African Journal of Computing & ICT

www.ajocict.net

African Journal of Computing & ICTs

Volume 5. No. 1. January 2012

www.ajocict.net

African Journal of Computing & ICTs

Volume 5. No. 1. January 2012

www.ajocict.net

A Journal of the Institute of Electrical & Electronics Engineers (IEEE)
Computer Chapter Nigeria Section

ISSN- 2006-1781

IEEE African Journal of Computing & ICT
Volume 5, No. 1. January, 2012

ISSN 2006-1781

CONTENTS

Editorial Board

Editor-in-Chief's Introduction

Disseminating Indigenous Computing and ICT Research & Development In Africa

Computing and ICT no doubt form important pillars in the present day global modern technological processes and development. Although Africa can be said to have improved substantially in the quest for knowledge in this age, much structural advancement and progress in the larger part of the continent have not been noticed. This can be partly ascribed to the relatively meager investment and financial commitment to providing infrastructure for sustainable research. The effect of this is that the technological state and needs of the general African continent are not the same as those of the developed world. As such, research output which may be important for the present socio-econo-technological state of the continent may not quite appear useful in the developed world, and are thus usually not considered fit for their journals. Apart from this, there are relatively few journals based in the African continent that are devoted exclusively to papers on computing and ICT.

It is in the light of the above that the *African Journal of Computing & ICT* is born for the purpose of publishing nontrivial relevant research results in the fields of computer science/engineering, information technology(IT) and allied fields. Every submitted paper is multiple blind-reviewed for quality, relevance, depth and accuracy. We are committed to excellence in publishing and desire that the Afr, J of Computing and ICT will be a prime avenue for the dissemination of cutting edge research report by Africans, for Africa and all lovers of research and development all around the world. We intend to promote indigenous Computing and ICT development through the dissemination of cutting edge research and development report with a view of reducing the heavy dependence on importation of ICT-related products. The journal welcomes papers from African scholars and also from non-Africans whose papers address important issues that are relevant to Africa.

Dele Oluwade, Senior Member IEEE
Professor & Dean of the College of ICT
Salem University
Lokoja
Nigeria

E-mail: bamideleoluwade@computer.org

Preface to Vol. 5. Issue 1

This volume of the African Journal of Computing & ICT contain six (6) articles that have been subjected to rigorous peer review by experts in the subject domain. The articles articulate issues of Research and Development related to programmable gate arrays, congestion control in Mobile networks, students' learning and performance, boi-informatics and cyber security both from a technical and social perspective (examining the impact of cyber crime on E-banklng using social theories). They definitely make for insightful reading.

In the first paper V.A. Akpan opined that the evolution of field programmable gate arrays (FPGAs) as custom-computing machines for digital signal processing (DSP) has led to the emergence of several modeling and design methodologies among which are the register transfer level (RTL) and electronic system level (ESL). Thus, the choice of FPGA selection has become a challenging task. The paper then discuss industry-standard FPGA system design methodologies; the modeling and development of FPGA-based embedded systems design based on the ESL design and verification methodologies from a higher level of abstraction view point rather than RTL; and presents a complete framework for an embedded processor system design, synthesis, implementation, simulation and verification targeting an FPGA.

In the second paper, Mughele et al attempted to determine the causes of congestion and evaluated congestion control mechanisms adopted by service providers in Nigeria using MTN Nigeria as a case study. A survey research design methodology using stratified random sampling technique was used in conjunction with automated data from MTN's data logging. Findings from the study showed that apart from the carrying capacity of the MTN network other factors such as the use of phones for data transfer and multimedia activities contributes to traffic congestion on the network. Recommendations were made based on these findings.

In the third article, Akinola et al characterize educational attainment as being related to discovering knowledge for predicting students' performances. This position was substantiated by mining undergraduate Computer Science students results obtained in some first year 100 level courses. Result from the study shows that a priori knowledge of Physics and Mathematics are essential in order for a student to excel in Computer Programming. The research has very strong implications for students' advising and academic performance related issues.

Makolo et al in the fourth article carried out a comparative analysis of the similarity check mechanism used in the most effective algorithm for mining simple motifs GEMS (Gene Enrichment Motif Searching) and that used in a popular multi-objective genetic algorithm, MOGAMOD (Multi-Objective Genetic Algorithm for Motif Discovery). Their research attempt to empirically proof the high sensitivity of the resulting algorithm, STGEMS in mining motifs from challenging sequences such as the *Plasmodium falciparum structure*.

Fadiora et al in the fifth paper employed pattern recognition in combination with Object Character Recognition (OCR) mechanisms to analyze and classify image-based spam e-mails. The resulting system is an enhancement of existing OCR systems that suffers from the problem of noise generated around spam messages embedded in images.

Wada and Odulaja assessed cyber crime and its impact on the banking institutions in Nigeria viz –a-viz existing policy framework and institutional countermeasures in combating cyber crime in the banking industry. This papers X-rays cybercrime policy issues and provide insight into how cybercrime impacts on E-banking from a Nigerian perspective. Social theories were then used to explain causation with a view of guiding policy makers on behavioural issues that should be considered when formulating policies to address cyber criminal activities in Nigeria.

We appreciate and welcome feedbacks, comments and rejoinders

With very best compliments

Longe O.B (PhD)
Managing/Production Editor

Model-Based FPGA Embedded-Processor Systems Design Methodologies: Modeling, Syntheses, Implementation and Validation

V.A. Akpan
Department of Physics Electronics
Federal University of Technology
Akure, Nigeria
avincent@ieee.org

ABSTRACT
The evolution of field programmable gate arrays (FPGAs) as custom-computing machines for digital signal processing (DSP), real-time embedded and reconfigurable systems development, embedded processors, and as co-processors for application specific integrated circuit (ASIC) prototyping has led to the emergence of several modeling and design methodologies among which are the register transfer level (RTL) and electronic system level (ESL). Moreover, due to the vast number of FPGA manufacturers and third-party partners support tools; the choice of FPGA selection can become a challenging task. This paper: 1) discusses industry-standard FPGA system design methodologies; 2) discusses the modeling and development of FPGA-based embedded systems design based on the ESL design and verification methodologies from a higher level of abstraction view point rather than RTL; and 3) presents a complete framework for an embedded processor system design, synthesis, implementation, simulation and verification targeting an FPGA. Finally, a brief conclusion is given with some discussions on future directions.

Keywords— AccelDSP, ChipScope Pro, embedded development kit (EDK), embedded system; electronic system level (ESL), field programmable gate array (FPGA), ModelSim, PlanAhead, register transfer level (RTL), System Generator for DSP, Xilinx integrated software environment (ISE), Xilinx platform studio (XPS), Xilinx software development kit (Xilinx SDK).

1. INTRODUCTION

AN embedded system is one with a built-in or embedded processor or computer, typically for carrying out some kind of real-time applications and should have some kind of connection to the outside world. It is the application rather than the hardware itself that defines the embedded system. Embedded Systems have been around since the early days of computers.

African Journal of Computing & ICT Reference Format:

V.A. Akpan (2012): Model-Based FPGA Embedded-Processor Systems Design Methodologies: Modeling, Syntheses, Implementation and Validation. Afr J. of Comp & ICTs. Vol 5, No.1 pp 1– 26.

- ISSN 2006-1781

However in those days, the number of such systems was not very large basically because of the hardware cost. Although there are infinite varieties of embedded systems, the principles of operation, system components and design methodologies are essentially the same but vary in complexity depending on the application. Embedded systems are used typically to carry out real-time applications. A real-time embedded system can be viewed as an embedded system in which the time at which the output (result) produced is significant and critical. The lag from input time to output time must be sufficiently small for acceptable timeliness.

A typical embedded system design for digital signal processing (DSP) and control systems

applications involves a significant amount of custom logic circuitry with pre-designed hardware components such as processor, memory units and various types of input/output (I/O) interfaces as well as an integrated software platform partitioned between the hardware components within the embedded system initiated and controlled by a real-time operating system [1]. For multivariable systems or systems with relatively short sampling time, the embedded system design becomes a complex and complicated task in meeting real-time constraints.

Traditionally, early systems were controlled by mechanical means using cams, gear, levers and other basic mechanical devices. As system complexity increases, the programmable logic controller (PLC) was introduced which provided an easy way for system control using "*ladder logic*" (and sometimes C) programs and can be reprogrammed rather than rewiring the control system. Unlike computers, PLCs are rugged computer that controls a part of an industrial control system and can typically withstand shock, vibration, elevated temperatures, and electrical noise which are the characteristics of industrial and manufacturing systems [2].

Furthermore, as multivariable systems dynamics coupled with nonlinearities and complexities increases with additional computational burdens, new computing platforms began to evolve as programmable logic devices which includes logic manipulations with varying computational efficiency. These computing platforms have evolved within four decades from simple programmable logic devices (PLDs) such as programmable array logic (PAL), generic array logic (GAL), complex programmable logic devices (CPLDs) to the currently most widely used field programmable gate arrays (FPGAs). The PLDs (PAL, GAL and CPLD) and FPGAs are all programmable devices which mean that they are integrated circuits that are used to create a circuit in which the internal design is not defined until after it has been programmed.

It is important to note here that PLC and PLD are two distinct entities. While the former is the brain of industrial manufacturing process; the latter is an integrated circuit that can be used to implement a digital logic design in hardware [3]. A similarity is that both are programmable. We also note that PLDs contains relatively limited number of logic gates and the functions they can be used to implement are much smaller and simpler when compared to FPGAs.

Moreover, the internal architecture of PLDs are predetermined by the manufacturer, but the PLDs are created in such a way that they can be configured in the field to perform a variety of different functions. On the other hand, FPGAs are digital integrated circuits that contain configurable (programmable) blocks of logic gates along with configurable interconnects between these blocks. Depending on how the FPGAs are implemented, some may only be programmed a single time (one-time programmable, OTP), while others may be reprogrammed several times. The term "*field programmable*" refers to the fact that FPGA programming takes place in the field as opposed to devices whose internal functionality is hardwired by the manufacturer such as application-specific integrated circuits (ASICs) and application-specific standard parts (ASSPs) [4]. Thus, FPGAs can be configured or reprogrammed while it is resident in a higher-level system or resident in an electronic system that has been deployed to the outside world.

ASICs and ASSPs are also programmable devices that can contain hundreds of millions of logic gates and can be used to implement incredibly large and complex functions. ASICs and ASSPs are based on the same design processes and manufacturing technologies. Both are custom-designed and tailored to address specific application. The only difference being that an ASIC is designed and built to order for use by a specific company, while an ASSP is designed for multiple customers. Although ASICs offers the ultimate in terms of logic gate, complexity, and performance but designing and building one is an extremely time-consuming and expensive process, with the added disadvantage that the final design cannot be modified once it has been fabricated (*frozen in silicon*) without creating a new version of the device.

FPGAs lies between PLDs and ASICs because their functionality can be customized in the field like PLDs, and they can contain millions of logic gates that can be used to implement

extremely large and complex functions that could previously be realized using ASICs only [5], [6]. In comparison to ASIC, FPGA is cheaper in small quantities, implementing design changes is much easier, and the time to complete the overall system design, implementation, verification and deployment (i.e., time to market) is faster.

FPGAs are a good choice [1], [7] for implementing digital systems because: *1)* FPGAs offer large logic capacity, exceeding several million equivalent logic gates, and include dedicated memory resources, *2)* they include special hardware circuitry that is often needed in digital systems, such as digital signal processing (DSP) blocks (with multiply and accumulate functionality) and phase-locked loops (PLLs) (or delay-locked loops (DLLs)) that support complex clocking schemes, and *3)* they also Support a wide range of interconnection standards, such as double data rate static read-only-memory (DDR SRAM), PCI and high-speed serial protocols. In addition to the above features, FPGAs provide a significant benefit as “off-the-shelf” chips that are programmed by the end user.

On an FPGA, hard- and soft- processors are available for implementation [1], [6]. A hard processor is a pre-designed circuit that is fabricated within the FPGA chip. A more flexible alternative is to use a soft processor. In this case, the processor exists as code written in a hardware description language (HDL), and it is implemented along with the rest of the system by using the logic and memory resources in the FPGA fabric. One advantage of this approach is that resources on the FPGA are consumed for processor only when these components are actually needed in the system. It is also possible to include multiple soft processors in the FPGA when desired [8].

FPGA are semiconductor devices containing programmable logic components and programmable interconnects. The programmable logic components can be programmed to duplicate the functionality of basic logic gates such as AND, OR, XOR, NOT or more complex combinatorial functions such as decoders or simple math functions. In most FPGAs, these programmable logic components (or logic blocks, in FPGA parlance) also include memory elements, which may be simple flip-flops or more complete blocks of memories. A hierarchy of programmable interconnects allows the logic blocks of an FPGA to be interconnected as needed by the system designer, somewhat like a one-chip programmable breadboard.

These logic blocks and interconnects can be programmed after the manufacturing process by the engineer/designer (hence the term "field programmable") so that the FPGA can perform whatever logical function is needed. Applications of FPGAs include DSP, software-defined radio, aerospace and defense systems, ASIC prototyping, medical imaging, computer vision, speech recognition, cryptography, bio-informatics, computer hardware emulation and a growing range of other areas. FPGAs now find applications in areas that require the use of massive parallelism offered by their architectures [4], [6].

In the traditional approach for designing such systems, a new integrated circuit (IC) chip is created for the custom logic circuits, but each pre-designed component is included as a separate chip. Because many products contains hardware and software components, the difficulty of generating a design from a set of requirements and specifications increases as the product becomes complex. These difficulties led to the development of electronic system level (ESL) design and verification ([9], [10]) which is an algorithm modeling methodology that focuses on a higher abstraction level using high-level languages such as C, C++, or MATLAB® to model the entire behaviour of the system with no initial link to its implementation. The ESL design and verification has evolved into an industry standard complementary methodologies that enable embedded system design, verification, and debugging for designing hardware and software implementation of custom system-on-a-chip, system-on-FPGA, system-on-board, and entire multi-board systems.

As discussed in the next Section 2, embedded systems design are further complicated by the number of design tools required to model, synthesize, implement, verify and the debug the complete embedded system. Section 3 introduces FPGA design methodology using the

Xilinx PlanAhead tool whereas Sections 4 and 5 focus on simulation and verification methodologies using ModelSim™ and ChipScope™ Pro tools respectively. The paper concludes in Section 6 with some discussions. This contribution shows how an embedded system design could be approached from an algorithmic model-based view point. It also shows how these embedded system development tools can be combined to achieve the design objectives. Although, Altera [11] and Xilinx [12] are the two world-leading FPGA manufacturers, the later is considered in this study. The reason being that a careful study of Xilinx's latest FPGA products list by function categories together with their data sheets and performance capabilities published in [1], [12], [13], [14] and [15] reveals that Xilinx's offers the best industry-standard and high-end FPGA design and verification tools [16], [17], [18], [19].

2. The Model-Based Design Flow Methodology

The integration of Simulink and MATLAB from The MathWorks, Inc. [20] and the Xilinx FPGA design suite of tools [12] as well as the Altera FPGA design tools [11] now allows embedded system development from a model-based view point which targets an FPGA. An important model-based implementation which is currently receiving attention is the implementation model predictive control (MPC) algorithms [21] which is a computationally intensive online optimization problem solved at each sample time [22], [23], [24]. Additionally, as nonlinearity is the characteristics of many industrial systems, FPGA implementation of neural network algorithms, which is efficient for modeling the dynamics of nonlinear systems, have also been reported [7], [25].

The proposed block diagram for the embedded system design flow is shown in Fig. 1. A related but reduced architecture from an ASIC point of view has been reported in [26]. As shown in the Fig. 1, four (4) design approaches can be identified for implementing an embedded FPGA-based design by an algorithm developer (AD), a system engineer (SE), a hardware/software engineer (HSE) and a non-DSP hardware engineer (NDHE) view points. The term "model-based designs" here refers to design problems formulated as algorithms and developed using

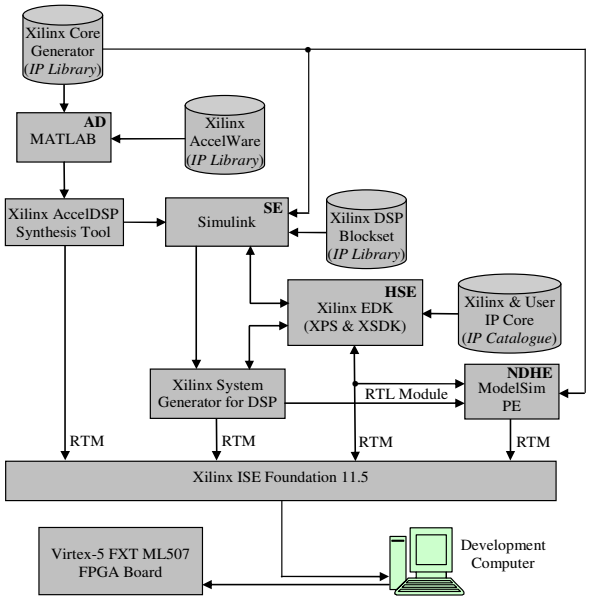


Fig.1. Embedded system design flow: IP – Intellectual Property, AD – algorithm developer, SE – system engineer, HSE – hardware/software engineer, NDSPHE – Non-DSP hardware engineer, EDK – Embedded Development Kit, XPS – Xilinx Platform Studio, XSDK – Xilinx Software Development Kit, RTM – RTL Top-Level Module, ISE – Integrated Software Environment.

Fig. 2. System modeling, development, Simulation and validation.

MATLAB and Simulink from The MathWorks [20].

2.1 Algorithm Development Using the Xilinx AccelDSP Modeling and Synthesis Tool

From an algorithm development and ESL design view point, the problem is specified, formulated and developed as a synthesizable MATALB algorithmic model using MATLAB and the Xilinx

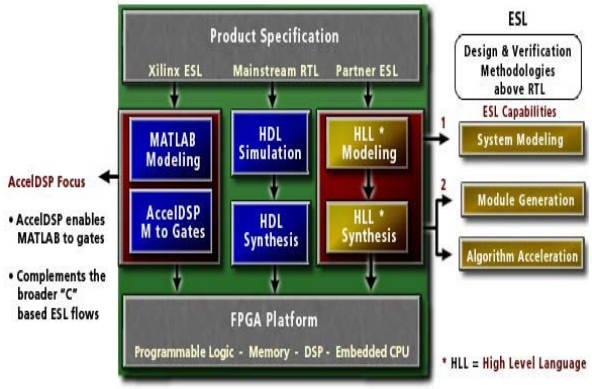


Fig. 3. AccelDSP design routine at the Electronic System Level (ESL).

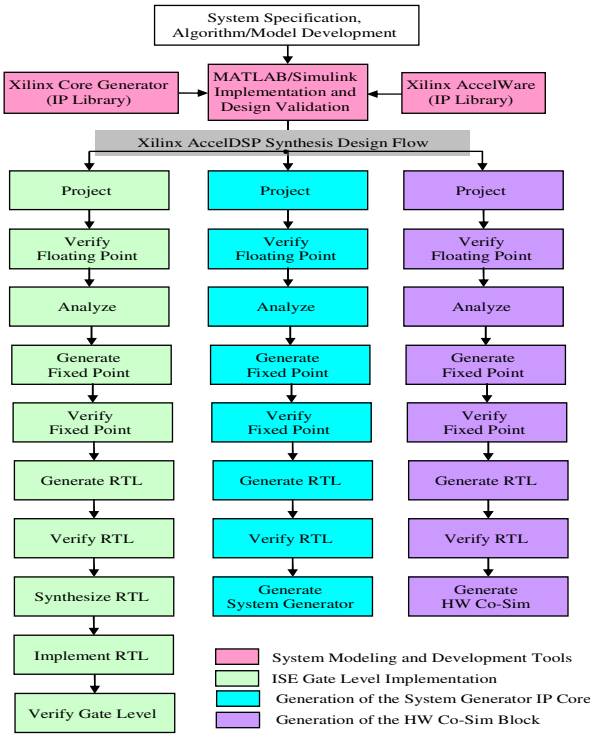


Fig.4: From system specification and algorithm/model development to Xilinx AccelDSP synthesis design flow option implementations

AccelWare functions. Once the MATLAB/Simulink algorithm is validated similar to the workflow shown in Fig. 2 and its performance satisfies the design specifications, the floating-point MATLAB algorithm is then synthesized using the Xilinx AccelDSP modeling and synthesis tool. As illustrated in Fig 3, the Xilinx AccelDSP tool is an advanced ESL design tool which converts a synthesizable floating-point MATLAB algorithm to a fixed-point MATLAB and C++ functions, and automatically generates a verified synthesizable RTL model directly from the fixed-point MATLAB model.

At this point, three AccelDSP implementation options are available as illustrated in Fig. 4, namely: the ISE Gate Level implementation for programming the FPGA via the Xilinx integrated software environment (ISE) foundation, the creation of an IP core for exporting and integrating with a Xilinx System Generator model, or performing a Hardware-in-the-loop Co-Simulation (HW Co-Sim).

