrieval of these records are also not possible. Researchers have also established that paper-based systems have no mechanism for ascertaining completeness, data integrity, accuracy and timeliness of data collection and entry (Douglas, et al., 2005; Njuguna, et al., 2014). This is a critical concern in the use of paper-based systems for facility-based mortality data collection. The use of information technology approaches is an optimal solution to these problems of paper-based data recording and reporting (Satterlee, McCullough, Dawson, & Cheung, 2015). The goal of this study is to contribute knowledge on techniques for improving the recording and reporting of deaths that happen in health facilities, using open source mobile information technology.

2. Literature Review

2.1 Paper-based approaches to recording and reporting of deaths

Reporting and management of facility-based mortality data can be done using different approaches. The predominant method used in many resource-constrained health institutions is the paper-based method. This method relies on paper forms, registers and books for recording hospital death information. With the establishment of the National Health Management Information System (NHMIS) by the Nigeria Federal Ministry of Health (FMOH) in 1991, health facilities have since employed the use of standardized paper-based Health Management Information System (HMIS) forms for facility data collection and reporting. HMIS is a mechanism for collecting and generating information required for operating health services and also for research and planning (Asangansi, et al., 2013). In terms of deaths, HMIS forms are used in health facilities to collect monthly summaries of facility-based data. Completed forms at health facility level are sent to the local government level for onward transfer to the state and then national. In Nigeria, the high cost of data aggregation and transmission using the HMIS forms led to the use of the DHIS2 software (HISP, 2016).

The software supports the electronic collection of HMIS data. Unfortunately, many health facilities are not able to migrate to the DHIS2 platform. Findings from the Nigeria Health ICT Phase 2 Field Assessment exercise conducted by the UN Foundation in collaboration with the Ministry of Health in 2014 showed that DHIS2 is operational in just a quarter of health facilities across selected representative clusters assessed (UN Foundation, 2015). This daunting result justifies the need for
alternative cost-effective platforms for collecting data on vital events like mortality.

2.2 Electronic approaches to recording and reporting of deaths

Having identified the problems of paper-based systems, health and information technology practitioners have, over the years, sought the use of electronic approaches as alternate solutions. In the early 1960s, hospitals in the United States used computer-assisted microfilm systems to index and store deaths and other vital events records for easy retrieval. Prior to the advent of this technology, vital information were stored in files in the form of ledgers (Logrillo, 1997). Electronic death registration system is another system developed for registering deaths. It is a secure system used in Washington to register and edit death certificates online. The web-based application allows reporting of deaths from health facility to local and national offices of government for management (Mirabootalebi, Mahboobi, & Khorgoei, 2011). Though it streamlines the recording process and reduces the time of data collection and filing, it does not support mobile devices and also requires users to have high-speed internet connection (New York State Department of Health, 2016).

In order to have detailed, complete and timely information on all violence-related mortalities among residents and non-residents, the United States Center for Disease Control and Prevention designed a surveillance system – National Violent Death Reporting System (NVDRS), which extracts information from crime laboratories, death certificates, medical examiner and law enforcement files and records (Paulozzi, Mercy, Frazier, & Annest, 2004). Notwithstanding the relevance of NVDRS in reporting and preventing violence related deaths such as suicides, homicides, deaths of undetermined intent, there exists crucial factors limiting its global adoption: (1) only a few low-, middle- and high-income countries have the required resources to implement the system, (2) collation of violence-related mortality data from different locations can only be done by a single agency with requisite technical competencies such as public health with vast knowledge in epidemiological surveillance (Butchart, 2006).

Africa is also not left out in the struggle to correctly use mortality information in assessing impact of health policies and planning. Electronic systems designed to capture comprehensive and timely data on mortality have been deployed in healthcare settings to boost the civil registration systems of some countries. In most cases, these systems are intervention-specific, hence cannot be used for the assessment of all death events in a health facility. For instance, the Malawian government in 2014 made concerted efforts at reducing maternal mortality through the implementation of a national Maternal Death Surveillance and Response (MDSR) system. The objectives of the project include: improving timely and quality reporting of all maternal deaths, systematize verbal autopsies and maternal death audits (Konopka, 2016). Obviously, it is pertinent to have a cost-effective, scalable and sustainable electronic system (preferably open source-based system) that is able to routinely report diagnoses and deaths information for administration and quality assurances purposes. This is because health facilities are the major source of mortality data needed for national and subnational policy and planning.