The AccelDSP IP Core Generators provide a direct path to hardware implementation for complex MATLAB built-in and toolbox functions, which when used with the AccelDSP synthesis tool, produces synthesizable and pre-verified intellectual property (IP) cores that enables and facilitate algorithmic synthesis for Xilinx FPGAs.

2.2 System Development Using the Xilinx System Generator for DSP Modeling and Simulation Tool

From the system development point of view, the Xilinx System Generator for DSP enables the use of MATLAB/Simulink modeling and simulation environment for FPGA design by providing a smooth path from initial design capture via the System Generator token (shown on the left side of Fig. 5) to Xilinx FPGA implementation and analysis.

Every System Generator for DSP design developed using Simulink must include at least one System Generator token shown on the top left side of Fig. 5. Double clicking on the System Generator token opens up the figure shown on the right in Fig. 5. As shown in this figure, the System Generator token provides six compilation options for exploring FPGAs, namely: *1)* HDL (hardware description language) Netlist generation, *2)* NGC Netlist generation (NGC means netlist with logical design data and constraints), *3)* Bitstream generation, *4)* EDK Export Tool for exporting an EDK processor IP core to a pre-designed embedded processor core or for importing a pre-designed embedded processor core, *5)* Hardware Co-Simulation, and *6)* Timing and Power Analysis.

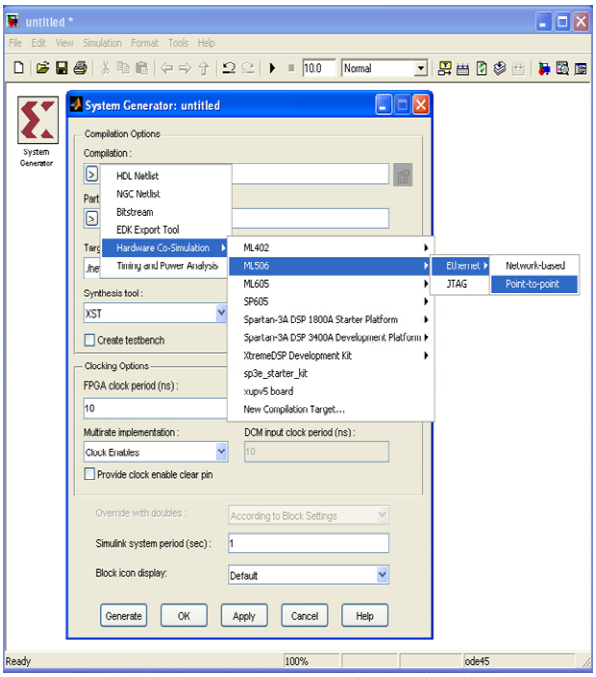


Fig. 5. System Generator token (left) and the six System Generator compilation options (right) with available Hardware Co-Simulation options without the Virtex-5 ML507 FPGA board.

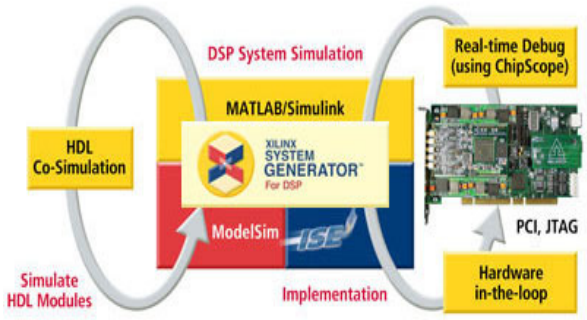


Fig. 6. HDL Co-Simulation with ModelSim and FPGA Hardware-in-the-Loop (HIL) Simulation with ISE using Xilinx System Generator for DSP in MATLAB/Simulink modeling and simulation environment.

As illustrated in Fig. 6, the Xilinx System Generator for DSP also seamlessly integrates with ModelSim for HDL Co-Simulation and the ISE for FPGA hardware-in-the-loop (HIL) simulation. Note that the real-time debugging using the Xilinx ChipScope is an optional tool. The HIL co-simulation is first performed to fix any bug and thereafter the ChipScope debugging tool is included to perform a complete real-time debugging for complete system verification and validation. Logic analysis and performance verification using Xilinx ChipScope™ Pro is presented later in Section 5. Unlike the AccelDSP that interacts with MATLAB as shown in Fig. 7, the Xilinx System Generator for DSP sits directly on top of MATLAB/Simulink as shown in Fig. 6 and Fig. 7.Note that the circled symbol (A) is used to illustrate the System Generator path for FPGA hardware-in-the-loop using the hardware co-simulation block without using the Xilinx ISE as a gateway for direct communication. However, it calls and uses the Xilinx ISE Foundation during all the compilation and synthesis processes, optimization routines, place and route, and generation of the hardware co-simulation block as indicated by the circled symbol (B). Although, the host computer is not counted as part of the five blocks, it is the main development platform upon which all developments, implementation, simulation and validation are performed.

The Xilinx DSP blocksets, provided with System Generator for DSP, contains over 90 DSP building blocks that are available from the Simulink modeling environment. These blocks include filters, correlators, trigonometric functions, modulation/demodulation, error correction, arithmetic, memories, co-processing operations, etc; and they produce optimized and highly synthesized logic for programming Xilinx FPGAs. The Xilinx System Generator for DSP provides three ways for importing a pre-designed embedded processor(s) into a System Generator model, namely: *1)* as a black box block, *2)* as a PicoBlaze Microcontroller block, or *3)* as an EDK Processor block mentioned earlier.

2.3 Xilinx EDK (XPS and XSDK) Development Tools

The Xilinx embedded development kit (EDK) is a suite of tools and collections of ready-to-use IPs (although some IPs can be modified to suite some specific use and/or application) that are used to design a complete embedded processor system for implementation in a Xilinx FPGA. The EDK is made up of the Xilinx platform studio (XPS) that is used for designing the hardware portion of the embedded processor system and the Xilinx software development kit (Xilinx SDK) which is an integrated development environment (IDE) that is used for the

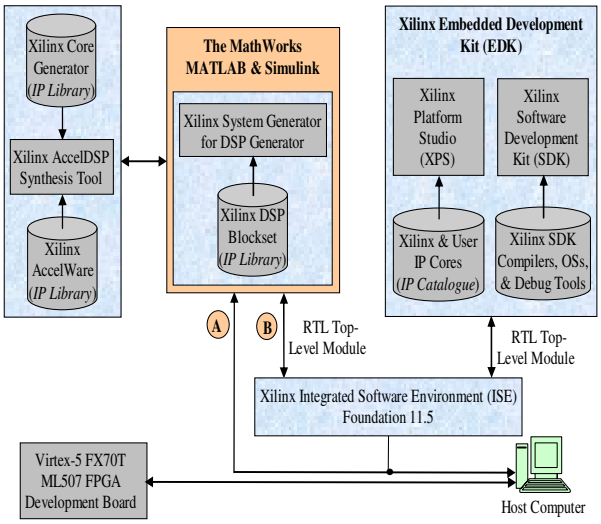


Fig. 7: The block diagram for the proposed model-based design flow for the FPGA implementation of the AGPC algorithm on Virtex-5 FX70T ML507 FPGA development board.

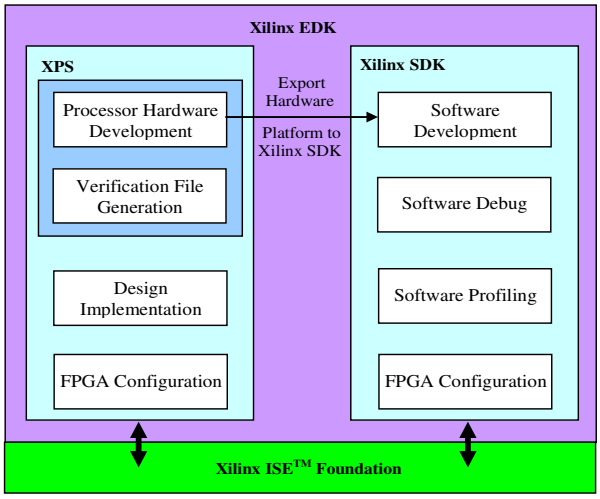


Fig. 8. The basic embedded system design flow using the EDK via ISE™.

development, compilation and verification of C/C++ embedded software portion of the embedded system.

The EDK also includes ready-to-use hardware IPs for Xilinx embedded processors, assorted hardware drivers and libraries for the embedded hardware/software development, GNU compiler and debugger for C/C++ software development targeting the Xilinx MicroBlaze™ soft processor and the IBM PowerPC™ hard processor systems as well as utilities to support all phases of the embedded processor(s) system design and development.

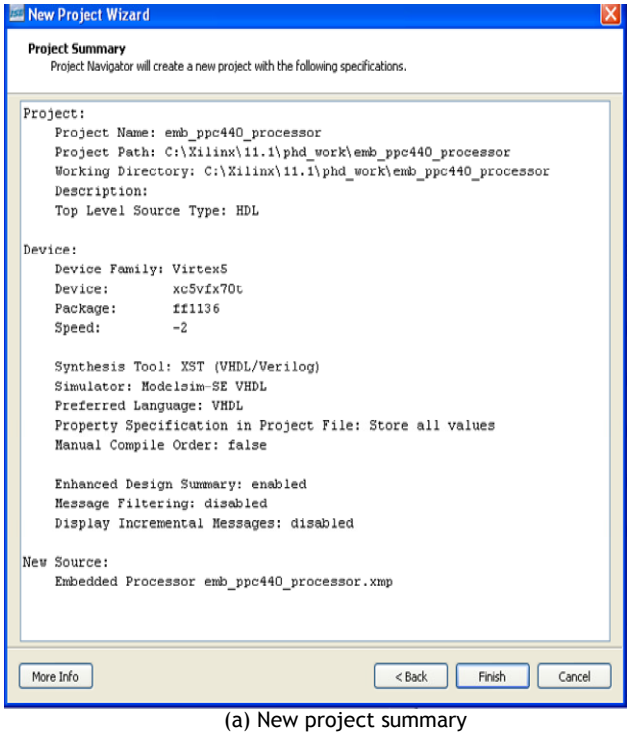
An embedded system design is a complex task since it consists of the hardware and software portions as shown in the basic EDK embedded system design flow in Fig. 8. The task becomes more complicated when integrating an imported algorithmic IP core or user-defined logic into the embedded processor system or integrating an exported embedded processor system into an existing system and/or IP core outside the XPS or the EDK environment. The communication between the processor and the IP core or user-defined logic often occurs over a shared bus. The information conveyed frequently consists of different types of data such as data for processing, data denoting the status of IP core or data affecting the mode of operation. Thus, the Xilinx EDK automates the embedded processor design process using the Base System Builder (BSB) as well as the import/export process to and from the System Generator or exporting to the Xilinx SDK.

2.3.1 The Base System Builder (BSB)

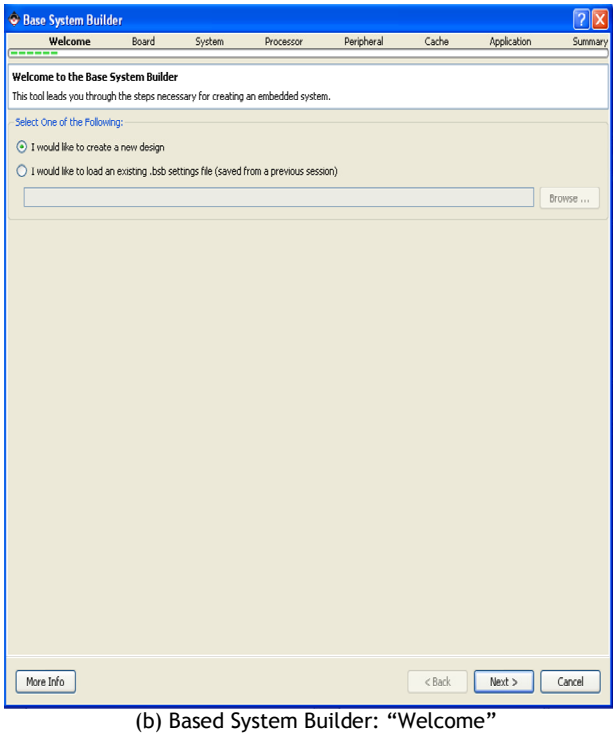
Embedded processor design using the base system builder (BSB) is an eight-stage automated embedded processor design wizard. The eight stages are: Welcome, Board, System, Processor, Peripheral, Cache, Application, and the Summary. Embedded processor design using the base system builder (BSB) is an eight-stage automated embedded processor design wizard. The eight stages are: Welcome, Board, System, Processor, Peripheral, Cache, Application, and the Summary. Even though the BSB wizard is automated, however, the designer must know the requirements of the application for which the embedded processor system is designed for. After the FPGA device, synthesis and simulation tools specification as shown in Fig. 9(a) using the ISE, a typical embedded processor system design flow using the BSB wizard via the EDK tools can be summarized as follows.

The “Welcome” allows new processor(s) to be design or an existing pre-designed processor system to be loaded as shown in Fig. 9(b). The “Board” stage allows the FPGA device family and package to be specified, if different from that specified in the “New Project Wizard”.

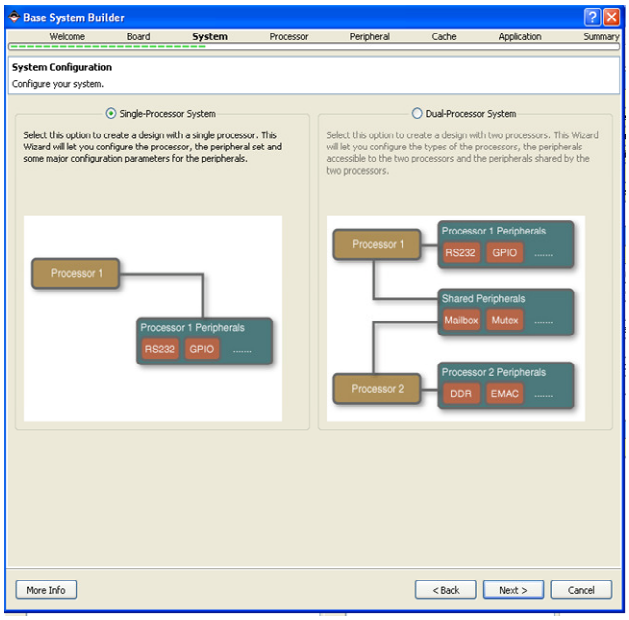
This is sometimes useful if a custom FPGA board



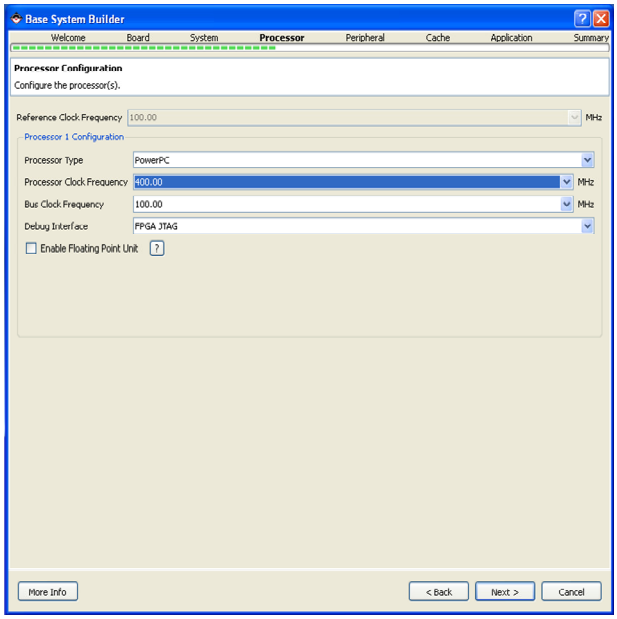
(a) New project summary



(b) Based System Builder: "Welcome"



(c) Based System Builder: "System"



(d) Based System Builder: "Processor"

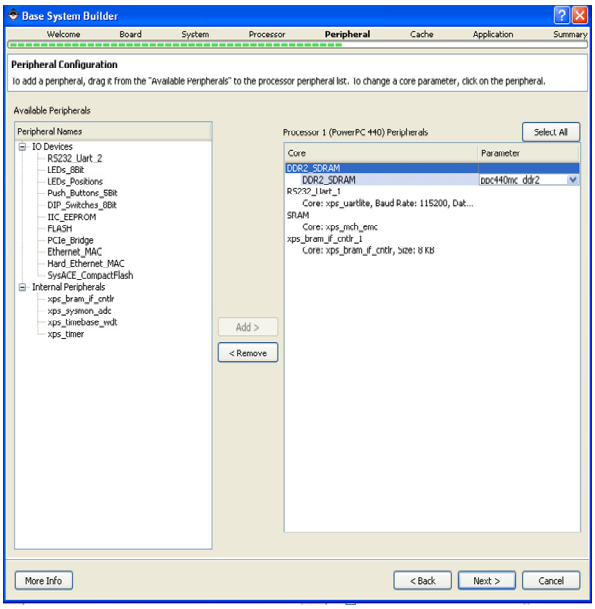
Fig. 9 . The Xilinx ISE "New Project Summary" and the BSB Welcome, System, and Processor design stages for the embedded PowerPC™440 processor system.

different from the pre-configured Xilinx FPGA development boards is to be used. It is also useful if the processor design was not initialized and started using the Xilinx ISE. The advantages of initializing and starting an embedded processor system design from the ISE are many as discussed in Appendix A of [13]. The "*System*" stage shown in Fig. 9(c) allows a single- or dual-processor system to be specified and designed. The Virtex-5 XC5VFX70T devices

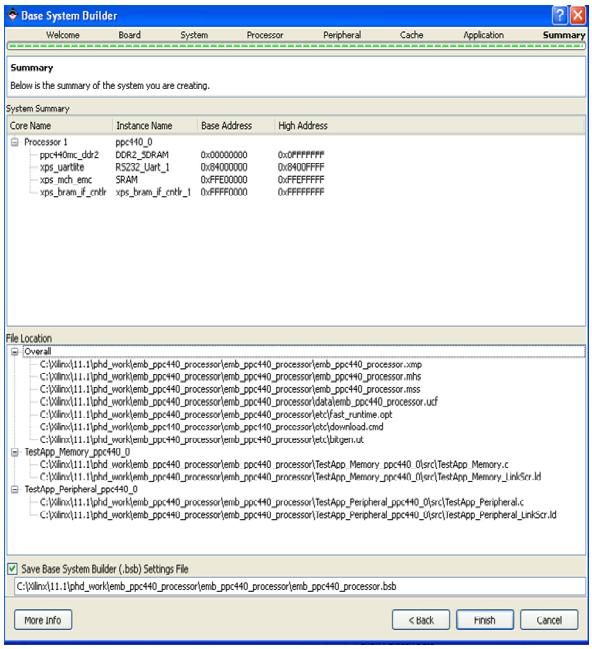
family currently supports single processor systems design. Thus, only a single processor system is discussed here. Next, in the "*Processor*" stage, the choice of selecting a PowerPC™ or a MicroBlaze™ processor are available. For this illustration, a PowerPC™440 is selected as the intended processor as shown in Fig. 9(d). Note that it is also possible to select the MicroBlaze™ processor in the same manner.

The "*Peripheral*" stage allows different memory types and peripherals to be added or removed from the proposed embedded processor system. Once a memory or peripheral is selected, the associated device controllers and drivers are automatically added to the design. Furthermore, if the "Interrupt" check box is selected, the interrupt controller is also included which must be configured in the XPS after the BSB have created the embedded processor system. As discussed in [13] Section 5.3.1.4.2 under memory types as well as hardware and optimization specific to an FPGA embedded processor in Section 5.3.1.4.3 of [13]; the choice of memory and hardware peripheral including their respective controllers have significant effects on the embedded systems performance. Here, peripherals that are not needed are removed. The actual size of the embedded program is yet to be known and this makes the choice of the memory difficult to select. In this regard, the embedded processor local memory is selected first. Next, the external DDR SRAM and the on-board SRAM are added. In this case, the serial port is needed to print out all results to the host development computer. Thus, the only peripheral added here is the *UART (RS323_Uart_1)* and it is configured as follows: *Buat Rate = 115200*, *Data Bits = 8*, *Parity = None*, and the *Interrupt* is not used (that is, it is left unchecked). The BSB dialog for the "Peripheral" stage and the selected memory types and peripherals is shown in Fig. 10(a).

The "*Cache*" stage allows the instruction and data caches memory types and controllers to be enabled. As mentioned earlier, the PowerPC™440 embedded in the Virtex-5 series of FPGAs provides 32-Kbit of caches which are built directly into the silicon of the hard PowerPC™440 core. Normally, these caches are enabled in software and can be configured to cache multiple memory regions. Here, both the instruction and data cache memory types are enabled, although this can also be configure in the software design part of the embedded processor system implementation using the Xilinx SDK.