2.3 The use of mobile technology in recording and reporting of deaths

With the gradual transition from paper-based systems to a more flexible, efficient and cost-effective means of recording facility-based mortalities, the use of mobile technology has been developed and used in many settings (developed and developing countries) to enhance data collection and entry. Feedback from projects that were implemented using mobile-driven technology shows the overwhelming advantages of mobile technology. A review by Hall et al. (2014) on the impacts of using mobile technologies in Low-and-middle-income countries to measure common and crucial health outcomes highlighted these benefits. The use of mobile technology minimizes data entry errors and reduces the rate of data loss. It also strengthens data collection, tracking and reporting of vital events.

2.4 Available open source tools for recording and reporting of deaths

The adoption of open source tools in the development of data collection and management
systems by health institutions, especially in developing countries, is a low-cost technological innovation that has improved the standard of hospital data collection and reporting systems. This is practically due to the flexibility of adapting and extending existing open source software to suit hospital-specific requirements, without unnecessary duplication of effort, as such preventing the lock-in complexities that are concomitant with proprietary systems (Reynolds & Wyatt, 2011).

Though there are really no established literatures reporting the use of open source solutions specifically for recording and reporting of facility-based mortalities, different countries have built or adapted some medical and health records open source-based applications that incorporate mortality data capturing features. A typical example is the DHIS2 software which is used as the health management information system in many developing countries. DHIS2 Tracker is a module built into the DHIS2 software, which supports longitudinal tracking, aggregation and exporting of one-time events such as deaths and births into DHIS2 core application for validation checks, mapping and reporting (HISP India, 2016).

In a review conducted by Meystre and Müller (2005), the authors outlined several customizable open source software presently in use in most health institutions around the world. Examples include: Care2x, OpenVistA, FreeMed, OpenEMed, OSCAR McMaster and OpenEMR. Results from a similar study conducted in September, 2010 by the National Opinion Research Center (NORC), Chicago on the use of open source electronic health record systems in medical institutions also demonstrated the robustness and flexibility of using open source solutions in healthcare settings. It was reported that open source medical software such as VistA electronic health record, OpenMRS, MedLynks and ClearHealth were customized and implemented by different health institutions to address the specific requirements of these institutions. (Goldwater, et al., 2013). The concern would be on the cost of hardware, ancillary software, technical expertise and other resources needed for the successful implementation and continuous support of the chosen open source solution. We addressed this concern by implementing a facility-based mortality data collection system using mobile tools built on simple free and open source software technology.

3. Methodology

3.1 Study Setting

The Cross River health and demographic surveillance system (Cross River HDSS) is a research platform operating two cohorts located within the southern senatorial district of Cross River State in south-south Nigeria. The two cohorts have a combined population of 33,446 persons in 8,508 households (48% of which are rural dwellers) continuously under surveillance. The rural cohort is located in the Akpabuyo Local Government Area (LGA) of the state and the urban cohort located in Calabar-Municipality; both in the southern part of the state. Data on community-based deaths is collected 4-monthly in each of these cohorts using verbal autopsy procedures (INDEPTH Network, 2008). However, there is a secondary and tertiary health facility respectively in the rural and urban cohorts, where facility-based deaths are recorded. There is the St. Joseph’s Hospital in Akpabuyo and the University of Calabar Teaching Hospital in Calabar-Municipality.

3.2 Conceptual Design

Figure 1, describes the conceptual design for collecting and reporting facility-based deaths. At one end, you have the end-users who are the health workers (HWs) using either mobile (Android) smartphones or web-forms (Internet-connected computers using HTTP Protocol) to collect and send mortality data from a health facility. This data is originally recorded on hospital registers. At the other end, there is a web server implemented on ODK Aggregate (Open Data Kit, 2016), hosted with cloud technology (Google Cloud Platform, 2016). The web server hosts XML-formatted (blank) electronic mortality forms, as well as receives data from the users (health workers) over GSM or Internet (web form HTTP) upload. This data is first saved in the cloud-based web server. Later, and when appropriate, technical/admin users download the data into different formats, (particularly CSV) for analysis and reporting. XML (extensible markup language) is a kind of markup
language used by application designers to define a set of rules for encoding documents in a format which is both human-readable and machine-readable (Quin, 2015).

Figure 1: Conceptual design for the mortality surveillance system

3.3 Use Case
In software systems design, a use case is a list of steps, typically defining interactions between a role (known in unified modeling language [UML] as an “actor”) and a system that supports a particular business goal. The actor can be a human, an external system, or time that triggers the use case. Use cases are used to show a system’s context and functionality (Booch, et al., 2007). The use case for this study is shown in Figure 2.