(a) Based System Builder: "Peripheral"



((b) Based System Builder: "Summary"

Fig. 10. The BSB: the Peripheral and Summary design stages for the embedded PowerPC™440 processor System

The "*Applications*" stage lists the readily available applications to be implemented by the embedded processor system. The applications are usually written in C programming language and users applications. The default Xilinx applications available under "*Application*" are the "Memory" and "Peripheral" test programs as shown in Fig. 10(b) under the *File Location* category: "*TestApp_Memory_ppc440_0*" and "*TestApp_Peripheral_ppc440_0*". Note that new software programs can be crated and added into this "*Application*" both from the XPS after the BSB must have finished creating the processor, and from the Xilinx SDK during the software design portion of the embedded processor system.

The "*Summary*" is the last stage of the BSB-guided steps for creating an embedded processor system. This stage lists all the available peripheral associated with the created embedded processor together with their instance name, base and high addresses as shown under *System Summary* in Fig. 10(b). The "*Summary*" stage also list the major software associated with the processor system as shown under *Overall* in the *File Location* category in Fig. 10(b). The components of the previous "*Application*" stage are also listed in the "*Summary*" stage dialog window.

2.3.2 Importing and Exporting an Embedded Processor Core

As mentioned earlier, an embedded processor can be imported to or exported from a design or a System Generator model. As shown in Fig. 1, Fig. 7 and Fig. 8, the EDK communicates interactively with the ISE™. The ISE™ can be viewed as an interface between the completed embedded processor system and the target FPGA device. It is good practice to start the embedded processor system design using the ISE™ because it manages the complete XPS project (see Fig. 8) including the EDK user constraint file (UCF) as well as the FPGA programming. It is necessary that the interface between the processor and the IP core or user-defined logic be specified via shared-memories. Shared-memories are used to provide storage locations that can be referenced by a name and allows a memory map and the associated software drives to be generated by the EDK tools during the embedded processor design phase.

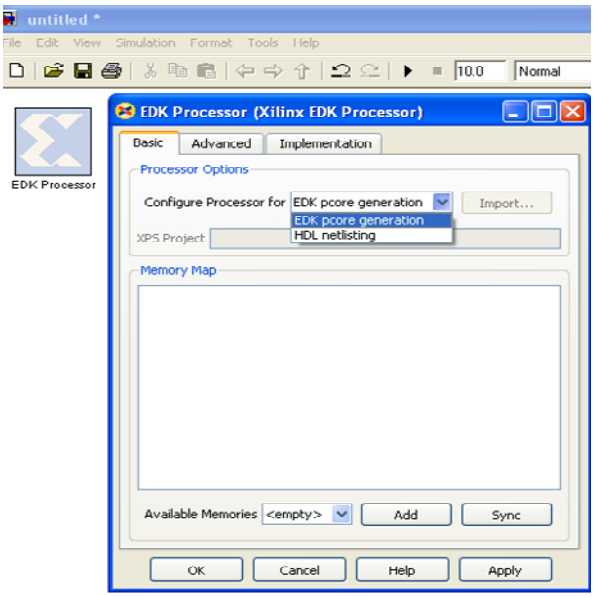


Fig. 11. EDK import and export options within the System Generator

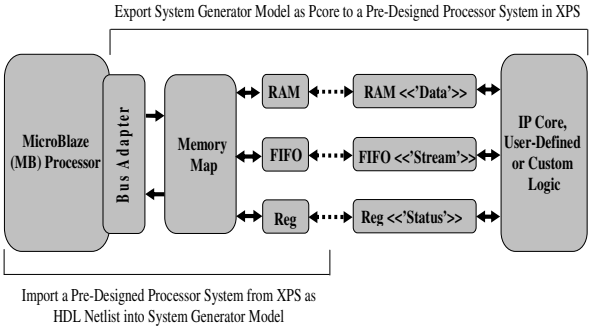


Fig. 12. Basic structure, interface and communication between an embedded processor system and an PI core, user-defined or custom logic.

As discussed in the previous Section 2.2, this contribution is concerned on how the HDL Netlisting and the EDK Export Tool can be used for importing and exporting EDK processor to and/or from the Xilinx System Generator for DSP respectively. The EDK processor development within the System Generator supports two modes of operation shown in Fig. 11, namely: the HDL netlisting mode for importing a pre-design EDK processor from the XPS to System Generator model as a netlist and EDK Pcore generation mode for exporting a System Generator model as a Pcore to the XPS. These two modes of operations are illustrated in Fig. 12. Note that as at the time of this report, the Xilinx EDK and System Generator only support importing a MicroBlaze™ embedded processor

system. The exported Pcore can be integrated into an existing pre-designed embedded MicroBlaze™ processor, embedded PowerPC™ processor or multi-processor systems.

2.3.2.1 Importing an EDK Processor into System Generator Model

When the HDL netlisting mode in Fig. 11 is selected and used with the EDK Export Tool selected via the System Generator token of Fig. 5, the processor system will be imported into System Generator as a black box. The assumption here is that the MicroBlaze™ processor has been designed prior to its importation into the System Generator model. The creation of a complete processor system is an automated process using the Xilinx BSB just discussed above.

The imported EDK processor is also augmented with bus interfaces necessary to connect the System Generator memory map to the processor system. During netlisting within the MATLAB/Simulink environment using System Generator for DSP, the MicroBlaze™ processor and the generated memory-mapped hardware are both nestlisted into hardware. The structure of the EDK processor system imported into the System Generator model is indicated by the portion enclosed lower bracket shown in Fig. 12.

It is necessary to note that once the EDK processor is imported into System Generator, modifications are made to the original embedded processor system in the EDK project to work inside System Generator (no matter the location of EDK project directory). Thus, it may be necessary to retain a copy of the original EDK project before the import process. On the other hand, an EDK processor exported to the Xilinx SDK similar to that illustrated in Fig. 8 does not alter the EDK project but rather the software application runs on the embedded processor system. However, in both the System Generator and Xilinx SDK cases, the processor system must be re-imported if any change(s) is/are made or are to be made after the initial import or export process.

2.3.2.2 Exporting a System Generator Model to the EDK

When the EDK pcore generation mode is selected in Fig. 11 and used with the EDK Export Tool selected via the System Generator token of Fig. 5, the System Generator for DSP is able to create a pcore from the given System Generator model. The structure of the model generated as a pcore to the EDK is shown by the portion enclosed by the upper bracket of Fig. 12.

In this mode, the assumption is that the embedded MicroBlaze™ processor core that has been added to the model is just a place-holder. Its actual implementation is elaborated and filled in by the EDK when the peripheral is finally linked into an EDK project. As such the pcore that is created consists of custom logic, the generated memory map and virtual connections to the custom logic, and the bus adaptor. The pcore also contains a collection of files describing the peripheral's hardware, software drivers, bus connectivity and documentations. Thus, the EDK allows peripherals to be attached to processor(s) created within the EDK based on the above peripheral information.

2.4 Xilinx ISE™ Foundation

The Xilinx ISE™ Foundation is used for design implementation. Design implementation is the process of translating, synthesizing, mapping, place and route, optimization, timing and power analysis, and generating a bitstream file for a given embedded design. With the Xilinx *Fmax Technology*, the Xilinx ISE™ Foundation provides the solution for optimal design performance in the least amount of time. The ISE™ learning curve is greatly reduced because the design implementation tools are conveniently available in ISE™ for easy access and project management, and this can significantly reduce the project completion time [12]. As shown in Fig. 1, the ISE™ is the bridge between the complete design and the FPGA device. It provides a complete RTL design environment for Xilinx FPGAs that includes: design capture using schematics, RTL language templates and RTL editors, simulations, synthesis, place and route, bitstream generation, and programming of the FPGA as well as supports for real-time in-circuit debugging of the programmed FPGA using Xilinx ChipScope™ Pro. The typical ISE™ design implementation is summarized in the flowchart of Fig. 13.

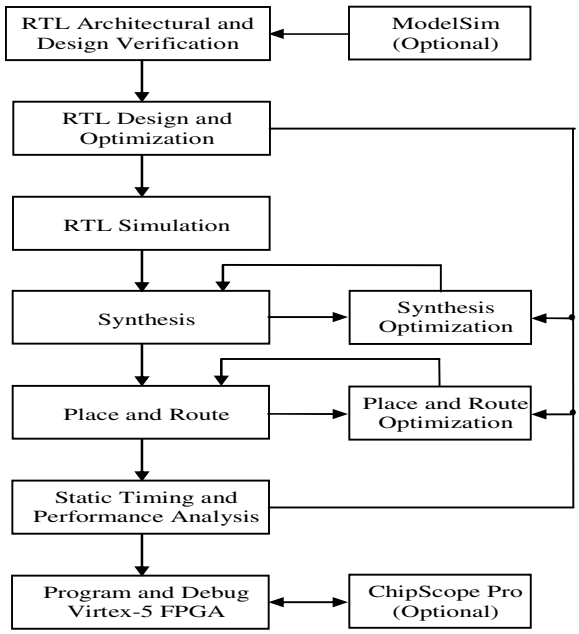


Fig. 13. Typical ISE™ design implementation flowchart.

The Xilinx's *Xplorer Script (XST)* tool allows for the observation of maximum design performance for any clock cycle in a design by running the various combinations of implementation tool options. The Xilinx *Timing-Driven Place and Route* tool is enhanced for optimal programmable logic prototyping and implementation. In *ISE™* when specifying timing requirements for critical paths, performance is dramatically improved through tools such as *Timing Analyzer, Constraints Editor* with *Time Specs*, *FPGA Editor*, and the *Floorplanner™*. The *PlanAhead™ Design Analysis Tools*, discussed in the next section, is also used in the design implementation and configuration as an optional hierarchical *Floorplanner™* and design analysis tool that decreases design time and increases performance by simplifying logic synthesis through the physical design analysis.

Translate and Map facility is also provided by the ISE™ which performs all the steps necessary to read a netlist file in EDIF format and creates an output file describing the logical design (here, a logical design is in terms of logic elements such as AND gates, OR gates, decoders, flip-flops, and RAMs). The ISE™ *Timing-Driven Map* technology helps to lower device cost. With an exclusive *Timing-Driven Map* option, better design utilization for FPGA devices can be achieved, particularly if the device is already more than 90% utilized. *Timing-Driven Map* is a next-generation enhancement to *ISE™* physical synthesis, and combines placement with *logic slice packing* to improve placement quality for "*unrelated logic*". *Device Configuration* is also supported when configuring the programmable logic device. As the last step in design methodology, a bitstream is generated from the physical place and route information and is transferred through cables to the target device as shown in Fig. 1 and Fig. 10.

The *ISE™* Project Navigator guides the designer through a simple *Push-Button Design Flow* to implement design automatically. For more complex designs, designer has complete control over every aspect of the design process flow. *High Speed Design* is fully supported because almost every high-performance logic system being developed today contains a high-speed memory interface to logic. With DDR and QDR clock timing and accuracy demands, the most robust design is achieved with no room for error in the interfaces. The *ISE™* also include *High Speed I/O* synthesis optimization of paths to/from the 10.3125 Gb/s I/Os. The ISE™ libraries include 1-, 2- and 4-byte versions of Xilinx high-speed protocols which can readily be included into HDL codes for high performance.

3. Advanced FPGA Design and Analysis Using Xilinx PlanAhead Design Tool

PlanAhead [27], [28], [29] is an advanced FPGA design and analysis tool used to design FPGA devices. It provides an integrated and intuitive environment for the entire FPGA implementation process. With PlanAhead, circuit performance improvements can be realized by analyzing the design RTL sources, synthesized netlists and implementation results. In addition, experiments with different implementation options, refine timing constraints and apply physical constraints and floorplanning techniques can be performed to help improve design results. Early estimates of resource utilization, interconnect delay and routing connectivity can assist with appropriate logic design, device selection and floorplanning.

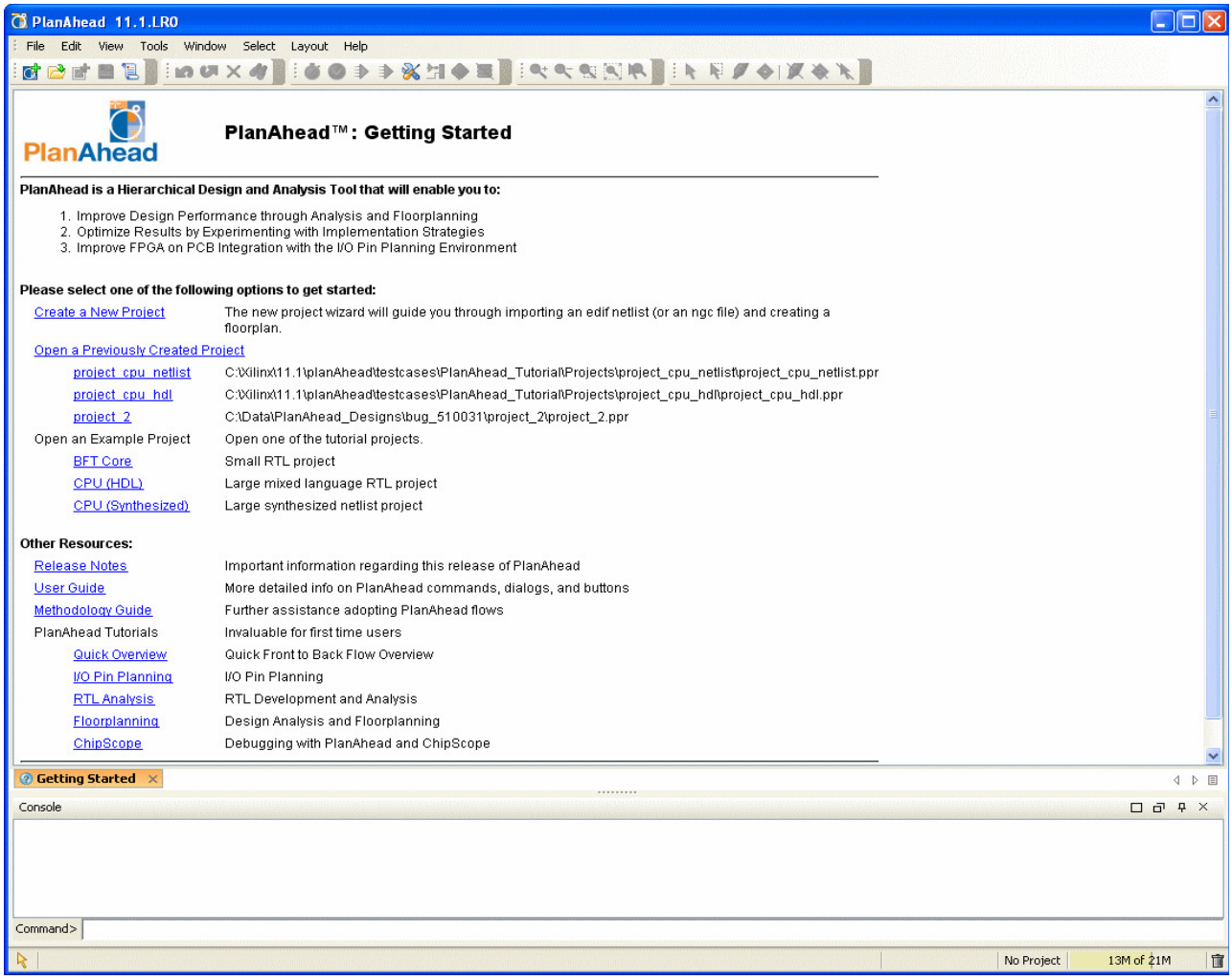


Fig. 14. The PlanAhead create new project, getting started, help and other resources page. Notice the PlanAhead Getting Started help page with links to open or create projects and to view documentation.

A hierarchical database also enables a block-based incremental design methodology that can reduce the run times and compute resources required to place and route the design.

Invoking the PlanAhead software brings up the window shown in Fig. 14 as a *Getting Starting* page with help and other resources as well as links to open or create project and to view documentation. A click on the "*Create a New Project*" option opens the "*Create a New PlanAhead Project*" wizard shown in Fig. 15 that can guide one through the PlanAhead project design and analysis setup process. PlanAhead features includes input/output (I/O) planning environment, RTL development and analysis environment, synthesis and implementation environment, design and placement analysis, static timing estimation (*TimeAhead*), timing constraint editor, robust floorplanning environment, block-based capabilities, etc.

The PlanAhead software tool provides an environment which can help to improve design results throughout the design flow. Integrated within the Project Navigator environment, PlanAhead is automatically launched at four different design process steps, namely: *1)* I/O pin planning (Pre-Synthesis), *2)* I/O pin planning (Post-Synthesis), *3)* Floorplan Area/IO/Logic (Post-Synthesis), and *4)* Analyze Timing/Floorplan Design (Post-Implementation). Each of these steps offers unique and powerful capabilities previously only available in the standalone PlanAhead environment. The PlanAhead software replaces PACE and Floorplanner for all pin planning, design

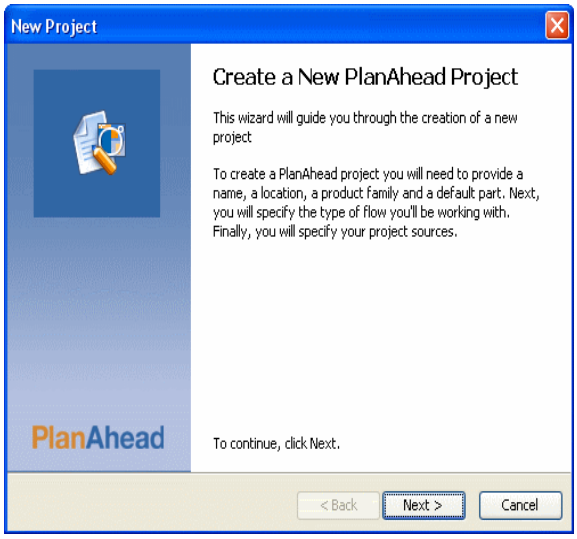


Fig. 15. PlanAhead New Project overview.

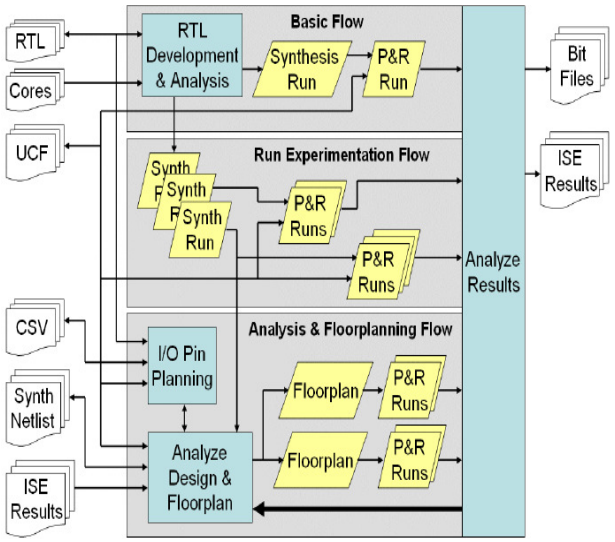


Fig. 16. PlanAhead design flows with inputs and outputs files.

viewing, and floorplanning flows for FPGA designs.

When PlanAhead is invoked from Project Navigator, the interface provides access to only the PlanAhead features specific to the selected task. This mode of PlanAhead is called ISE™Integration mode. The mode is displayed in the status bar at the bottom of the PlanAhead viewing environment.

3.1 PlanAhead Design Flows

The PlanAhead™ software can be used in various ways at different points in the FPGA design flow. It can also be used as a complete flow management tool from RTL development through bitstream generation, or for I/O pin planning, RTL netlist analysis, implementation result design analysis, floorplanning, or ChipScope™ tool core insertion and debugging (see Section 5 on ChipScope™ Pro). The complete PlanAhead design flows are illustrated in Fig. 16.

Through analysis and floorplanning, physical constraints are applied to help control the implementation of the design. The PlanAhead environment allows and enables exploration and experimentation with various implementation strategies. All of the implementation attempts and data are completely managed from within the PlanAhead environment. PlanAhead is also used after implementation to analyze the ISE™ software placement and timing results in order to improve the performance of the design. Additional physical constraints derived from the imported results may also be used to lock placement during subsequent implementation attempts. The flow chart of Fig. 16 illustrates the common design flows as well as the various input and output formats of PlanAhead. The addition of a logic synthesis environment and front-to-back process is one of the biggest additions to PlanAhead in Xilinx design suite release 11 (2009) onwards [12]. It is now possible to import RTL sources, synthesize the logic, implement the synthesized netlist, analyze the implementation results, floorplan and experiment with implementation options and generate bitstreams.

PlanAhead manages the design flow and data to accommodate experimentation with multiple synthesis and implementation runs simultaneously as illustrated in Fig. 16. Previously, PlanAhead projects consisted of a single imported netlist; the creation of a floorplan was always required; and the netlist was always loaded into memory during a PlanAhead session. However, as shown in Fig. 16, it is also now possible to create an integrated multiple floorplans for experimentation, which were basically constrained for a single netlist.

3.1.1 Basic Flow

The basic flow enable users to easily import RTL sources, synthesize and implement the design and view results. It is often used to implement the design initially. If further design analysis and floorplanning is desired, the design analysis flow should be used. We briefly list the input/output file formats and processes used while importing design data. These files include: RTL Source Files (Verilog, VHDL, or other design text files), Xilinx Cores (NGC / NGO), XST Constraint Files (XCF), I/O Port Lists (CSV), I/O Port Lists (HDL - Verilog or VHDL), Top-Level Netlists (EDIF), Module-Level Netlists (EDIF), Top-Level Netlists (NGC), Constraint Files (UCF / NCF), Xilinx ISE Placement Results (NCD / XDL), and the Xilinx TRCE Timing Results (TWX/TWR). While reading the input files, PlanAhead writes out any errors, warnings and messages in to the *planAhead.log* file. These messages are also displayed in the PlanAhead Console view.

3.1.2 Run Experimentation Flow

Experiment with multiple synthesis or implementation runs using different strategies can be performed. A set of proven PlanAhead strategies is bundled with the software but user-defined strategies can also be defined. The PlanAhead synthesis and implementation environment enables the designer to configure, launch, and monitor multiple synthesis runs using the Xilinx® XST synthesis tool. It is also possible to define reusable *strategies* for synthesis runs. As an example, one might create strategies for power, performance, or area optimization; and then assign these strategies to individual runs, and launch them simultaneously or serially. The synthesis run results are displayed interactively and report files are accessible.

Running multiple synthesis strategies results in multiple netlists being created, and stored within the PlanAhead project. PlanAhead then enables one to interactively load the various versions of the netlist into the environment for analysis. During the netlist import, Floorplans can be created for I/O pin planning, device analysis, floorplanning and implementation. Implementation runs can be created from any completed PlanAhead synthesis run, floorplan, or imported third-party synthesized netlist. Creating floorplans allows experimentation with various logical constraints, physical constraints, or alternate devices and then define reusable strategies for implementation runs. As an example, one might create strategies for the various map logic optimization options or par effort levels and then assign these strategies to individual runs and launch them simultaneously or serially. The implementation run results are displayed interactively and the command report files are accessible.

3.1.3 Design Analysis & Floorplanning Flow

PlanAhead has extensive design analysis and floorplanning capabilities that can be utilized at various stages of the design process. RTL can be elaborated and analyzed prior to synthesis. Synthesized netlists can be explored using target devices and constraints. Implementation runs can be imported into the PlanAhead environment for further analysis and floorplanning. Implementation results derived from using command line tools can also be imported. Specific flow features for design analysis and floorplanning are detailed in [27], [28], [29].

3.2 Using Platform Studio and EDK with PlanAhead

The Xilinx Platform Studio (XPS) tool is used by FPGA designers to build embedded systems on Xilinx FPGA devices. XPS provides a comprehensive environment for designers to integrate their hardware and software system components. To ease the integration process, XPS tries to make use of preset design tool options while implementing the embedded system. Using preset options makes the process of system implementation simpler but can result in the system not meeting desired performance goals. In such cases, designers need to have an option to enhance their system performance. The PlanAhead tool has been proven to help users in hardware design analysis and physical design for performance improvement. Design flow using XPS and PlanAhead is shown in the flow chart of Fig. 17.

In XPS [27], use either the *Tools → Generate Netlist* or *Tools → Generate Bitstream* options to create "*synthesis*" and "*implementation*" sub-directories. Synthesis scripts (*.scr*), project

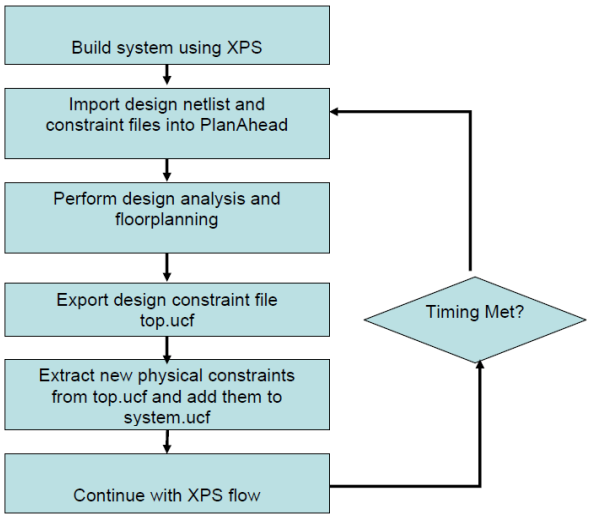


Fig. 17. PlanAhead design methodology using XPS and EDK.

files (*.prj*) and report (*.srp*) files are stored in the "*synthesis*" sub-directory. Design netlist files created as a result of the synthesis process are stored in the "*implementation*" sub-directory. XPS also creates *.bmm* files in the "*implementation*" sub-directory for configuring the Block RAMs on the device. The PlanAhead tool is inserted in the design flow after synthesis has been performed.

XPS uses the XST (Xilinx Synthesis Technology) tool for design synthesis. The synthesis netlists generated are in NGC format. The top-level file is named "*system.ngc*" and all other files have the nomenclature "*module_name.ngc*". The design constraints are stored in a file named "*system.ucf*". The following steps are used to illustrate this design flow:

1). Build the system using XPS.
2). Use the *Tools → Generate Netlist* or *Tools → Generate Bitstream* options in XPS to create the *system.ngc* and *system.ucf* files, along with the NGC and NCF files for sub modules.
3). Import the *system.ngc*, *system.ucf* and core files into the PlanAhead tool.
4). Use the analysis and physical design capabilities of PlanAhead to explore the design.
5). Create physical design constraints if necessary, as determined in Step. 4.
6). Use the *File → Export floorplan* command to export the top-level UCF file, top.ucf. Under the *File types to generate* header, clear the check mark next to *Netlist*.
7). Manually extract the new physical constraints from the *top.ucf* file and add them to the XPS generated *system.ucf* file.
8). Run place and route through XPS.
9). If the design performance has improved as desired, then continue with the post placement and routing XPS flow.
10). If design goals are still not met, then return to the PlanAhead environment for further analysis. Determine the best choice of RTL changes, floorplanning, etc. to proceed.

4. Verifications and Simulations Using ModelSim

ModelSim is a verification and simulation tool for VHDL, Verilog, SystemVerilog, and mixedlanguage designs [30], [31]. This section provides a brief and conceptual verification and simulation techniques that can be achieved using ModelSim. It is divided into four main work-flow: (a) Basic simulation flow, (b) Project flow, (c) Multiple library flow, and (d) Debugging tools.

4.1 Basic Simulation Flow

The flow chart shown in Fig. 18(a) shows the basic steps for simulating a design in ModelSim.

1). Creating the Working Library
In ModelSim, all designs are compiled into a library. We typically start a new simulation in ModelSim by creating a working library called "*work*". "*Work*" is the library name used by the compiler as the default destination for compiled design units.

2). Compiling the Design
After creating the working library, we compile the design units into it. The ModelSim library format is compatible across all supported platforms. The design can be simulated on any platform without having to recompile the design.

3). Loading the Simulator with the Design and Running the Simulation
With the design compiled, we load the simulator with the design by invoking the simulator on a top-level module (Verilog) or a

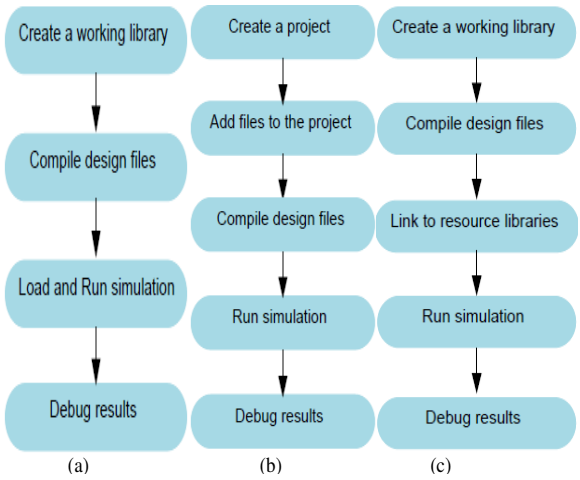


Fig. 18. ModelSim verification and simulation flows: (a) Basic simulation flow, (b) project flow, and (c) multiple library flow.

configuration or entity/architecture pair (VHDL). Assuming that the design loads successfully, the simulation time is set to zero, and one can enter a run command to begin simulation.

4). Debugging the Results

If the resulted obtained is different from that expected, then ModelSim's robust debugging environment can be used to track down the cause of the problem and result mismatch.

4.2 Project Flow

A ModelSim project is a collection mechanism for an HDL design under specification or test. Even though one does not have to use projects in ModelSim, they may ease interaction with the tool and are useful for organizing files and specifying simulation settings.

The diagram of Fig. 18(b) shows the basic steps for simulating a design within a ModelSim project flow. As one can observe in Fig. 18(a) and (b), the project design flow is similar to the basic simulation flow. However, there are two important differences:

(i) One does not have to create a working library in the project flow; it is done automatically by ModelSim.
(ii) Projects are persistent. In other words, they will open every time ModelSim is invoked unless specifically closed.

4.3 Multiple Library Flow

ModelSim uses libraries in two ways: *1)* as a local working library that contains the compiled version of the design; and *2)* as a resource library. The contents of the working library will change as the design is updated and recompiled. A resource library is typically static and serves as a parts source for the design. Although, one can create personal resource libraries, they may also be supplied by another design team or a third party (e.g., a silicon vendor).

One specifies which resource libraries will be used when the design is compiled, and there are rules to specify in which order they are searched. A common example of using both a working library and a resource library is one where the gate-level design and testbench are compiled into the working library, and the design references gate-level models in a separate resource library. The diagram of Fig. 18(c) shows the basic steps for simulating with multiple libraries.

One can also link to resource libraries from within a project. For example, if a project is being used, one would replace the first step above with these two steps: create the project and add the "*testbench*" to the project.

4.4 Debugging Tools - Automating Simulation using DO Files

ModelSim offers numerous tools for debugging and analyzing designs including using projects, working with multiple libraries, setting breakpoints and stepping through the source code, viewing waveforms and measuring time, viewing and initializing memories, creating stimulus with the Waveform Editor, and automating simulation.

Apart from executing a couple of pre-existing DO files, the previous discussions focused on using ModelSim in interactive mode: executing single commands, one after another, via the GUI menus or Main window command line. In situations where the tasks to be completed are repetitive, one can increase the productivity with DO files.

DO files are scripts that allows the execution of many commands at once. The scripts can be as

simple as a series of ModelSim commands with associated arguments, or they can be full-blown Tcl (tool command language) programs with variables, conditional execution, and so forth [32]. The DO files can be executed from within the GUI or can be run from the system command prompt without ever invoking the GUI.

4.4.1 Creating A Simple DO File

Creating DO files is as simple as typing the commands in a text file. Alternatively, the Main window transcript can be saved as a DO file. In the following illustration given below, the commands entered in the Main window transcript are used to create a DO file that adds signals to the Wave window, provides stimulus to those signals, and then advances the simulation. An example is given below for illustration.

1). Load the *design_name* design unit (where *design_name* specifies the design name)
   a. Start ModelSim.
   b. Change the default directory to the current working directory.
   c. Load the design unit with the design (*design_name*).
   d. Enter *vsim design_name*.
2). Enter commands to add signals to the Wave window, force signals, and run the simulation.
   a. Select File → New → Source → Do to create a new DO file.
   b. Enter the following commands into the source window:

```
add wave count
add wave clk
add wave reset
force -freeze clk 0 0, 1 {50 ns} -r 100
force reset 1
run 100
force reset 0
run 300
force reset 1
run 400
force reset 0
run 200
```

3). Save the file.
   a. Select *File → Save As*.
   b. Type *sim.do* in the File name: field and save it to the current directory.
4). Load the simulation again and use the DO file.
   a. Enter *quit -sim* at the VSIM> prompt.
   b. Enter *vsim test_counter* at the ModelSim> prompt.
   c. Enter *do sim.do* at the VSIM> prompt.
   ModelSim executes the saved commands and draws the waves in the Wave window.
5). After completing the above, one can select *File → Quit* to quit ModelSim.

4.4.2 Running in Command-Line Mode

We use the term "*command-line mode*" to refer to simulations that are runned from a DOS/UNIX prompt without invoking the GUI. Several ModelSim commands (e.g., *vsim, vlib, vlog,* etc.) are actually stand-alone executables that can be invoked at the system command prompt. Additionally, one can create a DO file that contains other ModelSim commands and specify that file when the simulator is invoked as follows.
1). Create a new directory and copy the project files into it.
   Start by creating a new directory for the project. Create the directory and copy any existing source files into the newly created directory.
2). Create a new design library and compile the source file.
   Again, the following commands are entered at the DOS prompt in the new directory that was created in step 1.
   a. Type *vlib work* at the DOS prompt.
   b. For Verilog, type *vlog design_name.v* and for VHDL, type *vcom design_name.vhd* at the DOS prompt.
3). Create a DO file.
   a. Open a text editor.
   b. Type the following lines into a new file:

```
# list all signals in decimal format
    add list -decimal *
# read in stimulus
    do stim.do
# output results
    write list design_name.lst
# quit the simulation
    quit -f
```

   c. Save the file with the name *sim.do* and place it in the current directory.
4). Run the batch-mode simulation.
   Type *vsim -c -do sim.do design_name -wlf design_name.wlf* at the DOS prompt.
   The *-c* argument instructs ModelSim not to

invoke the GUI. The *-wlf* argument saves the simulation results in a WLF file. This allows one to view the simulation results in the GUI for debugging purposes.

5). View the list output.

Open *design_name.lst* and view the simulation results. Note that the output produced by the Verilog version of the design may appear slightly different from that based on VHDL version.

6). View the results in the GUI.

Since the simulation results are saved in *design_name.wlf*, they can then be viewed in the GUI by invoking VSIM with the *-view argument*.

a. Type *vsim -view design_name.wlf* at the DOS prompt.

The GUI opens and a dataset tab named "*design_name*" is displayed in the Workspace

b. Right-click the *design_name* instance and select Add → To Wave → All items in region.

The waveforms display in the Wave window.

7). After the results must have been viewed, one can select File → Quit to close ModelSim.

4.4.3 Using Tcl with the Simulator

The DO files discussed and used in previous illustration contained only ModelSim commands. However, DO files are really just Tcl scripts [32]. This means one can include a whole variety of Tcl constructs such as procedures, conditional operators, math and trig functions, regular expressions, and so forth.

In this discussion, a simple Tcl script that tests for certain values on a signal can be created which then adds bookmarks that zooms the Wave window when that value exists. Bookmarks allow one to save a particular zoom range and scroll position in the Wave window as in the example below.

1). Create the script.

a. In a text editor, open a new file and enter the following lines:

```
proc add_wave_zoom {stime num}
 {
   echo "Bookmarking wave $num"
   bookmark add wave "bk$num" "[expr
                $stime - 50] [expr $stime + 100]" 0
 }
```

These commands do the following:

- Create a new procedure called "*add_wave_zoom*" that has two arguments, *stime* and *num*.
- Create a bookmark with a zoom range from the current simulation time minus 50 time units to the current simulation time plus 100 time units.

b. Now add these lines to the bottom of the script:

```
add wave -r /*
  when {clk'event and clk="1"}
  {
      echo "Count is [exa count]"
    if {[examine count]== "00100111"}
        {add_wave_zoom $now 1}
    elseif {[examine count]=="01000111"}
            {add_wave_zoom $now 2}
  }
```

These commands do the following:

- Add all signals to the Wave window.
- Use a when statement to identify when *clk* transitions to 1.
- Examine the value of *count* at those transitions and add a bookmark if it is a certain value.

c. Save the script with the name "*design_name.do*".

Save it into the current working directory created previously.

2). Load the *design_name* design unit.

a. Start ModelSim.

b. Select *File → Change Directory* and change to the directory the DO file was saved to in step 1(c) above.

c. Enter the following command at the QuestaSim> prompt:

vsim – voptargs = "+acc" *design_name*

3). Execute the DO file and run the design.

a. Type do *design_name.do* at the VSIM> prompt.

b. Type run 1500 ns at the VSIM> prompt.

The simulation runs and the DO file creates two bookmarks.

c. (i) If the Wave window is docked in the Main window, click somewhere in the Wave window and select Wave → Bookmarks → bk1.

(ii) However, if the window is undocked, select View → Bookmarks → bk1 in the

Wave window.

(iii) The Wave window can be zoomed (in or out) and scrolled to the time when *count* is 00100111.

5. Logic Analysis and Design Performance Verifications Using Xilinx ChipScope™ Pro

5.1 ChipScope Pro Tools Overview

As the density of FPGA devices increases, so does the impracticality of attaching test equipment probes to these devices under test. The ChipScope™ Pro tools integrates key logic analyzer and other test and measurement hardware components with the target design inside the supported Xilinx FPGA devices listed in the ISE™ Design Suite Product Table [33]. The tools communicate with these components and provide the designer with a robust logic analyzer solution.

The ChipScope™ Pro 11.4 Serial I/O Toolkit provides features and capabilities specific to the exploration and debugging of designs that use the high-speed serial transceiver I/O capability of Xilinx FPGAs. The internal bit error ratio tester (IBERT) core and related software provides access to the high-speed serial transceivers (referred to as MGTs in the Xilinx's ChipScope™ Pro documentations) [12] and perform bit error ratio analysis on channels composed of these MGTs. The IBERT core supports the high-speed serial transceivers found in the Xilinx Virtex®-4, Virtex-5, Virtex-6, and Spartan®-6 FPGA devices listed in the ISE Design Suite Product Table [33]. Table 1 gives a brief description of the various ChipScope™ Pro software tools and their IP cores as well as their description.

5.2 ChipScope™ Pro System Development, Implementation and Debugging Platform

Fig. 19 shows a block diagram of a system containing debug cores added using the

TABLE I

CHIPSCOPE™ PRO TOOLS DESCRIPTION

	Tool	Description
1	Xilinx CORE Generator Tool	Provides core generation capability for the ICON, ILA, VIO, and ATC2 cores targeting all supported FPGA device families, and IBERT cores. Also provides core generation capability for the IBERT core targeting the Virtex-5, Virtex-6, and Spartan-6 FPGA families.The Xilinx CORE Generator™ tool is part of the Xilinx ISE Design Suite software tool installation.
2	IBERT Core Generator	Provides full design generation capability for the IBERT core targeting the Virtex-4 and Virtex-5 devices. The user chooses the MGTs and parameters governing the design, and the Core Generator uses the ISE design suite to produce a configuration file.
3	Core Inserter	Automatically inserts the ICON, ILA, and ATC2 cores into the user's synthesized design.
4	PlanAhead Design Analysis Tool	Automatically inserts the ICON and ILA cores into the design netlist. For more information on this feature, go to PlanAhead™ Design Analysis Tool [27], [28], [29], [34].
5	Analyzer	Provides device configuration, trigger setup, and trace display for the ILA, IBA/OPB, IBA/PLB, VIO, and IBERT cores. The various cores provide the trigger, control, and trace capture capability. The ICON core communicates to the dedicated Boundary Scan pins. The Analyzer tool also provides device configuration, project management, and control over the IBERT core, including monitoring status and controlling variables.
6	ChipScope Engine Tcl (CSE/Tcl) Scripting Interface	The scriptable CSE/Tcl command interface makes it possible to interact with devices in a JTAG chain from a Tcl shell[(1)][35].

Note: ICON - Integrated Controller core, ILA - Integrated Logic Analyzer core, VIO - Virtual Input/Output core, IBERT - Integrated Bit Error Ratio Test core, ATC2 - Agilent Trace Core 2, and IBA - Integrated Bus Analyzer cores.

(1) *Tcl* stands for *Tool Command Language*. The CSE/Tcl interface requires the Tcl shell program (called xtclsh) that is included in the ChipScope Pro and ISE 11.4 tool installations or in the ActiveTcl 8.4 shell available from ActiveState [34].

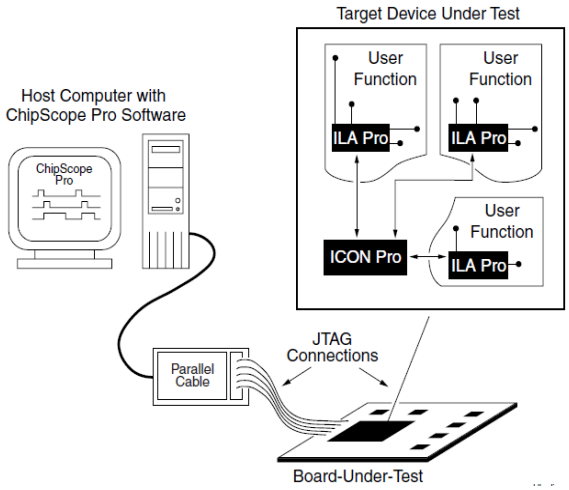


Fig. 19. ChipScope Pro system block diagram with the host development computer, available ChipScope Pro cores, connection interface and the FPGA board under test.