Figure 2: Use Case diagram for the Mortality Surveillance System
From the Use Case diagram of Figure 2, the following actions take place in the system:

- **Create Form & Upload:** The technical/admin user creates the mortality surveillance data collection forms in XML-compatible format and uploads to the ODK Aggregate Web Server.
- **Download Form:** The health worker, using an Android mobile phone configured to point to the ODK Aggregate Web Server, downloads the blank mortality form into the phone.
- **Fill & Upload Form:** The health worker fills the form with individual mortality episodes. S/he can view and edit the entries for the form fields before uploading the form to the ODK Aggregate Web Server.
- **View Submitted Data:** The technical/admin staff logs onto the ODK Aggregate Web Server to view the submissions from health workers who submitted data from different remote locations.
- **View Form & Form List:** Both the health worker and technical/admin staff can view the list of available forms for data entry. While the health worker views the forms list with his/her mobile device, the technical/admin user can also view the forms from the Web Server, including forms that are not yet available for use by the health worker.
- **Download Data for Analysis:** The essence of every data collection exercise is to report findings from the data. After submissions, the technical/admin staff can download the data from the ODK Aggregate Server, mostly in CSV format, for analysis and reporting. Data can also be downloaded into KML format for geo-mapping of health facilities.

4. Implementation

4.1 Results

The implementation of this study was piloted at the University of Calabar Teaching Hospital (UCTH) Calabar, which is one of the health facilities collaborating with the Cross River HDSS. Android Tablet PCs and Android smartphones installed with ODK Collect v1.44 were used to capture and transmit retrospective mortality data from the Records Department of UCTH to ODK Aggregate server v1.4.4 deployed on Google App Engine Cloud. The data was exported into CSV (comma separated values) format. Data was analyzed using an open source software, the R Statistical Computing Software v3.3.1 (The R Foundation, 2016). Figures 3 – 5 show some of the results from the analysis.

![Figure 3: Proportion of deaths by age group and sex](image)

Figure 3: Proportion of deaths by age group and sex
4.2 Discussion

The use of electronic formats to capture mortality records made data input simpler and faster, as about 95% of the input fields are mere dropdown lists from which choices could be made. This reduces, to the barest minimum, errors due to user text input, thus enforcing some level of system input controls. The ability of the ODK technology to export data in compatible formats, like CSV, facilitates the analysis and reporting of data using any analysis software (Creativyst Software, 2010). The availability of data in the cloud enhances access and facilitates the ability to generate reports even from remote locations. The reports shown in Figures 3 – 5 of this study are seldom promptly available in most health facilities in Nigeria, due to reasons ranging from inconsistent data formats based on paper storage, to lack of resources to implement information technology solutions. The use of free and open source tools takes away costs associated
with software and enhances interoperability (Walli, Gynn, & Rotz, 2005; Shaikh & Cornford, 2011).

5. Conclusion

This study has demonstrated that, the lack of availability of mortality data from health facilities due to poor recording-keeping from paper-based approaches can be reduced to the barest minimum, when electronic data collection and reporting tools are introduced. It has also shown that, the use of free and open source technology will ensure affordability and enhance interoperability when these solutions interface with other systems in these facilities. The prompt availability of information on facility-based deaths (especially on cause of death) can go a long way in assisting health policy and interventions by hospital administrators and other stakeholders in the healthcare sector.

The merits notwithstanding, the introduction of electronic data collection and reporting method as presented in this study is limited by the difficulty with which text can be typed using smartphones, especially on mortality issues, where sometimes a narrative on symptoms during illness may be informative. However, these narratives can be captured as photo inputs to form part of the individual record during data collection. Besides, there are other facility-based electronic data reporting methods, like DHIS2 that reports aggregated data. Future research could explore the integration of the mortality surveillance system with the DHIS2.

References


LEARNING PATTERN DISCOVERY: IMPACT OF USER-CENTRIC DESIGN APPROACH TOWARDS ENHANCEMENT OF E-LEARNING SYSTEMS

Kingsley Okoye
School of Architecture Computing and Engineering
University of East London
Docklands Campus, 4-6 University Way, E16 2rd, London, UK
U0926644@uel.ac.uk

ABSTRACT

Purpose: The purpose of this paper is to propose a system for automated learning directed towards discovering and enhancement of the set of recurrent behaviours that can be found within a learning knowledge base and how they can be modelled to enable a more effective reasoning and tactical strategies for personalized adaptation and decision making.

Design/Methodology/Approach: The paper introduces a methodology that makes use of information readily available within a learning process to explore and analyse the data to obtain inference knowledge capable of enhancing individual learning performance. The author evaluated this approach through a user-centric design prototype and a model developed using Business Process Model Notations to corroborate the focus on improving performance of e-learning systems.

Findings: The technological response to satisfying the increasing demand for richer and more precise depiction of e-learning applications capable of providing platforms for pattern exploration where users can browse for knowledge they might consider as interesting is by providing a personalized adaptive learning system for the users. Adoption of such technological developments will spark a great success for many learners and needed to provide continuous intelligent recommendation, guidance and feedbacks on learner’s performance especially in achieving the overall learning objective.

Research limitations/implications: To meet the overall 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 do not happen automatically. Such process 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).

Practical implications: The work in this paper considers the implication of the user centered design approach, and to this end, identify some of the common design problems with e-learning systems as means to revealing the implications for designers to stick to user-centric standards when developing automated learning systems in order to ensure learners satisfaction and reliability.

Originality/value: The approach and prototype described in this paper provides an enhanced model for learning which is useful towards the development of e-learning systems that are more intelligent, predictive and robotically adaptive, which also aid in discovery of new learning patterns and enhancement of existing learning processes.

Keywords: Automated learning, pattern discovery, learning data, user profiles, e-learning system.

Paper Type: Research Paper

1. Introduction

E-learning systems should dynamically support different adaptive learning paths and contents to fit learner’s diverse needs based on the user’s 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). E-learning systems should enable users to create, share and collaboratively edit contents to suit their individual learning needs and styles. 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 et al, 2011). Recent challenges in literature shows
that contemporary tools for collecting, modelling and analyzing of captured learners data are proving to be more complex in nature due to the increasing demand for systems capable of providing adaptive platform for pattern exploration where users can browse for knowledge they might consider as interesting and useful towards achieving their learning goal (Kriegel et al, 2007). This growing complexity is proved by the need for richer and more precise description of real-world objects and depiction of systems that allows for flexible exploration of learning patterns/data types. According to Kriegel et al (2007) future development will be to find richer patterns by developing systems which derive understandable patterns as well as making the discovered patterns explicable.

The current innovations and challenges is what motivates software engineers in recent years to provide personalized adaptive e-learning systems which takes the advantage of the readily available learning data, as well as how to explore and analyse the captured data in order to provide inference knowledge capable of enhancing individual learning performance. 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.

The work in this paper focuses on exploring some of the common design problems with e-learning systems performance and user requirements that needs to be solved in order to provide a user-centred experience, and effective use of such systems. We reveal the implications for developers to stick to certain design standards when developing e-learning applications and then propose an adaptive learning model to help address some of the design problems with present systems in order to improve learners’ engagement and outcome on performance. This paper further provide road map for future improvements. We look at the determining factors that impact the development of an enhanced personalized adaptive learning systems based on three probes - Accessibility, Presentation Design, and Standard Compliance - bearing in mind that if e-learning systems developers do not take into consideration these factors along with the user experience, most developed systems fail desolately and users tend to turn away from the readily available products looking for easier-to-use and intuitive e-learning platforms.

In this paper, we propose a learning model directed towards discovering and improving the sets of recurrent behaviours that can be found within a learning executing environment. The proposed model is developed in order to address the problem of determining the presence of different learning patterns within a learning knowledge base and how they can be modelled to enable a more effective reasoning and tactical strategies for personalized learning adaptation and decision making. The standpoint for our approach is based on the following objectives:

- To show how user data from learning processes can be extracted and transformed into useful harmonization for improved learning performance.
- Provide a more enhanced model for learning which is useful towards the development of e-learning systems that are more intelligent, predictive and robotically adaptive, which also aid in discovering of new and enhancement of existing learning processes/pattern.

The focus is on identifying data about different users within a learning knowledge base and enriching the information values of the resulting model based on the captured user profiles.