ChipScope™ Pro tools. Users can place the ICON, ILA, VIO, and ATC2 cores (collectively called the ChipScope™ Pro cores) into their design by generating the cores with the Core Generator and instantiating them into the HDL source code. One can also insert the ICON, ILA, and ATC2 cores directly into the synthesized design netlist using the Xilinx Core Inserter or PlanAhead™ tools. The design is then placed and routed using the ISE™ 11.4 implementation tools. Next, the user downloads the bitstream into the device under test and analyzes the design with the Analyzer software.

The Analyzer tool supports the following download cables for communication between the PC and the device in the JTAG Boundary Scan chain: Platform Cable USB and Parallel Cable IV. The Analyzer and cores contain many features that FPGA designers need for thoroughly verifying their logic designs. User-selectable data channels range from 1 to 4,096 and the sample buffer sizes range from 256 to 131,072 samples. Users can change the triggers in real time without affecting their logic. The Analyzer leads designers through the process of modifying triggers and analyzing the captured data.

5.3 ChipScope™ Pro Tools Design Flow

The ChipScope™ tools design flow illustrated in

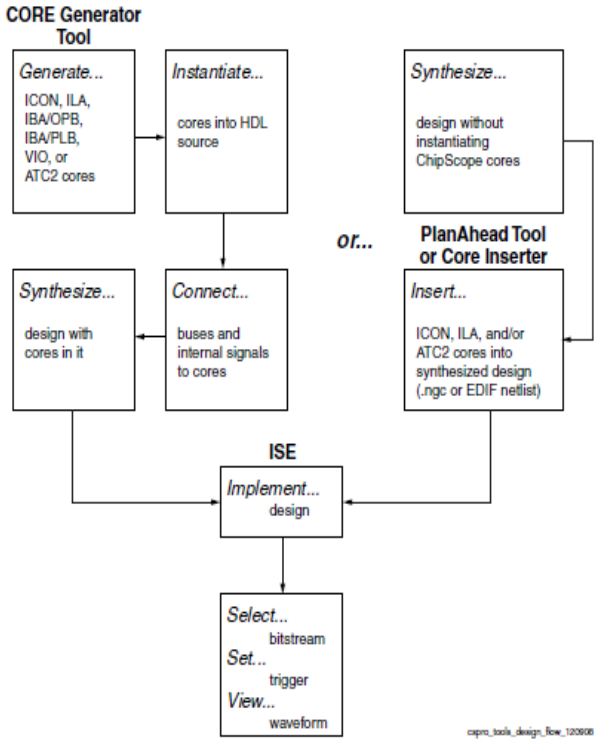


Fig. 20. ChipScope™ Pro design flow.

Fig. 20 merges easily with any standard FPGA design flow that uses a standard HDL synthesis tool and the ISE 11.4 implementation tools. The cores (ICON, ILA, IBA, VIO, and ATC2) can also be used in the EDK and System Generator for DSP tool flows for embedded processor and DSP designs, respectively. For more information on how to use the ChipScope™ Pro cores, see the EDK Platform Studio [36] and System Generator for DSP [37] documentation.

5.3.1 ICON Core

All of the cores use the JTAG Boundary Scan port to communicate to the host computer via a JTAG download cable. The ICON core provides a communications path between the JTAG Boundary Scan port of the target FPGA and up to 15 ILA, IBA, VIO, and/or ATC2 cores (as shown in Fig. 19). For devices of the Spartan®-3, Spartan®-3E, Spartan®-3A, and Spartan®-3A DSP families, the ICON core uses either the USER1 or USER2 JTAG Boundary Scan instructions for communication via the BSCAN primitive. The unused USER1 or USER2 scan chain of the BSCAN primitive can also be exported for use in the application, if needed. For all other supported

devices, the ICON core uses any one of the USER1, USER2, USER3 or USER4 scan chains available via the BSCAN primitives. It is not necessary to export unused USER scan chains because each BSCAN primitive implements a single scan chain.

5.3.2 ILA Core

The ILA core is a customizable logic analyzer core that can be used to monitor any internal signal of the design. Since the ILA core is synchronous to the design being monitored, all design clock constraints that are applied to the design are also applied to the components inside the ILA core. The ILA core consists of three or three major components:

1). Trigger input and output logic:
- Trigger *input* logic detects elaborate trigger events
- Trigger *output* logic triggers external test equipment and other logic

2). Data capture logic:
- ILA cores capture and store trace data information using on-chip block RAM resources

3). Control and status logic:
- Manages the operation of the ILA core.

5.3.3 IBA/OPB Core

The Integrated Bus Analyzer for the CoreConnect® On-Chip Peripheral bus (IBA/OPB) core is a specialized logic analyzer core specifically designed to debug embedded systems that contain the IBM CoreConnect® On-Chip Peripheral Bus (OPB). The IBA/OPB core consists of four major components:

1). A protocol violation monitor:
- Detects and reports up to 32 violations of the IBM CoreConnect® OPB bus protocol.

2). Trigger input and output logic:
- Trigger *input* logic detects OPB bus and other user-defined events.
- Trigger *output* logic triggers external test equipment and other logic

3). Data capture logic:
- Captures and stores trace data information using on-chip block RAM resources.

4). Control and status logic:
- Manages the operation of the IBA/OPB core.

The IBA core for the IBM CoreConnect® On-Chip Peripheral Bus (IBA/OPB) is used to monitor the CoreConnect® OPB bus of embedded MicroBlaze soft processor or Virtex-4 FX and Virtex-5 FXT families PowerPC™ hard processor systems. Up to 16 different trigger groups can be monitored by the IBA/OPB core at any given time. The IBA/OPB core can also implement the same trigger and storage qualification condition equations as the ILA core. Note that a description on how to generate and use the IBA/OPB core in an embedded processor design can be found in the *ChipScope™ OPB IBA data sheet* [38] and the EDK Platform Studio online help [36].

5.3.4 IBA/PLB Core

The Integrated Bus Analyzer for the CoreConnect® Processor Local Bus v3.4 (IBA/PLB) core is a specialized logic analyzer core specifically designed to debug embedded systems that contain the IBM CoreConnect® Processor Local Bus (PLB). Note that the IBA/PLB core is used only for PLB versions prior to PLB v46. For PLB v46 buses, a customized ILA core (*chipscope_plbv46_iba*) is attached to the PLB v46 bus using the Xilinx Platform Studio tool. The IBA/PLB core consists of three major parts components:

1). Trigger input and output logic:
- Trigger *input* logic detects PLB bus and other user-defined events.
- Trigger *output* logic triggers external test equipment and other logic.

2). Data capture logic:
- Captures and stores trace data information using on-chip block RAM resources.

3). Control and status logic:
- Manages the operation of the IBA/PLB core.

The IBA core for the IBM CoreConnect® Processor Local Bus (IBA/PLB) is used to monitor the PLB bus of embedded MicroBlaze soft processor or Virtex-4 FX and Virtex-5 FXT families PowerPC™ hard processor systems. Up to 16 different trigger groups can be monitored by the IBA/PLB core at any given time.

The IBA/PLB core can also monitor other generic design signals (using the TRIG_IN trigger group) in addition to the PLB bus signals. This capability allows the user to correlate events that are occurring on the PLB bus with events elsewhere in the design. The IBA/PLB core can also be connected to other capture cores using

the TRIG_IN and TRIG_OUT port signals to perform cross-triggering operations while monitoring different parts of the design.

The IBA/PLB core is also able to implement the same trigger and storage qualification condition equations as the ILA core. Note again that a description on how to generate and use the IBA/PLB core in an embedded processor design can be found in *ChipScope™ PLB IBA data sheet* [39] and the EDK Platform Studio online help [36].

5.3.5 VIO Core

The Virtual Input/Output (VIO) core is a customizable core that can both monitor and drive internal FPGA signals in real time. Unlike the ILA and IBA cores, no on- or off-chip RAM is required. Four kinds of signals are available in a VIO core:

1). Asynchronous inputs:
- These are sampled using the JTAG clock signal that is driven from the JTAG cable.
- The input values are read back periodically and displayed in the Analyzer.

2). Synchronous inputs:
- These are sampled using the design clock.
- The input values are read back periodically and displayed in the Analyzer.

3). Asynchronous outputs:
- These are defined by the user in the Analyzer and driven out of the core to the surrounding design.
- A logical 1 or 0 value can be defined for individual asynchronous outputs.

4). Synchronous outputs:
- These are *defined* by the user in the Analyzer, *synchronized* to the design clock and *driven out* of the core to the surrounding design.
- A logical 1 or 0 can be defined for individual synchronous outputs. Pulse trains of 16 clock cycles worth of 1's and/or 0's can also be defined for synchronous outputs.

Activity Detectors

Every VIO core input has additional cells to capture the presence of transitions on the input. Since the design clock will most likely be much faster than the sample period of the Analyzer, it's possible for the signal being monitored to transition many times between successive samples. The activity detectors capture this behavior and the results are displayed along with the value in the Analyzer. In the case of a synchronous input, activity cells capable of monitoring for asynchronous and synchronous events are used. This feature can be used to detect glitches as well as synchronous transitions on the synchronous input signal.

Pulse Trains

Every VIO synchronous output has the ability to output a static 1, a static 0, or a pulse train of successive values. A pulse train is a 16-clock cycle sequence of 1's and 0's that drive out of the core on successive design clock cycles. The pulse train sequence is defined in the Analyzer and is executed only one time after it is loaded into the core.

5.3.6 ATC2 Core

The Agilent Trace Core 2 (ATC2) [40] is a customizable debug capture core that is specially designed to work with the latest generation Agilent logic analyzers. The ATC2 core provides external Agilent logic analyzers access to internal FPGA design nets (as shown in Fig. 21).

ATC2 Core Data Path Description

The data path of the ATC2 core consists of:

1). Up to 64 run-time selectable input signal banks that connect to the user's FPGA design.

2). Up to 128 output data pins that connect to an Agilent logic analyzer's probe connectors.

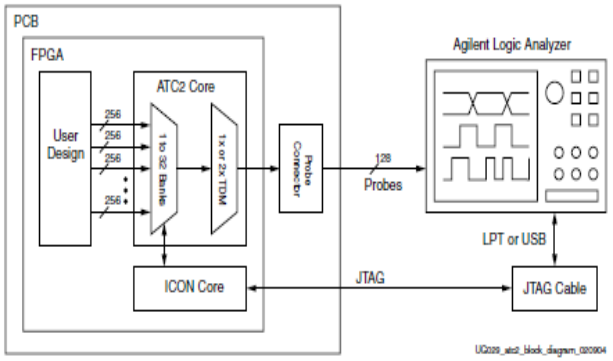


Fig. 21. ATC2 core and system block diagram.

3). Optional twice time-division multiplexing (TDM) available on each output data pin that can be used to double the width of each individual signal bank from 128 to 256 bits.
4). Supports both asynchronous timing and synchronous state capture modes.
5). Supports any valid I/O standard, drive strength, and output slew rate on each output data pin on an individual pin-by-pin basis.
6). Supports any Agilent probe connection technology [35].

The maximum number of data probe points available at run time is calculated as:
(64 data ports) * (128 bits per data port) * (2x TDM) = 16,384 probe points.

ATC2 Core Data Capture and Run-Time Control

The external Agilent logic analyzer is used to trigger on and capture the data that passes through the ATC2 core. This allows one to take full advantage of the complex triggering, deep trace memory, and system-level data correlation features of the Agilent logic analyzer as well as the increased visibility of internal design nodes provided by the ATC2 core. The Agilent logic analyzer is also used to control the run-time selection of the active data port by communicating with the ATC2 core via a JTAG port connection (as shown in Fig. 21).

5.3.7 IBERT Core

The IBERT core has all the logic to control, monitor, and change transceiver parameters and perform bit error ratio tests. The IBERT core has three major components:

1). IBERT Logic
The IBERT logic instantiated the actual transceiver component, and contains the pattern generators and checkers. A variety of patterns are available, from simple clock-type patterns to full PRBS patterns to framed counter patterns utilizing commas and comma detection.
2). Dynamic Reconfiguration Port (DRP) Logic
Each transceiver has a Dynamic Reconfiguration Port (DRP) on it, so that transceiver attributes can be changed in system. All attributes and DRP addresses are readable and writable in the IBERT core. Each transceiver's DRP can be accessed individually.
3). Control and status logic
Manages the operation of the IBERT core.

IBERT Design Flow

Since the IBERT is a self-contained design, the design flow is very simple. When using the ChipScope™ IBERT Core Generator to generate IBERT core designs for Virtex-4 and Virtex-5 devices, the design directory and bit file name are specified, options are chosen, and the Generator runs the entire implementation flow, including bitstream creation, in one step.

The design flow for generating IBERT core designs for Virtex-6 and Spartan-6 devices are very similar except the Xilinx CORE Generator tool is used. The main difference is that the design directory and device information is specified in the Xilinx CORE Generator project. In both cases, it is not mandatory or required to explicitly run any other Xilinx software to generate an IBERT core design bit file.

6. Conclusion

Model-based approaches toward embedded processor system design have been presented in this paper. The major advantage of the proposed model-based design approach is that a physical design can be modeled without a prior need for the knowledge regarding the hardware design, architecture and implementation routines provided the MATLAB® floating algorithms describing the model-based design is synthesizable.

The model-based embedded processor system design approach presented in this paper begins from the algorithmic modeling and synthesis using the Xilinx AccelDSP and a complete model-based modeling and further synthesis using Xilinx System Generator for DSP. The Xilinx embedded development kit (EDK) which bundles the Xilinx platform studio (XPS) and the Xilinx software development kit (SDK) constitutes the tools for developing the hardware and software portions for the completed embedded system.

Three very important advanced FPGA design, simulation and real-time verification tools have

been introduced and discussed, namely: *1)* PlanAhead™ design and analysis tool, *2)* ModelSim™ verification and simulation tool, and *3)* ChipScope™ Pro logic analysis, real-time debugging and performance verification tool.

None of all the FPGA related articles considered in this paper investigates the performance capability of any of these three tools. Despite the fact that embedded processors can improve the over all system performance as reported in [11], [12], [13], [15], [41]; none of the articles reviewed in this paper is concerned with any embedded processor(s) in their designs.

In line with the evolving field of FPGAs and their high-end applications, three works are currently under development based on the methodologies proposed and presented in this paper. First, the trade-off between a hard- and soft- processor core will be investigated. On the basis of their performance comparison, a complete FPGA-in-the-loop implementation will be performed, and then an integrated embedded multi-processor system will be investigated. The target FPGA device is the Xilinx Virtex-5 FX70T ML507 FPGA board designed specifically for industrial embedded systems applications.

REFERENCES

[1] V. A. Akpan (Nov., 2009), FPGA Embedded Systems Design Technologies: with an Overview of Xilinx Systems Design Tools, Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Greece, pp. 1 – 31. [Online] Available: http://users.auth.gr/~iosamar/technicalreports.htm.

[2] H. Jack, "Automating Manufacturing System with PLCs," Version 4.2, April 3, 2003, http://clay-more.engineer.gvsu.edu/~jackh/books.html

[3] J. Ganssle and M. Barr, "Embedded Systems Dictionary," CA 95020, USA: CMP Books, 2003

[4] C. Maxfield, "The Design Warrior's Guide to FPGAs: Devices, Tools and Flows," Oxford, UK: Newnes – Elsevier, 2004.

[5] R. Dubey, "Introduction to Embedded System Design Using Field Programmable Gate Arrays," London: Springer-Verlag, 2009.

[6] S. Kilts, "Advanced FPGA Design: Architecture, Implementation, and Optimization," New Jersey, USA: John Wiley & Sons, 2007.

[7] E. O. Cardenas and R. J. Romero-Troncoso, "MLP neural network and on-line backpropagation learning implementation in a low-cost FPGA," GLSVLSI'08, May 4 – 6, 2008, Orlando, Florida, USA, pp. 333 – 338.

[8] Virtex- FXT FPGAs, http://www.xilinx.com/products/virtex5/fxt.htm.

[9] G. Moertti, System-level design merits a closer look: the complexity of today's designs requires system-level, EDN Asia, February, 01 2002, pp. 22-28. [Online] Available: http://www.ednasia.com/article-1129-systemleveldesignmeritsacloserlook-Asia.html.

[10] G. Martin, The future of high-level modelling and system level design: Some possible methodology scenarios, Cadence Design Systems. [Online] Available: http://www.eda.org/edps/edp02/PAPERS/edp02-s1_1.pdf.

[11] Altera, www.altera.com, 2011.

[12] Xilinx, www.xilinx.com, 2011.

[13] V. A. Akpan (2011), "*Development of new model adaptive predictive control algorithms and their implementation on real-time embedded systems*", Ph.D. Dissertation, Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Greece, pp: 1 – 517, [Online] Available: http://invenio.lib.auth.gr/record/127274/files/GRI-2011-7292.pdf.

[14] S. Guccione, "List of FPGA-based computing machines", [Online] Available: http://www.io.com/~guccione/HW-list.html.

[15] B. H. Fletcher, "FPGA embedded processors: Revealing true system performance", Embedded Systems Conference, San Francisco, 2005, pp. 1 – 18.

[16] E. O. Cardenas and R. J. Romero-Troncoso, "MLP neural network and on-line backpropagation learning implementation in a low-cost FPGA," *In Proc. of the 18 ACM Great Lakes Symposium on VLSI (GLSVLSI'08)*, Orlando, Florida, USA, May 4 – 6, 2008, pp. 333 – 338.

[17] A. Malinowski and H. Yu, "Comparison of embedded system design for industrial applications". *IEEE Transactions on Industrial Informatics*, vol. 7, no. 2, pp. 244 – 254, May 2011.

[18] E. Monmasson, L. Idkhajine, M. N. Cirstea, I. Bahri, A. Tisan and M. W. Naouar, "FPGAs in industrial control applications". *IEEE Transactions on Industrial Informatics*, vol. 7, no. 2, pp. 224 – 242, May 2011.

[19] M. Pérez, M. Vásquez, J. Rodríguez and J. Pontt, "FPGA-based predictive current control of a three-phase active front end rectifier". *IEEE Int'l Conf. on Industrial Technology*, Gippsland, 10 – 13 Feb., 2009, pp. 1 – 6.

[20] The MathWorks Inc., MATLAB & Simulink R2009a, Natick, USA. www.mathworks.com.

[21] L. Wang, "Model Predictive Control System Design and Implementation Using MATLAB," London: Springer-Verlag, 2009.

[22] Y. Shoukry, M. W. El-Kharashi and S. Hammad, "MPC-on-chip: An embedded GPC coprocessor for automotive active suspension systems," IEEE Embedded Systems Letters, vol. 2, no. 2, pp. 31 – 34, Jun. 2009.

[23] Y. Shoukry, M. W. El-Kharashi and S. Hammad, "Networked embedded generalized predictive control for an active suspension system," 2010 American Control Conference, Baltimore, MD, USA, June 30 – July 2, 2010, pp. 4570 – 4575.

[24] K. V. Ling, B. F. Wu and J. M. Maciejowski, "Embedded model predictive control (MPC) using a FPGA," Proc. 17th World Congress, *The International Fed. of Aut. Control, Seoul, Korea*, July 6 – 11, 2008, pp. 15250 – 15255.

[25] Y. Maeda amd M. Wakamura, "Simultaneous perturbation learning rule for recurrent neural networks and its GPGA implementation," IEEE Trans. Neural netw., vol. 16, no. 6, pp. 1664 – 1672, Nov. 2005.

[26] P. Meloni, S. Secchi and L. Raffo, "An FPGA-based framework for technology-aware prototyping of multi-core embedded architectures," *IEEE Embedded Systems Letters*, vol. 2, no. 1, pp. 5 – 9, Mar. 2010.

[27] Xilinx PlanAhead Methodology Design, Release 11.4, UG633 (v11.1.0) March 3, 2009.

[28] Xilinx PlanAhead User Guide, UG632 (v11.4), December 1, 2009.

[29] Xilinx PlanAhead Tutorial: Qick Front-to-Back Flow Overview, Release 11.2, UG673 (v11.2) June 24, 2009.

[30] Mento Graphics Corporation, ModelSim Tutorial, Software Version 6.4b, 2008. www.mentor.com.

[31] Mento Graphics Corporation, ModelSim User's Guide, Software Version 6.4b, 2008. www.mentor.com

[32] Tcl Developer Xchange, http://www.tcl.tk/.

[33] ISE Design Suite Software Matrix, www.xilinx.com/publications/matrix/Software_matrix.pdf

[34] PlanAhead™ Design Analysis Tool, http://www.xilinx.com/tools/planahead.htm, 2009.

[35] ActiveState, http://www.activestate.com/

[36] EDK Platform Studio Online Help, http://www.xilinx.com/products/design-tools/ise-design-suite/embedded-edition.htm

[37] System Generator for DSP, www.xilinx.com/ise/optional_prod/system_generator.htm, 2011.

[38] DS282, ChipScope OPB IBA Data Sheet, http://www.xilinx.com/bvdocs/ipcenter/data_sheet/chipscope_opb_iba.pdf, 2009.

[39] DS283, ChipScope PLB IBA Data Sheet, www.xilinx.com/bvdocs/ipcenter/data_sheet/chipscope_plb_iba.pdf, 2009.

[40] Agilent Technologies, www.agilent.com/find/logic.

[41] J. Young and B. Machesney, "Reality check: A guide to understanding optimized processor core", *Synopsys Inc., White Paper*, pp. 1 – 10, September 2011.

Vincent A. Akpan (M'08) obtained the B.Sc. degree in Physics from the Delta State University (DELSU), Abraka, Nigeria in 1997; a Master of Technology (M.Tech.) degree in Instrumentation from The Federal University of Technology, Akure (FUTA), Nigeria in 2003; and a Ph.D. degree in Electrical & Computer Engineering from the Aristotle University of Thessaloniki (AUTH), Thessaloniki, Greece in 2011.

Between 1998 and 2004, he was a Graduate Assistant with the Department of Physics, DELSU. Since 2004 he has been with the Department of Physics Electronics, FUTA where he is currently Lecturer II. His integrated research interests include: computational intelligence, system identification, adaptive predictive control, real-time embedded systems, signal processing & machine vision, and Robotics & Automation. He is the co-author of a book and has authored or co-authored over 20 articles in refereed journals and conference proceedings. He is a regular reviewer in 6 international scientific/academic journals and several IEEE sponsored conferences. He is also an editorial board member for the American Journal of Intelligent Systems.

Dr. Akpan is a member of the IEEE, USA and The IET, UK. He was one among two Nigerian recipients of the 2005/2006 Greek State Scholarship (IKY) for a Ph.D. programme tenable in Greece.

Congestion Control Mechanisms and Patterns of Call Distribution in GSM Telecommunication Networks: The Case of MTN Nigeria

[1]**Mughele.E. Sophia & Wole Olatokun**
Africa Regional Center for Information Science (ARCIS)
University of Ibadan
Ibadan, Nigeria

Tunde Adegbola
Africa Languages Technologies Initiative (ALT-i)
Ibadan, Nigeria

[1]**Corresponding Author**
Doctoral Student (ARCIS) - prettysophy77@yahoo.com

ABSTRACT
Congestion remains a major challenge to telecommunications service provision both to the service providers as well as the subscribers. We attempted to determine the causes of congestion and evaluated congestion control mechanisms adopted by service providers in Nigeria using MTN Nigeria as a case study. A survey research design methodology using stratified random sampling technique was employed for selecting six states out of 36 states in Nigeria that formed the target population. Automated data from MTN's data logging machine was the main data collected and supplemented by unstructured interview. Six research questions were stated and tested in line with the set objectives. Data analysis was done using MS Excel and results presented using descriptive statistics. Findings from the study showed that apart from the carrying capacity of the MTN network other factors such as the use of phones for data transfer and multimedia activities contributes to traffic congestion on the network. Other external factors such as vandalization of network equipments, weather and high rise buildings also contributes negatively in varying degrees to service sustainability. Recommendations were made based on these findings.

Keywords, ISDN, **QOS,** Transceiver Station, Mobile Switching Station & Base Station Subsystem

1. INTRODUCTION

According to the International Telecommunication Union, by 1996 Nigeria's teledensity ratio was a mere 0.3. It rose slightly to 0.4 by 1999. Nigeria's teledensity is a far cry from the African average of 1.67. As at 1999, Nigeria had limited telephone network with a long waiting list of subscribers.

African Journal of Computing & ICT Reference Format:

E.S. Mughele, W. Olatkun & T. Adegbola, T. (2012) - Congestion Control Mechanisms and Patterns of Call Distribution in GSM Networks – The Case of MTN Nigeria. Afr J. of Comp & ICTs. Vol 4, No. 3. Issue 2 pp 29-- 42.

- ISSN 2006-1781

Even the NCC admits that Nigeria had a very limited telephone network for many years, and the waiting list is estimated at over 10 million people, who have applied to the Nigeria Telecommunications Limited, NITEL, for services. However, with the liberalization of the telecommunication industry in 2001, the story changed dramatically. The teledensity ratio had tripled within just five years of GSM operations. By May 2005, Nigeria with an estimated population of 128,771,988 had more than 9 million GSM subscribers, making the country one of the fastest growing GSM markets in the world (Ajala, 2005).

Nigerian telecommunication market looks forward to achieve a tele-density of 100% by 2020 driven by massive telephone and mobile communication improvements, calling for a huge growth in the ICT sector. Nigeria, with a population of nearly 140 million people being serviced by 60 phone lines, a 45% success rate in a development where in it will boost the Nigerian economy and infrastructure. No modern economy can thrive without an integral information technology and telecommunications infrastructure. This is because ICTs provide the veritable platform for development across the economic and other sectors if well harnessed, (Ndukwe, 2008).

According to (Ndukwe, 2008), if the goal is achieved, Nigeria would then have as many phone lines which can impressively improve the lives of the local people in as far as communication through telephone, cell phone and internet is concerned. Today the ratio stands at 1:1 implying that a Nigerian individual living in an urban centre owns three communication lines in the sense that at work and home an individual will be serviced by telephone lines and a mobile cellular phone, to make it three gadgets of communication

At the moment, there are five major Global System for Mobile (GSM) Telecommunication operators in Nigeria: MTN, Airtel, GloMobile, Etisalat and MTEL. MTN enjoys the greatest patronage, with over 35.1 million subscribers. It was predicted that between 2003 and 2006, Nigeria's GSM market would be Africa's fastest-growing mobile market, and this prediction had long been fulfilled. The competition is getting fiercer by the day as operators have to compete desperately for the same potential subscribers (Ajala, 2005).

Congestion is a problem all GSM service providers are facing and trying to solve; the first issue that needs to be tackled by a G.S.M operator is the provision of network coverage to the target population. Calls cannot be made or received in areas where there is no network. Again, where a network exists but has poor connectivity; calls may be difficult to make or receive. However, poor connectivity of call may occur as a result of a number of factors chief among, which is congestion. Congestion is a situation that arises when the number of calls emanating or terminating from a particular network is more than the capacity that the network is able to cater for at a particular time.

It causes call signals to queue on the transmission channel. Consequently, the rate of transfer of voice signals is reduced or quality of signals received become distorted or both. At worse, the calls will not connect at all. There are technical mechanisms along the transmission link that tend to create or worsen congestion. When a number is dialed on a G.S.M phone, the call is routed to the nearest base station where the Base Transceiver Station (BTS) receives, amplifies, and reroutes the call to the Base Station Controller (BSC). The BSC controls and manages single or multiple BTSs and communicates directly with the Mobile service Switching Center (MSC) with an interface called the “A” Interface .The MSC finally routs the calls to its destination after the credit status of the call have been confirmed by the pre-billing software. The balance of the capacity of these interfacing devices vice a vice the demand from subscribers determine the condition of congestion. Except where there are emergency calls and power outage, calls are transmitted at the required speed. The study intends to determine the causes of congestion in the GSM network, and to proffer likely solution that will reduce the problems associated with network congestion.

Today, operators are fast realizing that they are in a highly competitive environment where subscribers can make or break them. Dissatisfaction by subscribers gives rise to a high rate of subscriber churn and low revenue for the operator. The performance of the network has a direct impact on the revenues. The NCC (Nigerian Communication Commission) is bringing pressure to the operators to step up the quality of service offered to Nigerians and had even gone a step further to award contracts to private companies to conduct comparative analyses of the quality of service offered by each of the operators. The NCC is further threatening to sanction any operator that fails to pay attention to quality (Ajala, 2005). When congestion occurs it has two implications, firstly on the part of the subscribers, this can be seen on the time wasted. When there are limited servers to serve subscribers it result to heavy “traffic” which is termed as congestion. When calls are congested the wasted time on the queue can only be evaluated by the importance and urgency of the call, as such calls could depend on the life of an organization at that particular time of call, or it could be a life critical security issue.

Secondly, congestion has also serious implication on the service subscriber because when calls are congested it implies that the service providers are losing in terms of resources. The more calls get connected, the higher the increase in their resources, better maximizations of their bandwidth and service capacity. This study took a critical look at G.S.M architecture, its operations and the remote causes of congestions in the network, the control mechanism in use by MTN Nigeria, and recommended a framework for better congestion control mechanisms. The following research questions guided the study.

1. What is the system capacity of the BTS, BSC and MSC deployed to various MTN locations throughout the country?
2. How does the system capacity of MTN compare with the number of activated telephone lines?
3. What is the average number of incoming calls during the peak periods of the day? How does this compare with the system capacity?
4. What is the average number of call congested per day?
5. What other technical hitches associated with equipment failure and/or power failure, which contribute directly or indirectly to congestion in the MTN network?
6. Does MTN apply any control mechanisms to check the hindrances occurring in the signaling network? If yes, what are the types of control mechanisms been applied? And at what levels are they applied?

2. RELATED WORKS

As Defined by Kuboye (2010), Congestion is the unavailability of network to the subscriber at the time of making a call. Congestion occurs when there are limited resources at the service point, this leads to queues, called traffic or congestion. It is the situation when the blocking occurs and no free path can be provided for an offered call (Syski, 1986). That is, when a subscriber cannot obtain a connection to the wanted subscriber immediately. The ideal telephone system is a situation where it is possible for all subscribers to talk in pairs simultaneously.

If one connecting device be allocated for a pair of subscribers, then the number required will be too high to be reasonable (Syski, 1986). Such an ideal system is impracticable because of its enormous size, very high cost and maintenance difficulties. It is therefore necessary to reduce the number of connecting devices which means that the subscribers are confronted with the possibility that some of their calls may be unsuccessful. The reduction in number of connecting resources consequently leads to reduction in the number of conversation which can take place simultaneously.

2.1 Causes of congestion

Congestion simply means 'full of traffic'. It occurs when too many things or people are seeking for use of a resource that is limited. The Oxford-Advanced Learner's Dictionary (year) defines congestion as the state of being crowded, blocked or too full of traffic. Congestion is the unavailability of network to the subscriber at the time of making a call. It is the situation when the blocking occurs and no free path can be provided for an offered call (Syski, 1986).

It manifests in various ways, hence it is usually an everyday experience. It therefore, makes an economic sense for congestion to occur because naturally, in every real world situation resources most time are limited and it is necessitated by an unequal relationship between services required and the facilities used for rendering the service. Congestion often results in queues. Queuing is seen as a line of things or people waiting to use a particular resource. For instance, a traffic jam usually occurs when there are many vehicles prying a road at the same time.

The road will be filled (congested) with vehicles that are moving in succession. This scenario is replayed almost in every human activity. It occurs when patients wait to be attended to in a hospital; when bank customers queue up in banking halls waiting for service and; when travelers board aircraft, etc. However, congestion is not restricted to human activities as it could also take place in data transmission. Congestion in signaling network may well be the visible symptoms caused by signaling traffic being redirected due to failure of the network component, (Lars, et al 2001) and (Stefan, and Ake, 2002).

Manfield, (1993) posited that congestions are more likely to arise from traffic redirections at network component failure or by an extremely high call density to one specific node. Congestion inhibits the normal flow of data signals thereby resulting in these signals queuing up in order to be allocated resources by the server. The time it takes these signals to wait in the queue is a determinant of the capacity of the server. The capacity of the network equipment, literally, has been the number of data signals a system can handle.

2.2 Network Capacity

Various attempts have been made to define the capacity of the network; the Erlang capacity, which has been used in telephony networks, is a probabilistic definition. It specifies the arrival rate of calls that the system can allow so that the probability of blocking of calls on arrival is lower than some threshold (Erlang, 2000 in Altman, 2004). Altman (2004) in giving another definition to capacity, which is more appropriate for wireless phone, declares that capacity is taken to be the rate of calls that the system can allow so that the probability that the quality of service not attained is sufficiently minimal. Here calls are not blocked when exceeding the limit of the system to provide the required quality of service. From the above definitions, the transmission rate used by a cell is a fixed constant, which may be class dependent. Most packet losses observed over a unicast audio connection are caused by congestion in the network.

As mentioned earlier that GSM were used for data transmission during the 2.5G and 3G era, so when many packets of messages are sent they are routed through their various transmission channels, collision may occur due to congestion and such packets can be destroyed. On other occasions, such packet of messages can be on the queue for a considerably long period of time; and when they are not routed to their destination because of lack of resources such packets are destroyed. The state of congestion might have been created by the audio connection, by other connections sharing common resources with the audio connection, or by both. Therefore GSM service providers should make appropriate specification for the sources and the bandwidth requirement for such source in order to reduce collision and packet loss during routing of packets in the transmission channels.

Typically, sources react to congestions by decreasing their bandwidth requirements in an attempt to reach a state where network utilization is high. If all sources use the same bandwidth control mechanism, and if this mechanism is designed adequately, then all resources share the resources of the network fairly (Kelly, 1991).In practice, however, not all resources use the same mechanism. Furthermore, some sources do not use any control mechanism at all. As a consequence, an audio source, which decreases its bandwidth requirements upon detecting congestion, might not observe any decrease in its loss rate as a result of its action. Thus, it is difficult to evaluate and to control the impact of a bandwidth (or rate) control action taken by a source on the state of the network in general and on the loss rate for this source-destination pair in particular.

2.3. The Stateless Architecture of the Internet

The problem essentially stems from the current stateless architecture of the Internet illustrated by the (FIFO) first in first out discipline at the switches, which makes the delay, and loss processes of a connection strongly dependent on the arrival processes of other connections. Some wireless network allows for multimedia application and new services such as file transfer Internet browsing. These applications are non-interactive and as such are less sensitive to the assigned throughput. The applications can be considered as part of best effort service whereby the Base station can assign the rate of transmission. For a given rate of arrival of best effort calls, the system's capacity will depends on the assigned throughput. However, it is assumed that the total volume Vs of traffic created by an application S does not depend on the assigned traffic, the time it takes this session and the resources that is occupied in the network is given by Vs/R (S) where R (S) is the assigned transmission rate. It is assumed that the power control of the energy per bit of the best-effort application does not depend on its transmission rate.

Adugu, (2004) opines that the effects of non-ideal power control however can be included in the model by a simple transformation of the problem's parameters. The capacity of best effort traffic that is expressed is obtained with restriction to a single cell. There is also the combination study of best effort with real time (non best - effort) cell classes as well as the

discussion of extensions to the multi – cell case. Lowering transmission rates has two conflicting impacts: on one hand, one less resource is needed at a given time for handling the connections whose transmission rates were slowed down. On the other hand, (in contrast with real time connections), the duration of transmission of such cells becomes larger because the amount of information to be transmitted does not depend on the transmission rate. The overall result of these competing effects is that slowing the transmission rates increases capacity suggesting that there is a limit capacity that can be approached when slowing down the transmission rates.

As long as the arrival rate of information is below some level, blocking probability can become arbitrarily small by sufficiently slowing down the transmission rates (this is similar to the behavior of the well known Shannon capacity in information theory, in which we can get an arbitrarily low error probability if we use sufficiently long codes, as long as we transmit at a rate below the capacity) (Eifel, 2003).
From the view point of the network, four basic elements are related to congestion or indicate that a call could not be completed, these are itemized below.

1. Traffic channels congestion (TCHC): Traffic channels (TCH) represent a voice channel and each call uses TCH. There are eight channels defined for each radio frequency carrier and most are used for traffic channels and some for control channels (Mehrotra, 1997). When there is no free voice channel (TCH), then, we have traffic channels congestion (TCHC).

2. Dedicated control channel congestion (DCHC): Standalone dedicated control channel (SDCCH) is to provide authentication to mobile station, location updating and assignments to voice channel (TCHs) during idle periods (Mehrotra, 1997). When making a call or responding to paging message for the allocation of an SDCCH for authentication, if there is no vacant SDCCH to use at that time, the call will be terminated. This failure is called the dedicated control channel congestion.

3. Common control channels congestion (CCCHC): Common control channel is a group of control channels that support the establishment and maintenance of communication links between the mobile stations and base stations (Harte et al., 1999). It consists of random access channel (RACH), paging channels (PCH) and access grant channel (AGCH). RACH is used to make request for network assignment, PCH is used to alert the mobile station of incoming call and AGCH is used to assign mobile station to a specific DCCH or SDCCH for onward communication. When any of these three control channels is congested, there cannot be any call establishment between the sender and receiver, then, we have CCCH congestion

4. Pulse code modulation congestion (PCMC): Pulse code modulation (PCM) or E1 is the link required to connect the base station (BS) and mobile-switching center (MSC) together. Each PCM can carry between 1 and 32 calls. When there is no free PCM to carry the call signals between the BS and MSC, then we have pulse code modulation congestion.

3. METHOD

A pilot study was conducted on October 6, 2006, at the MTN national networking headquarters. The Radio Frequency Analyser Manager was interviewed, and the researcher also interacted with the Network Geographic Manager, who explained how poor whether condition could impact on congestion and measures put in place by MTN to maximize the weather in order to improve service rate. The researchers also observed the data logging system. Information obtained from the pilot study opened up new areas for the study. Most importantly, the pilot study was very useful as at the time it was conducted because the study was at its preliminary stage. All relevant data and information needed to sharpen the focus of the study were obtained during the pilot study. Thus, the pilot study opened up new areas and ideas for the research work, which add to understand that, apart from studying the data gotten from the automated system, there could be other factors that can also cause congestion in the network. This will help to ensure that every variable within the framework of congestion and control mechanism are properly under studied.

3.1 Data Integrity

The data used for the study was obtained from an automated system, and it was machine generated therefore it is devoid of interference. Although there were missing data, hence, the use of implementation of interpolation method. This was due to restricted access to the data logging system; interpolation was possible because it has been observed that there is a pattern in the behaviour of the data.

3.2 Unstructured Interview

This study also employs unstructured interview, the Radio Frequency Analyzer Manager in charge of the data logging system of MTN was interviewed in order to obtain other variables that causes congestion and measures taken by MTN to reduce congestion.

3.3 Sampling techniques and procedures

The study determined the effectiveness of the congestion control mechanism in MTN GSM network nationwide; hence MTN Nigeria is the target population for this study. The complete list of units for this study was an alphabetical list of the 36 states in Nigeria including the Federal capital territory covered by the MTN network. The complete listing served as the sample frame for the study. The network coverage of MTN is all the 36 states in Nigeria, in line with the topic under study; the capacity of the network was determined in relation to quality of service rendered in respect of congestion reduction in the network. Stratified sampling method was adopted with the 36 states in the country divided into six strata. A strata is equal to a zone and one state each will be randomly selected from each stratum.

The random method used for selection for each of the stratum was the hat method. The first stratum consisted of 5 states, the numbers 1-5 is written in a piece of paper i.e. each of the numbers are written ten times, all together gives fifty (50) numbers all in one hat. The numbers were selected five different times randomly, the number that occurred most frequently is selected (mode) and any state attached to the number is also selected to represent that strata. Meanwhile, for each of the strata the states are arranged in an array, and numbers are attached to each state in an ascending order starting from number 1.

This process was repeated for each of the strata, the second strata consist of seven states, numbers 1-7. Ten of each number was represented in the hat, altogether makes 70, random selection was conducted five times and the mode was selected with the state attached to it to represent the strata, this procedure was done for the remaining strata. The six geopolitical zones were North-west, North-east, North-central, South-west, South-east, and South-south. Each state was randomly picked from each of the geopolitical zones; in total five states including the federal capital

The randomly selected states from the six geopolitical zones were Kano, Adamawa, Abuja (F.C.T), Lagos, Enugu, and Rivers states, the state capital of each state was be used to analyse the behaviour of calls made in each of the locations with respect to congestion of calls and its control in the MTN network. Analysis of the network performance of Nigeria was undertaken, in order to determine the capability of the network. This was be done by comparing the number of calls that the network receives on a daily basis, to the number of congested calls.

4. DATA COLLECTION AND ANALYSIS

The data used for this study was primary, and was collected directly from MTN network system from the data logging system. The method of data collection guarantees the integrity of the data. In order to ensure data integrity, the researcher, had the privilege to observe the network data logging system with the organisation's network Engineer for six (6) hours daily for 3 consecutive days. This period, including the knowledge gathered during the pilot study conducted on the data, prior to the 3 days of observation, allowed for adequate knowledge of the capability of the network to connect or congest calls at various locations around the country at designated periods of the day. The knowledge acquired from the pilot study also informed other questions on the causes of congestion in the network, apart from the data that could be got from the data logging software.