The rest of the paper is structured as follows: in the next sections, we identify some of the key design criteria considered 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 user-centric adaptive learning system. Next, we propose an automated learning system that is capable of detecting changing trends in learning behaviours and abilities of users in order to help address some of the design problems with present e-learning systems. In addition, we discuss the importance of the described design methods and stipulation towards the achievement of an enhanced user performance and effectiveness in e-learning systems. In the last section, we discussed and analysed appropriate related works in relation to this area of research, followed by the conclusion and road map for future research.
2. E-Learning Design Practice and Criteria

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

2.1. Presentation Design

The quality of the components embedded in an e-learning system is highly influenced by the way the contents are designed and presented (Leacock and Nesbit, 2007). Design presentation 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 are expected to incorporate aesthetical design values and development of learning information in formats that are efficient, consistent with the fundamental essence of research, and knowledge of educational multimedia products by demonstrating standards in the development of e-learning applications. The principle 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 e-learning products (Tufte, 1997). Further, findings in support of Tufte’s 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 and 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 deployment, and is expected to efficiently contribute to intellectual and learning development. E-learning systems that have poor presentation design can result in extraneous information representation, which tends to reduce the quality and amount of information available for learners’ understanding.

The International Standards Organisation (ISO) are engaged in developing a new standard called the ISO/AWI 23973 "Software ergonomics for World Wide Web user interfaces" (Bevan, 2005). The organization 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 have been focused 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.

To this end, we propose some design principles that can serve as an effective guide to minimizing the extraneous information content through:

![Fig 1. Presentation Design Framework for Human-System Interaction](image-url)
• 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, 2014a).

In Fig 1 we propose a design framework which is useful for software developers to guide their design towards an improved performance in e-learning applications.

There is strong evidence that presentations which combine graphics with text, most of the time, tend to intuitively ease and improve learning when compared with text only designs (Saldanha et al, 2013). Such design approach 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. Designers have to understand the implication of graphical representation in presenting speech notions so prominent from their more self-explanatory use in presentation of useful data and informational intellectual maps consisting of nodes, connected by links captioned with relational terms serves as powerful projections or alternatives to presentations with text only. Such systems need to be designed in a way that makes them intuitive or comprehensive enough to impact knowledge or intellect, but concise enough to be able to allow flexibility and consistency in usage.

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. We concur 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. Take for instance, rather than explaining how to clone a computer system by text only, why not have a step by step animation that shows how all the components fit together? Also, why not enable the user to click on the animated components to find out their part names and components? Similarly, an electronic encyclopedia might use many words to describe the sound that parrots make without really communicating those sounds, but enabling the user to click buttons to hear the sound is a huge benefit.

2.2. Accessibility Design

An indispensable user factor to consider when building e-learning systems is Accessibility. This is an important feature as learners will not be able to interact and complete tasks if they find it difficult to use the learning platform. Studies have shown that many e-learning applications do not provide fully sufficient levels of accessibility to the users (Okoye et al, 2014a; Saldanha et al, 2013; Ali et al, 2008). These studies suggests that designing and developing a reasonably accessible e-learning 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 interfaces 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 include 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 (2014a) 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. An example is 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 (2006) provided guidelines for developing accessible e-learning platforms. They stated specifically that, in dealing with e-learning applications, these guidelines address 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 require a detailed understanding or knowledge about these standards and guidelines.

The International Standards Organisation (ISO), in part of ISO 9241 listed in Bevan (2005) and Travis (2009), provided 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 e-learning product with the intended user group.

From all standard arrangements, 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 educational background, 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.

2.3. Standard and Compliance Design

In e-learning application design, standards promote 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 to proper and ethical design than simply making use of standards. In spite of all ideas, usability standards are still critical and plays an essential role in the development of intuitive e-learning systems 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 International Standard Organization 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 e-learning system? 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
To accomplish a high level of standard compliance towards solving these common problems, designers of e-learning applications should stick to all applicable standards and should make resources available for the intended users. The quality of how information is described and how it closely conform to the learning objective is critical in assisting learners advance from searching to discovering (Duval and Hodgins, 2006; Okoye et al, 2014b). 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 means of information representation 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.

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 e.g. “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 system designers to procure a fit for purpose user-centered planning and instructional design. We opine that designers must understand who their targeted audience are before moving into their design because many learners tends to turn away from some e-learning systems if they find it difficult to intuitively use the platform.

Undoubtedly, it may be seen that without clear standards by which usability in e-learning system 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 e-learning technologies. Confidently, with the wide spread of the rate at which these evaluation standards would be adopted for design, we believe a large scale in proportion to e-learning systems will exhibit rich usability and prove effective for learning.