An unstructured interview and an automated data logging software were the major instruments employed for collection of data. These methods were used because they were the only means to get accurate data on every activity relating to the capacity of the network in relation to quality of service rendered and service rate. Such vital information as: the number of calls the network receives on a daily basis, the number of calls congested, number of calls connected, maximum number of calls the system could handle at peak periods, were observed, evaluated and collected.
Moreover, the use of unstructured interview was also adopted to reduce or eliminate the problem of likely omission of relevant factors other than systems capacity to handle calls adequately, that can also cause congestion directly or in directly in the network. The unstructured interview complemented the data collected through the automated machine (machine generated). The two data collection instruments were however, constructed such that they will adequately and sufficiently cover every aspect of the research questions and the objectives of the study.
Percentages and ratio distribution methods were employed in order to determine the efficiency of the congestion control mechanism used in MTN network nationwide and at specified locations. The hypotheses formulated for the study were tested using correlation analysis.

5. RESULTS AND DISCUSSION OF FINDINGS

5.1 System capacity of the BTS, BSC, and MSC deployed to various locations.

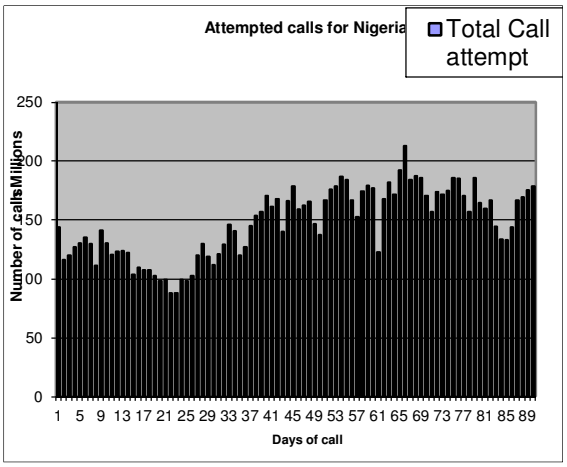
The system's capacity of the MTN network was obtained from the Radio Frequency Analyser Manager, Engr. Wasiu Otukoya of the MTN National Networking Headquarters in Victoria Island Lagos. The information was obtained through unstructured interview. Otukoya (2007) explained that the base transceiver station (BTS) cabinet of the MTN network can take twelve transceivers .On the average, the BTS of the MTN network deploys two cabinets and each cabinet contains twelve transceivers. It therefore consists of twenty-four transceivers and each of them is sectorized into three parts. There are eight transceivers per sector. This implies that one transceiver can accommodate eight-time slot. Meaning eight calls can be received simultaneously at a time, this can be expressed as 8*8 = 64, but not all the 64 are available for call communication, others are used for call set up, so four of the time slot are used for the call set up, hence we have 60 slots available, there are 60*3 time slot for each of the sectors, by implication, one MTN site can handle 180 calls simultaneously. This is the average capacity of BTS. On the other hand, the base station controller (BSC) capacity depends on how busy the transceivers are. Typically on the average, a BSC can handle between 30 to 50 BTSs'. One mobile switching centre can handle one base station controller. On the average, it is a ratio of 1: 1. Each of the BTS, BSC, and MSC in the various parts of the country has the same capacity except Ojota, which is MTN's biggest switching centre in Nigeria. (Otukoya, 2007)

5.2 Extent of adequate available capacity.

Number of calls attempted, established and congested in the MTN network for 91 days in Nigeria N = 91

Presented in Table 1 in the appendix, was the number of calls attempted, established and congested on a daily basis throughout Nigeria for 91 days (3 months). From Table 1, the total number of calls attempted in the MTN network for 91days is 13,249,678,058, the total number of call established is 12,965,155,325, and the number of call congested is 284,522,733. The average number of call attempted per day is 147,218,645, average number of calls established is 144,057,281 and congested calls are 3,161,364. Comparing the above data with the maximum capacity of the MTN network; which is the number of calls the BTS deployed all over the country can handle within a time slot. The MTN network has about 2,500 BTS nationwide; this implies that at peak period the BTS can accommodate 180*2,500 calls. This gives 450,000, which is the maximum capacity of the BTS at a time nationwide. When every transmission link is utilized, the extent of adequate available capacity can be evaluated when we compare the number of calls the network is capable of taking within a time slot which, as 450,000, and the average number of call that the network receives per day which, is 147,218,645. This implies that on the average per day 3,161,364, number of calls or subscribers suffer from congestion resulting from the MTN network.

5.3 Number of incoming calls compared with systems capacity



Average	147,218,645

Figure1: Attempted calls made in Nigeria for 91 days

Figure 1 shows the distribution of total numbers of attempted calls at the MTN network over 91days, which was 154,000,000 but the average number of call attempted at the network nationwide is 147,218,645. In order to evaluate the number of attempted calls at peak period of the day in relation to the MTN system capacity were determined based on the pattern of attempted calls per hour within a day. This captures the particular hour of the day that is regarded as peak period. The graph below shows the behaviour of calls per hour of the day, and the capacity of the network in Lagos.

Fig.2: Pattern of call distribution per hour in Lagos State

From Figure 2, Tassall represents total call attempted while Tcassall represents total calls established. It is also observed from the above graph that the peak period of the day is usually between 19hour and 20hour of the day. An auto correlation was conducted using the data that are represented in the Figure. The test shows that the data were repeated after 24hours, which implies that the Figure2 represents eight days and each repeated cycle of 24hour represents a day. Also the numbers of calls for the peak period for each of the eight days above are presented thus: 3,750,000, 3,650,000, 3,850,000, and 4,000,000 4,100,000 3,900,000 3,700,000 and 3,650,000.

The average number of incoming or attempted calls at peak period is 3,818,750. When this number is compared to the capacity of the network, (fig2), the total number of calls the network capacity could take at peak periods of the day for eight days is as follows 3,550,000, 3,500,000, 3,650,000, 3,850,000, 3,950,000, 3,750,000, 3,600,000, and 3,540,000. On the average, the capacity of the MTN network at peak period of the day is 3,673,750. When this is subtracted from the number of attempted calls at peak period (3,818,750 – 3,673,750 = 145,000), this implies that at peak periods of the day 145,000 calls are being congested in the MTN network in Lagos.

5.3 Hourly distribution of calls in Yola

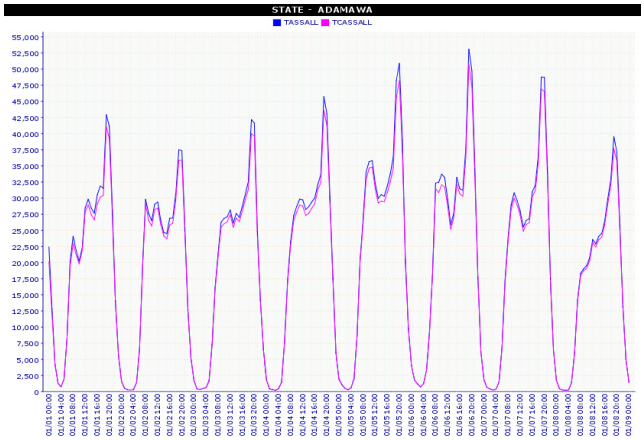


Fig. 3: Pattern of call distribution per hour in Yola, Adamawa State

From Figure 3, Tassall represent total call attempted while Tcassall represents total calls established. Observation from the Figure shows that the peak period of the day is usually between 19hour and 20hour of the day. Each successful crest represent a day, and the number of calls for the peak period in each of the eight days above are presented thus: 43,000, 37,500, 42,300, 46,250 51,000, 53,000, 48,750 and 39,800. The average number of incoming or attempted calls at peak period is 45,263. When this number is compared to the capacity of the network, from fig 4.3, the total number of calls the network capacity could take at peak periods of the day for eight days is as follows 41,250, 36,000, 40,000, 43,750, 48,500, 51,000, 47,000, and 37,500. On the average, the capacity of the MTN network at peak period of the day is 43,125. When this is subtracted from the number of attempted calls at peak period (45,263 – 43,125 = 2,138), this implies that at peak periods of the day 2,138 calls are being congested in the MTN network in Yola. This is very negligible compared to Lagos. There is still a pattern in the repetition of behaviour of the number of calls, which can be seen as a periodic behaviour from Monday to Sunday.

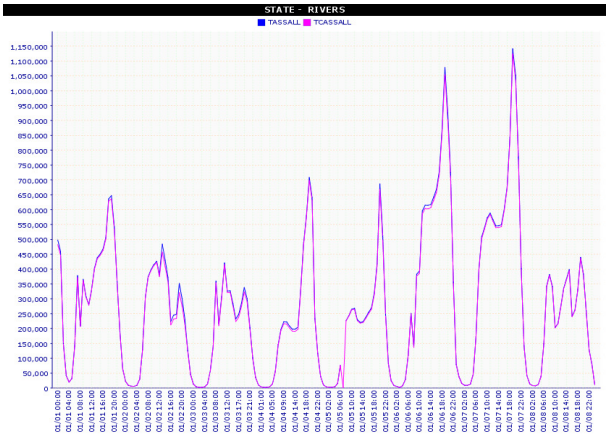


Fig 4: Pattern of call distribution per hour in Port-Harcourt River State

From Figure 4, Tassall represents total call attempted while Tcassall represents total calls established. It is also observed from the Figure that the peak period of the day is usually between 19hour and 20hour of the day. Each cycle represents a day, and the number of calls attempted at the peak period for each of the eight days above are represented thus: 650,000, 490,000, 425,000, 710,000 690,000, 1,080,000, 1,147,000 and 448,800. The average number of incoming or attempted calls at peak period of the day is 705,000. When this number is compared with the capacity of the network, from Fig 4.4, the total number of calls the network capacity can take at peak periods of the day for eight days is as follows 630,000, 460,000, 420,000, 700,000, 670,500, 1,060,000, 1,130,000, and 440,000. On the average, the capacity of the MTN network at peak period of the day is 688,750. When this is subtracted from the number of attempted calls at peak period, (705,000 – 688,750 = 16,250) this implies that at peak periods of the day 16,250 calls are being congested in the MTN network Port-Harcourt.

The Figure above displayed a different behaviour from the first two. From observation, the highest numbers of calls attempted and established were during the weekend especially Friday and Saturday. This could be attributed to the fact that in Port-Harcourt, social activities is usually very heavy from Friday evening to Saturday and this is accounted for because of the culture and life style of the people in Port-Harcourt.

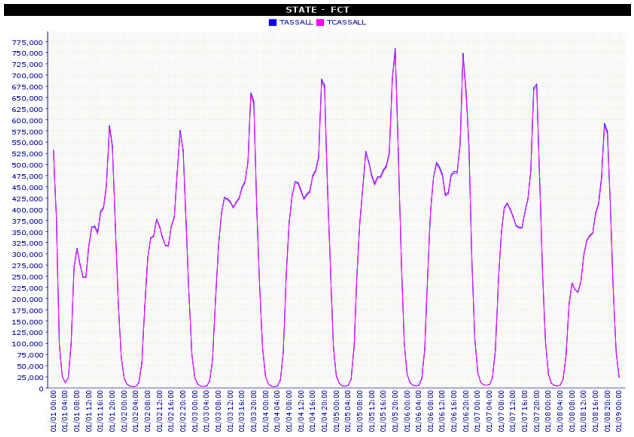


Fig 5: Pattern of call distribution per hour in Abuja Federal Capital Territory

From the Figure 5, Tassall represents total call attempted while Tcassall represents total calls established. It is also observed from the above graph that the peak period of the day is usually between 19hour and 20hour of the day. The graph above shows a regular repeated pattern and from the auto correlation data, it shows that each repetition represents a day.

The number of calls attempted at the peak period for each of the eight days above is presented thus: 585,250, 575,000, 660,000, and 695,500 765,000, 750,000, 680,000 and 595,800. The average number of incoming or attempted calls at peak period of the day is 663,219. When this number is compared to the capacity of the network, from fig 4.5, the total number of calls the network capacity take at peak periods of the day for eight days is as follows 580,000, 570,000, 650,000, 685,000, 750,000, 745,000, 675,000, and 585,000. On the average, the capacity of the MTN network at peak period of the day is 656,125. When this is subtracted from the number of attempted calls at peak period, (663,219 – 656,125 = 7,094), this implies that at peak periods of the day 7,094 calls are being congested in the MTN network Abuja. From the Figure above, there is a uniform and consistent periodic pattern of the behaviour of calls.

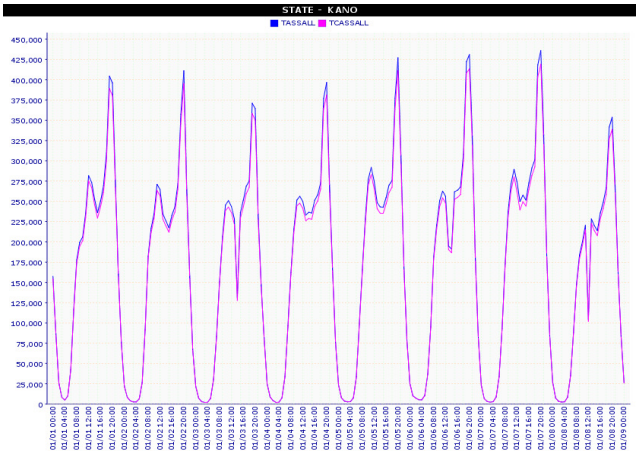


Fig. 6: Pattern of call distribution per hour in Kano

As shown in figure 6,, Tassall represents total calls attempted while Tcassall represents total calls established. It is also observed from the above graph that the peak period of the day is usually between 19hour and 20hour of the day. The above shows a regular repeated pattern and from the auto correlation data, it shows that each repetition represents a day. The number of calls attempted at the peak period for each of the eight days above is represented thus: 405,000, 412,500, 372,500, 397,500, 430,000, 435,000, 437,500 and 355,000.

The average number of incoming or attempted calls at peak period of the day is 405,625. When this number is compared to the capacity of the network, from fig 4.6, the total number of calls the network capacity can take at peak periods of the day for eight days is as follows 390,000, 397,500, 360,000, 385,000, 412,500, 415,000, 422,500, and 340,000. On the average, the capacity of the MTN network at peak period of the day is 390,313. When this is subtracted from the number of attempted calls at peak period, (405,625 – 390,313 = 15,312), this implies that at peak periods of the day 15,312 calls are being congested in the MTN network Kano.

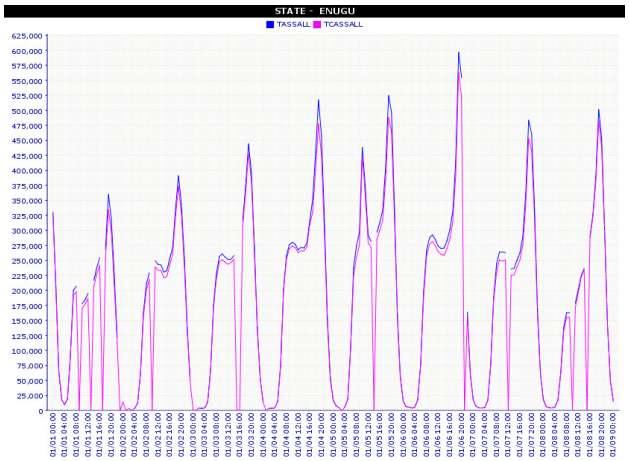


Fig.7: Pattern of call distribution per hour in Enugu, capital city of Enugu State

As illustrated in figure 7, Tassall represents total calls attempted while Tcassall represents total calls established. It is also observed from the above graph that the peak period of the day is usually 19hour to 20hour of the day. The graph above shows a regular repeated pattern and from the auto correlation data, it shows that each repetition represents a day. The number of calls attempted at the peak period for each of the eight days above is represented thus: 362,500, 390,000, 445,000, 520,000, 525,000, 600,000, 485,000 and 500,000. The average number of incoming or attempted calls at peak period of the day is 478,438. When this number is compared to the capacity of the network, from fig 4.7, the total number of calls the network capacity can take at peak periods of the day for eight days is as follows 335,000, 375,000, 427,500, 476,500, 487,500, 565,500, 451,500, and 485,000.

On the average, the capacity of the MTN network at peak period of the day is 450,438. When this is subtracted from the number of attempted calls at peak period, (478,438 – 450,438 = 28,000), this implies that at peak periods of the day, 28,000 calls are being congested in the MTN network Enugu.

5.4 Average number of calls congested daily

From the result, average number calls congested in the MTN network on a daily basis compared to the number of calls attempted in all of the 91 days. This number of calls congested indicates the number of attempted calls that suffers from congestion due to the inability of the system's capacity to meet the demands of the activated telephone lines nationwide.

5.5 Other factors affecting smooth flow of transmission

There were other various factors that affect the smooth flow of the network, the electronics can malfunction, and at times, it could be power failure, when this occurs there are site engineers to fix any equipment malfunction. Normally, there are two generators at each site, which are powered automatically once there is power failure. In some other cases, it can also be human vandals, who can vandalize these equipments. When such happens, it could lead to congestion, but then there are always engineers that can fix and attend to such technical issues.

6.0 ADOPTED MEASURES BY MTN TO OVERCOME CONGESTION

1) MTN has adopted various methods to reduce congestion, especially in the areas of poor network and dead spot; MTN controls congestion in the following ways:
2) Deployment of repeaters. This is a situation where there are two separate sites and in between is an area that is not within the network coverage. Instead of providing a site for that area, MTN mounts a repeater in order to maximize the capacity of both sites, the use of repeater is actually cheaper in term of cost and maintainability, rather than installing another BTS equipment.
3) Deployment of extended range site cell range. A GSM cell is limited to a range of 35km, it is a site that is slightly modified in that it can beat the normal 35km. Infact extended range can cover as long as 120km. This method helps to reduce the problem of dead spot.
4) Up grading site capacity of a cell that is badly impaired with the problem of congestion.

Some cell/site in the MTN network are sometimes badly impaired, example of such site is Cell54DB. A congestion control mechanism was applied in order to reduce congestion of calls. The diagram below shows the condition of the site before upgrading and after the upgrading is conducted on the site, the duration from when it was badly impaired to when it was upgraded was from August 5th to September 16th. (Otukoya, 2006).

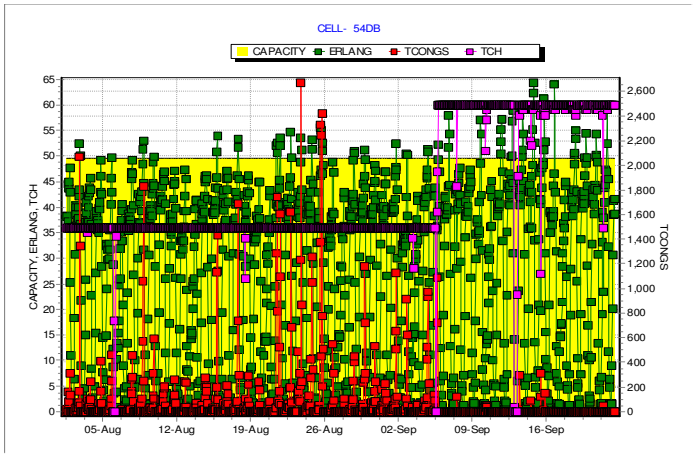


Figure 8: Upgraded MTN Cell/site, Cell54DB

The illustration in figure 8, shows a typical MTN site that was upgraded in order to reduce congestion. This demonstrates congestion control mechanism adopted by MTN at the Cell/site level. This site was badly impaired, so most calls connected to the site are congested because it has exceeded its capacity. Increasing traffic channel resources from 36 to 60 resulted in an increase in traffic from an average of 50 erlang/hour to 60erlang/hour. Where 50erlang/hour represents communication within the traffic channels between two users for 50minutes, after the upgrading of the site it was increased to 60erlang/hour thereby improving the capacity of the number of calls the cell can carry and congestion was also reduced to the beariest minimum. Also the number of calls congesting per hour reduced from a maximum of 2400 to less than 200 calls congesting per hour.

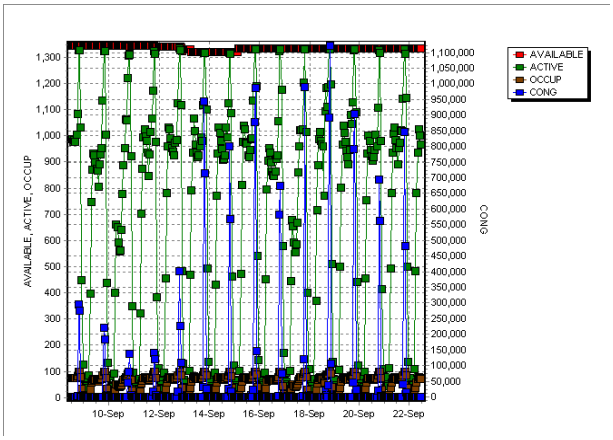


Fig 9: Congestion control at the BSC level

Figure 9 illustrates a typical MTN site that was upgraded in order to reduce congestion. This congestion control mechanism was adopted at the BSC level. There is congestion on this BSC and there is a need to install another BSC to reduce the congestion and to increase the capacity of the system in order to increase the number of BTS that the BSC can handle.