In this section of the paper, we propose a personalised adaptive learning system that is expected to collect routines and monitor changes in user’s behaviour during a learning process. This is to determine which adaptations technique is suitable or may be required progressively through time. The learning model is expected to take into account users profile i.e. prior knowledge of learners’ background, learning behaviour and actions when using the system. The system should also be capable to dynamically update the representation of users performance taking into account the changing state of the leaners and the variations in information that is relevant to each user considering 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). According to Huang and Shiu (2012) and Nganji et al (2011), the key challenge in developing automated system for learning is to build effectively represented user profile, learning styles and goal to help support reasoning about
each learner. Our proposed model is an automated learning system that is capable of detecting changing trends in learning behaviours and abilities of individual users. The goal is to discover user interaction patterns, and respond by making decisions based on adaptive rules centred on captured user profiles.

The approach described in this paper focuses on the personalization of learning contents for the users which makes it possible to efficiently generate learning patterns based on the sequence or control-flow of each individual learning patterns/behaviours. The proposed approach uses process adaptation and discovery technique (Okoye et al., 2014b; Van der Aalst, 2011) to allow for traces not present in an existing learning path 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. According to Han et al. (2008), Decision Tree Learning (DTL) is one of the scientifically proven technique that focuses on the classification of activities within a learning knowledge-base to predict patterns based on discoverable variables.

The technique uses uncompromising response variables by classifying the learning activities and arranging the resulting value in form of a tree. The process consist of nodes that correspond to the possible values (leaf nodes) and the predictive variables (non-leaf nodes). Each learning concept (class) within the tree splits a given set of node into two or more subsets (sub-class). In essence, all instances within the tree are represented as a sub-class of a domain class referred to as root node.

Based on the attribute of the activities within the domain class of the learning model tree, the learning concepts splits into ones that are leaf nodes and ones that are non-leaf node. Activities within the learning model are classified according to their corresponding attributes by ontologically describing the domain concepts, subclasses and the associated instances of the various learning concept classes. Attributes may appear multiple times within the tree but not twice on the same path. This is aimed to handle values based on supposed semantics and to ensure that no class can have the same instances, as we show using the Knowledge Interchange Format (KIF) syntax below:

\[
\forall ( ?X ?Y ) \\
\rightarrow ( \text{Class } ?X ?Y ) \\
( \not ( \text{Instance1 } ?X ?Y ) \\
( \text{Instance2 } ?X ?Y ) ) )
\]

Decision tree learning uses a recursive top-down algorithm expressed in terms of the root node, r, and all associated instances to the root node. Where \( x = \{r\} \), the set of nodes to be traversed (Van der Aalst, 2011; Han et al., 2008). Thus

\[
\begin{align*}
\text{IF } x = \emptyset & \text{ THEN END} \\
\text{Else} & \\
\hspace{1cm} \text{// Select and extract all the subset of x based on Entropy} \\
X: & = X/\{x\}, \text{ where } x \in X \ (x \text{ is a subset of } X \text{ based on entropy}) \\
\hspace{1cm} \text{// Check if splitting is possible?} \\
\hspace{2cm} \text{IF } X: & = \emptyset \text{ THEN END} \\
\hspace{2cm} \text{Else} \\
\hspace{3cm} \text{// create a set of Child nodes Y and Add} \\
\hspace{3cm} \text{Y to } X \\
\hspace{3cm} Y: & = X \cup Y \\
\hspace{3cm} \text{// and Connect x to all Child nodes in Y} \\
\end{align*}
\]

Fig 2(a). Decision Tree Learning Algorithm.

Fig 2(b). Top-down Flowchart of the Decision Tree Learning Algorithm

As shown in Fig 2(a) and 2(b), one of the basic function of the Decision Tree Learning algorithm is to help designers to define how the sequence flow of the learning activities split and also decide when to stop adding nodes.
This function is achieved by using the //Check if splitting is possible? IF \( X = \emptyset \) THEN END, Otherwise, the nodes continues to split through the enabling function; Else //create a set of Child nodes \( Y \) and Add \( Y \) to \( X \), \((Y: = X \cup Y)\), until the value of \( X: = \emptyset \).

The algorithm is beneficial in improving the resulting nodes or to restrict the learning tree to a certain level. It is necessary in order to determine the variation of the smallest unit (node) within the learning activity sets by splitting the set of activities into subsets using the idea of Entropy.

\[
E = \sum_{i=1}^{K} P_i \log_2 P_i. \tag{1}
\]

The more we split a node the lower the entropy until the overall root node reaches a definition value equals to zero. Entropy (eqn. 1) provides itself as a measure that is used to define and quantify the diversity in a leaf node to determine if splitting is possible or needed. The decision tree learning algorithm is useful for Learning Pattern Discovery to help locate all decision points in the process model, paths taken, and the attributes of learning activity sets known at (or prior) the decision point. The 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 such approach as described in this paper 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 and a more enhanced system for learning capable of increasing learners’ performance or goal in an unswerving path as shown in Fig. 3.

In Fig. 3 we show that modelling of data about the learning process is an important tool towards unlocking the information value of the various activities within the learning knowledge-base, by way of finding useful and previously unknown links between the activity concepts. The motivational perspective is the search for explanatory and predictive patterns within the learning process especially with regards to the large volume of users that are involved. Martin and Majidian (2013) refers to this tactics as Creative Knowledge Discovery which is concerned with the creation of new and effective patterns either by generalization of existing patterns or by analogy to patterns embedded in other domains. According to Martin and Majidian (2013), an important prerequisite for the approach is that we understand the relations within the learning data, thereby allowing us to find paths that are hidden in the knowledge base and to extract novel patterns that can be utilized for subsequent modelling and enhancement of the learning system.

The ability to analyse the learning information and create concepts is fundamental to representation and modelling of the various
learning activities and paths that users follow. This techniques can be applied towards automation of learning processes and the extraction of useful models, as we describe below:

- Create the Learning concepts that will be harmonized within the Learning Model.
- Provide Process Descriptions and pathways to accommodate all user Object and Data Types.
- Create the gateways and conditions for activity sequence flows to map the user objects with the Learning concepts.
- Check for Consistency of prediction, monitoring and recommendations for all defined Learners within the Model.

4. Related Work

There has been a combination of factors that affect learners’ experience with e-learning systems in recent years which include the level of efficiency in use and effortlessness learning, lack of satisfaction with the system, confusing navigation, slowness in loading and download time, frequency of errors and difficulty for first time users in learning the interface quickly. There have been recommendations 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. Accordingly, Okoye et al. (2014a) proposed a theoretical method towards a rich and enhanced usability in e-contents and validated the 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 no-frills 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).

There is evidence in the design of e-learning platforms 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 to achieve an intuitive, in-depth and effective use of learning technologies. The work in Okoye et al. (2014a) puts forward a novel framework for practice that enables a deeper understanding of accessibility requirements for dynamic web applications centred around 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 and applications with accessibility issues in mind as well as considering the user groups.

Niu and Kay (2010) mention that an important pragmatic concern is that it must be inexpensive to create e-learning systems and its presentations for different learners. The authors argue that e-learning platforms should dynamically update the representation of knowledge to take into account the changing state of the learners and changes in the information that is relevant to each user over time. They further suggest that such a 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 Okoye et al. (2014b, 2015, 2016) the authors used the idea of process mining augmented with semantic reasoning and concepts 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 knowledge-base. 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 Logics and programming 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.

According to Huang and Shiu (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 propose 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.

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 or a BPMN model and describes causal dependencies between learning activities (Van der Aalst, 2011; Fahlad 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; Munoz-Gama and Carmona, 2011; Rozinat and Van der Aalst, 2008). 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 Miani and Junior (2015) most of the existing techniques for analysing large growing knowledge bases such as Learning process base focus on building algorithms to help the knowledge-base automatically or semi-automatically extend. The authors note that the use of an association rule mining algorithm to populate knowledge base and to improve the relations between the various users within the knowledge base is a useful approach considering the fact that most systems constructing large knowledge bases continuously grow without containing all facts for each category, resulting in missing value dataset. To resolve this challenge, the authors developed a new parameter called Modified Support Calculation Measure which generates new and significant rules. They also developed a structure, based on pruning obvious item sets and generalized association rules which decreases the amount of discovered rules in order to help maintain the large growing knowledge base and rules.

In Okoye et al (2016), we mention that Association Rule Learning aims at finding rules that can be used to predict the value of some response variables that have been identified as being important but without focusing on a particular response variable. This association aims at creating rules of the form: If X Then Y, where X is often called the antecedent and Y the consequent. Thus, X ⊃ Y. According to the work in Okoye et al (2016) we show that this rule is similar and can be related to the Semantic Web Rule Language, SWRL, which is a useful language designed for process description especially to provide an improved learning ontology and enhancement of the learning process model. The SWRL rule has the form: \( atom \wedge atom \ (\text{antecedent}) \ldots \rightarrow atom \wedge atom \ (\text{consequent}) \). Association rule learning strongly supports the use of such metrics frequently expressed in the form of support and confidence.

These expressions help in measurement of the strength of the association between learning objects. Support determines how often a rule is applicable to a given data set which means the fraction of instances for which both antecedent and consequent hold. Hence, a rule with high support is more useful than a rule with low support. A rule that has low support may occur simply by chance and is likely to be irrelevant from a learning perspective because it may not be profitable to monitor, recommend and promote learning activities or patterns.

Elhebir and Abraham (2015) notes that pattern discovery algorithms use statistical and machine-learning techniques to build models that predict behaviour of captured data. According to the authors, one of the most pattern discovery techniques used to extract knowledge from pre-processed data is Classification. They observe that most of the existing classification algorithms attain good performance for specific problems but
are not robust enough for all kinds of discovery problems. The authors in Elhebir and Abraham (2015) propose that a combination of multiple classifiers can be considered as a general solution for pattern discovery because they obtain better results compared to a single classifier as long as the components are independent or have diverse outputs. The approach compares the accuracy of ensemble models, which take advantage of groups of learners to yield better results using the Meta Classifier (Staking and Voting) alongside other Base classifiers: Decision Tree algorithm, k-Nearest Neighbour, Naive Bayesian and BayesNet. Explicitly, the problems of modelling learning processes can be solved by transforming ontology population problem to a classification problem where, for each entity within the ontology, the concepts (classes) to which the entities belong to have to be determined i.e, classified (d’Amato et al, 2008; Okoye et al, 2014b, 2016).

The approach described in this paper builds from the related works in several aspects. First, we provide an automated learning system that ensures performance through user-centric design. We focus 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 supports e-learning processes bearing in mind the design factors that allow for usability and performance from three perspectives; 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 also expected to be useful in providing a learning path and guidance based on individual differences. This is achieved by collecting user’s initial capabilities and preferences on interaction and then determining which adaptations or further assistive measures are best suited or may be required through time. We achieve this by proposing a user-centred learning model which is useful in design, discovering and enhancement of the set of recurrent behaviours that can be found within a learning knowledge base and how they can be modelled to enable a more effective reasoning and tactical strategies for personalized adaptation and decision making.

5. Conclusion and Future Work

The work in this paper reveals some of the key design criteria necessary towards achieving a user-centered experience in e-learning systems and provide measures that are of relative importance when considering the performance and usability of e-learning systems. The work shows that the technological response to satisfying the increasing learning demands is by providing a personalized adaptive learning system for the users. To this effect, adoption of the new technological developments will result in a great success for many learners. It is therefore, of great significance for e-learning 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 reliability. To prove the mentioned practice, we propose an adaptive e-learning model 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 the overall learning objective.

Future work could focus on extending the approach described in this paper by applying the technique to a different process domain. This will help in analyzing the streams of activities that are involved in the process in order to produce inference knowledge, which can then be used to load a more enhanced model within the process domain.

Acknowledgments

The Author is most grateful to the tutors and supervisory team at the School of Architecture Computing and Engineering, University of East London, UK, for their wonderful supervision and support.

References


About the author

KINGSLEY OKOYE received his Bachelors degree in Computer Science from Nnamdi Azikiwe University, Nigeria in 2007, a Masters in Technology Management at the University of East London, UK in 2011. He is currently a PhD candidate in Software Engineering at the School of Architecture Computing and Engineering, University of East London. He is MIET member at the Institution of Engineering and Technology, UK. Graduate Member at the Institute of Electrical and Electronics Engineers, IEEE. His Research interests includes Learning Process Automation, Process Mining, Semantic Technologies, Process Modelling, Internet Applications and Ontology. Kingsley can be contacted via his e-mail; u0926644@uel.ac.uk.
Table of Contents

AN ANFIS-BASED FRAMEWORK FOR DETECTING DDOS ATTACK
Gabriel O Ogunleye, Stephen G Fashoto and Teresa O Ojewumi..............................................................1

DEPLOYING ECOMMERCE SOLUTIONS WITH CLOUD AND OPEN SOURCE TECHNOLOGIES:
HIGH AVAILABILITY APPLICATION MODELS
David Alphonsus Inyang-Etoh....................................................................................................................15

IMPROVING THE RECORDING AND REPORTING OF FACILITY-BASED MORTALITY
USING OPEN SOURCE MOBILE TECHNOLOGY: LESSONS FROM
CROSS RIVER HDSS, NIGERIA
Iwara Arikpo, Idongesit Eteng, Anthony Okoro and Uchenna Nnabuko ..............................................27

LEARNING PATTERN DISCOVERY: IMPACT OF USER-CENTRIC DESIGN APPROACH
TOWARDS ENHANCEMENT OF E-LEARNING SYSTEMS
Kingsley Okoye..............................................................................................................................................37