DISCUSSION

Findings from the study showed the capacity of each component in the MTN network. The BTS deployed to every location has the capacity to handle or receive 180 calls simultaneously within a time slot. Also one BSC has the capacity to handle between 30-50 BTS depending on busy nature of the BTS, hence the capacity of a BSC is between 5,400 and 9,000 call averagely within a time slot. One MSC has the capacity to handle one BSC. The extent of available capacity of network revealed that on the average, the capacity of calls the network can handle nationwide within a time slot is 450,000 calls, since MTN has about 2,500 BTS and one BTS can handle 180 calls at a time.

The network is able to handle 450,000 calls within a time slot, which is quite moderate, compared to the number of calls attempted. This result agrees with (Erlang, 2000 in Altman 2004) that stated that the arrival rate of calls that the system can allow is the capacity of system, so that the probability of blocking of calls on arrival is lower than the threshold. This is also confirmed in (Altman, 2004) in explaining the availability of the capacity of a network being adequate in a wireless phone or GSM network.

He declares that capacity is taken to be the rate of calls that the system can allow so that the probability that the quality of service is not attained is sufficiently minimal. Also, comparing the system capacity of MTN network with the number of calls that the network receives on a daily basis, showed that the system capacity is moderate in respect to quality of service considering the number of calls it is capable of handling within a time slot. On the capacity of the network of the six-location, understudied, the average capacity of the network in Lagos at peak period of the day is 3,673,750 calls; this is the average service rate of the capacity of the network for eight days.

When this value is compared with the average number of calls attempted for the same period which is 3,818,750 calls, congested call will be 3,818,750 –3,673,750 = 145,000. Also callers tend to attempt a particular number more than once when there is congestion in the network, and the MTN data logging system registers every call attempts made. The study found that peak period of the day is not during the working hours but between 7.00pm to 8.00pm, this is the period when the network experiences heavy traffic.

Furthermore, one of the basic causes of congestion is that GSM phones are now used for data transfer, Internet browsing and other multimedia applications. Such activities, affect the network's available capacity, this can lead to congestion. This was supported by (Adugu, 2005) who stated that some wireless networks allows GSM phone for multimedia applications and new services such as file transfer and Internet browsing, such applications are non-interactive and as such are less sensitive to the assigned throughput. Landstrom et al, (2005) in a related study also observed that 1G was developed purely for speech or voice communication alone; they noted that, there was less traffic in GSM network during the era of the 1G. During the advent of 2G, the use of GSM became more sophisticated because GSM was not only used for voice communication but for data transfer and internet connectivity. This led to heavy traffic and hence collision, which is as a result of congestion. These activities place heavy demand on the system capacity of the GSM network.

Also, the system capacity for Yola attempted calls is 45,263, as against what is obtainable 43,125 which is the established calls and congested calls is 2,138. For Port-Harcourt average calls attempted, established and congested are 705,000, 685,650 and 16,250 respectively. For Abuja we have 663,219, 656,125 and 7,094 respectively, Kano the values are 405,625, 390,313 and 15,312. Finally average number of attempted, established and congested calls in Enugu is 478,438, 450,438 and 28,000. The pattern in the behaviour of the data is periodic and repetitive. Figures 4.2, 4.3, 4.4, 4.5, 4.6 and 4.7 showed a sharp drop in the number of calls at a particular time every week day, this usually occurs between the 16hrs and 17hrs (4.00pm – 5.00pm) of the day, this is probably when people are closing from work.

At this time of the day, there is less traffic in the MTN network. Findings also show that calls reduce as it tends towards weekend, especially on Sundays. This is because, people do not go to work during the week end hence, there is less interaction. There is also a sharp drop in the number of calls on Sundays this drop is observed at about 12.00pm, this can be attributed to the fact that most people are in church and most phones will likely be switched off.

7. CONCLUSIONS

There are problems of congestion in Nigeria telecommunication industry with respect to GSM, such as difficulties in connecting subscribers and losing of resources by the service providers. It was observed from the findings that the peak period in the MTN network is usually between 9.00pm to 10.30pm of the day. This is surely an issue because it implies that Nigerians do not use GSM majorly for business but for pleasure and social activities. From the results it is apparent that Nigerians make more calls after work period, when they are relaxed at home after the day's job. The issue of priority calls is also very critical to the Nigeria factor. The concept of priority call or emergency call does exist in Nigeria. But there is no policy on ground to support the implementation of priority call. As a matter of fact every GSM service provider should make provision for priority lines for life critical security issues.

The GSM specialization does recognize emergency line. For instance, 112 is an emergency number that is recognized internationally. In developed countries, priority calls are accorded a dedicated line; this can be seen as part of development. Therefore there is need for Nigerian Communications Commission and every stakeholder involved to resolve the issue of priority call, perhaps the government can enact or make policy that will enforce every GSM subscriber in Nigeria to implement priority calls.

The other side of the coin is that, when priority calls are been implement though it is a good idea but it will cause traffic in the network during peak periods because some of the lines must have been prioritized hence, a whole lots of callers will be left with no choice but to contend for the remaining available line, hence giving priority to emergency calls has congestion factor. MTN being a business venture like any other profit making organization should make every effort to combat and minimize congestion to its beeriest minimum. This will help the organization minimize all available resources in order to maximize profit. Further study could be conducted to determine the level of GSM for social-economic activities and how it impacts on congestion

REFERENCES

[1] Adugu .E (2004) A Combined rate and error control mechanism in cellular network. (Accessed Internet on 17th August 2006)
Ajala I., (2005). Article on GIS and GSM Network Quality Monitoring: A Nigerian Case Study

[2] Anderson, P. (2003), Management Information System. GSM Association (2005). GSM World – 3GSM Platform. http://www.gsmworld.com/technology/3g/index.shtml

[3] Hutchinson E. and Sawyer C, (2000) Computer Communication and Information (McGraw-Hill Press) Seventh Edition.

[4] Harte L, Levine R, Livingston G (1999). "GSM Superphones", McGraw Hill 71: 45-47.

[5]Kelly F., (1991) Loss networks. The Annals of Applied probability 1(3): 319-370.

[6] Kuboye, B. (2010) Optimization models for minimizing congestion in Global System for Mobile Communications (GSM) in Nigeria. Journal Media and Communication Studies Vol. 2(5), pp. 122-126, May 2010 Available online http://www.academicjournals.org/jmcs ISSN 2141 – 2545 © 2010 Academic Journals

[7] Kühn, P.J, Pack, C.D. and Skoog, R. (1994), "Common Channel Signaling Networks: Past, Present, Future", IEEE Journal on Selected Areas in Communications, Vol. 12, No. 3, pp. 383-394,

[8] Landstorm S., Lars-Ake L., and Bodin U., (2004) Property of TCP-like congestion control. In proceedings of the Swedish National Computer Networking Workshop, Pp 13-18, Karstad, Sweden. Published and printed by LTU Press. http://www.csee.ltu.se/~saral.

[9] Landstorm S., Lars-Ake L., and Bodin U., (2004) Congestion control in a high-speed radio environment. In proceedings of the International Conference on Wireless Networks, Pp 617-623, Las Vegas, Nevada, USA, 21-24 June. Published and printed by University press. http://www.csee.ltu.se/~saral

[10] Lars A., Stefan P, and Ake A., (2001.) A network approach to signaling network congestion control Dept. of Telecommunications and Mathematics, Hogskolan Karlskrona/Ronneby, Blekinge Institute of Technology. http://www.Bth.com. And http://www.bth.se/fou/forskinfo.nst

[11] Mehrotra .A (1997). "GSM System Engineering", Artech house, Inc., pp: 70-73
MTN Customer Confidential Report (2005) Components of GSM Architecture

[12] Ndukwe E. (2008), ICT Statistics Newslog-Nigeria dreams of 100% teledensity by 2020. News related to ITU Telecommunication/ICT Statistics. http://www.itu.int/ITU-D/ict/newslog/Nigeria+Dreams+Of+100+Teledensity+By+2020.aspx

[13] Nhleko.P. (2010), MTN gets over 4m new Nigerian subscribers in first half of 2010. August 19, 2010 http://234next.com/csp/cms/sites/Next/Home/5608924-146/mtn_gets_over_4m_new_nigerian.csp

[14] Peter BE (2002). "GSM Raw Capacity Solution" Ericsson Review (2). Rajkummar F (1996). Performance issues of cellular Networks Http: //www.eic uk/surprise_96/ Journal

[15] Piasecki. D. (2002), Inventory Operations Consulting L.L.C http://www.inventoryops.com/ADC.htm

[16] Stefan P., and Ake A., (2002). Economical Aspects of a congestion control mechanism in signaling network, 2002. Dept. of Telecommunications and Mathematics, Hogskolan Karlskrona/Ronneby, Blekinge Institute of Technology. http://www.Bth.com.

[17] Syski R (1986). "Introduction to Congestion Theory in Telephone Systems" Elserier Science Publishers B. V.

[18] Stojanovic I. Airy M, Gesbert D. and Saran H. (2001) Performance of TCP/IP Over Next Generation Broadband Wireless links Supporting TCP Traffic. In IEEE VTC, pages 675-679,

[19] Theodore R., (2001) Wireless communications: principles and practice. Pp 141-145 Prentice Hall, 2nd edition, December.

[20] Verhonic and Seaman (1978) in Aina. L, (2002), Research in Information Science Page 18, Stirling-Horden Pu

A Data Mining Model for Predicting Computer Programming Proficiency of Computer Science Undergraduate Students

[1]**O.S. Akinola**
[2]**B.O. Akinkunmi**
[3]**T.S. Alo**
Department of Computer Science
University of Ibadan
Ibadan, Nigeria
[1]solom202@yahoo.co.uk [2]bo.akinkunmi@mail.ui.edu.ng [3]tosyno2001@yahoo.com

ABSTRACT
Providing quality education to its students is the main objective of higher education institutions. One way to achieve highest level of quality in higher education system is by discovering knowledge for prediction regarding enrolment of students in a particular course, alienation of traditional classroom teaching model, detection of unfair means used in online examination, detection of abnormal values in the result sheets of the students, prediction about students' performance and so on. The knowledge is hidden among the educational data set and it is extractable through data mining techniques. Prequalification ordinary level results of Computer Science students as well as their results obtained in the first year 100 level Physics, Mathematics and a programming course done at 200 Level were collected and suggested to data mining tasks. This is done in order to predict the performance of students in Computer programming. Result from the study shows that a priori knowledge of Physics and Mathematics are essential in order for a student to excel in Computer Programming. This work will be of considerable usefulness in identifying students at risk early, especially in very large classes, and allow the instructor to provide appropriate advising in a timely manner.

Key words: Data Mining, ANN, Computer Programming, Students' Performance, Undergraduates.

1. INTRODUCTION

One of the biggest challenges faced by both tutors and students of computer science today is whether it is compulsory for all students of the discipline to master computer programming and be a 'guru' in it or not. Computer programming is the art and science of writing instructions for the computer hardware to perform (Akinola, 2011).

African Journal of Computing & ICT Reference Format:

O.S. Akinola, B.O. Akinkunmi & T.S. Alo (2012): A Data Mining Model for Predicting Computer Programming Proficiency of Computer Science Undergraduate Students. Afr J. of Comp & ICTs. Vol 5, No.1 pp 43- 52

- ISSN 2006-1781

Usually, computer students undergo training in computer programming right from their first year up to probably third year of their degree or diploma in higher institutions. Several languages both procedural and object-oriented are learnt by them in this period. At the end of all these trainings, all computer science students are deemed to have mastered the art of developing computer programs. Critical observations show that this assertion is not usually the case.

No doubt, computer programming is a difficult and challenging subject area which places a heavy cognitive load on programmers. The big question is what factors are responsible for the disparity in proficiency of students in computer programming despite all the efforts and machinery put in place by their tutors?

One way to effectively address these students' programming proficiency disparity challenge is through the analysis and presentation of data, or data mining. Data mining enables organizations and institutions to use their current reporting capabilities to uncover and understand hidden patterns in vast databases [2]. These patterns are then built into data mining models and used to predict individual behaviour with high accuracy. As a result of this insight, tutors of computer programming at schools are able to understand the background of their students and device ways of teaching them more effectively. Undisputedly, several factors: social, family background, personal interest, etc., may affect a student's performance in his/her educational career; our motivation for this study deviates away from all these non-quantifiable factors. Our study on the other hand explores the effect of the students' a priori knowledge of the basic qualifying subjects on their proficiency in computer programming.

Usually, students of computer science at the University of Ibadan, Nigeria are admitted if they have O/L credit passes in English language, Mathematics, Physics, Chemistry and one other science subjects in addition to the University Matriculation Examination (UME) conducted by the Joint Admissions and Matriculations Board (JAMB) in Nigeria. They are also requested to pass at least three Physics, two mathematics and one Statistics courses besides the Introduction to Computer Science course at their 100 level (first year) before they would be allowed to continue their studies in Computer Science. In this study, we applied Artificial Neural Network (ANN), a data mining tool on a five-session data sets collected from computer science students' basic qualifying Ordinary Level (O/L) subjects, scores obtained in their first year basic courses (Physics and Mathematics) as well as their scores in a computer programming course (CSC 232 – Structured programming with Java). Neuro Shell Classfier was employed as the ANN data mining tool in the study.

In the rest of this paper, we present related works in section 2 of this work. The methodology used in this study is presented in section 3, while results and discussion of results are presented in section 4. Section 5 concludes the paper.

2. RELATED WORKS

The quest for patterns in data has been studied for a long time in many fields, including statistics, patterns recognition and exploratory data analysis [4, 5]. Analyzing data can provide further knowledge about a business by going beyond the data explicitly stored to derive knowledge about the business. This is where data mining has obvious benefits for any enterprise.

Data mining, also called Knowledge Discovery in Databases (KDD), is the field of discovering novel and potentially useful information from large amounts of data. Data mining has been applied in a great number of fields, including retail sales, bioinformatics, and counter-terrorism. In recent years, there has been increasing interest in the use of data mining to investigate scientific questions within educational research, an area of inquiry termed educational data mining. Brijesh and Saurabh [6] summarize the stages involved in data mining as in Figure 1.

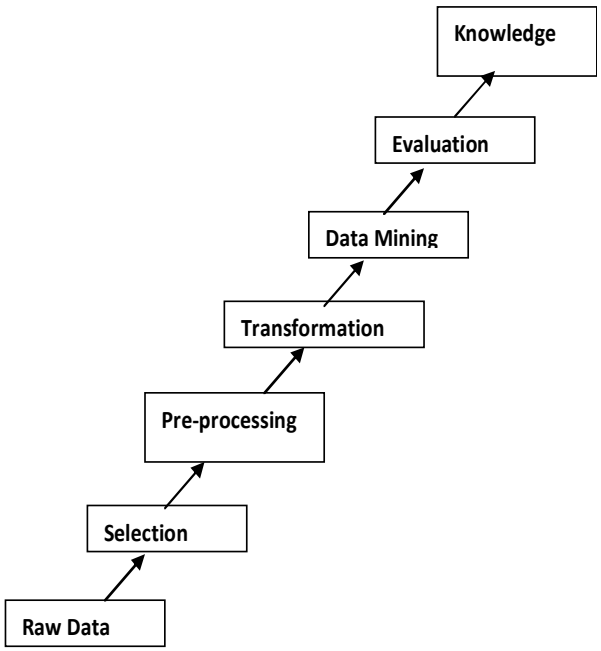


Figure 1: The steps of Extracting Knowledge From Data

Various algorithms and techniques like Classification, Clustering, Regression, Artificial Intelligence, Neural Networks, Association Rules, Decision Trees, Genetic Algorithm, Nearest Neighbour method etc., are used for knowledge discovery from databases. To describe a few of them, according to Brijesh and Saurabh [6], Classification is the most commonly applied data mining technique, which employs a set of pre-classified examples to develop a model that can classify the population of records at large. This approach frequently employs decision tree or neural network-based classification algorithms. The data classification process involves learning and classification. In Learning the training data are analyzed by classification algorithm. In classification test data are used to estimate the accuracy of the classification rules. If the accuracy is acceptable the rules can be applied to the new data tuples.

Neural network is a set of connected input/output units and each connection has a weight present with it. During the learning phase, network learns by adjusting weights so as to be able to predict the correct class labels of the input tuples. Neural Networks (NN) are a class of systems modeled after the human brain. As the human brain consists of millions of neurons that are interconnected by synapses, neural networks are formed from large numbers of simulated neurons, connected to each other in a manner similar to brain neurons. Like in the human brain, the strength of neuron interconnections may change (or be changed by the learning algorithm) in response to a presented stimulus or an obtained output, which enables the network to “learn”.

Neural networks have seen an explosion of interest over the last few years, and are being successfully applied across an extraordinary range of problem domains, in areas as diverse as finance, medicine, engineering, geology and physics. Indeed, anywhere that there are problems of prediction, classification or control, neural networks are being introduced [3]. Neural networks have the remarkable ability to derive meaning from complicated or imprecise data and can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques.

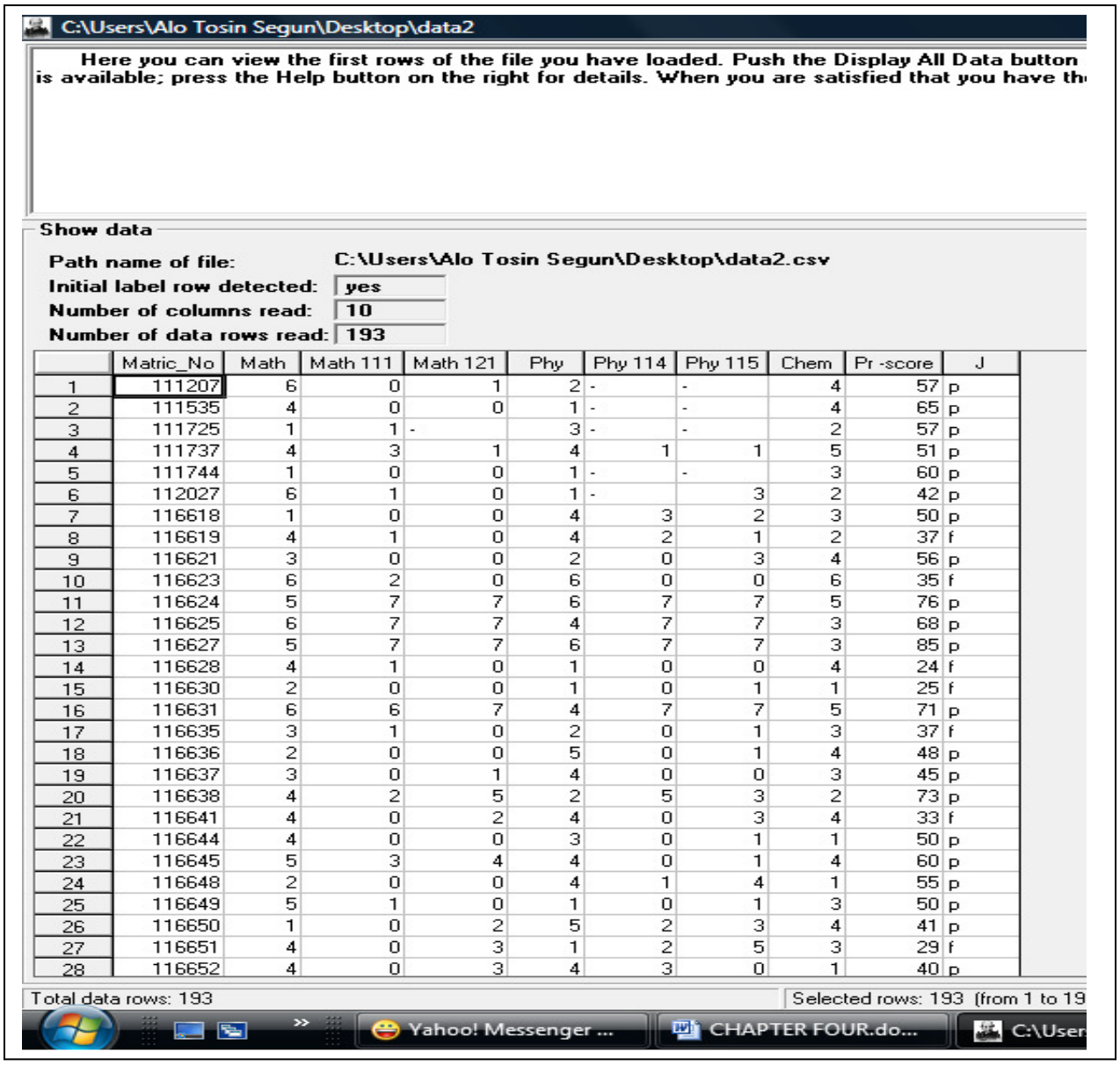
These are well suited for continuous valued inputs and outputs. Neural networks are best at identifying patterns or trends in data and well suited for prediction or forecasting needs[6]. The ultimate goal of data mining is prediction [3] - and predictive data mining is the most common type of data mining and one that has the most direct business applications.

Researchers are on to determining the various factors affecting students’ performance using data mining techniques. For instance, Brijesh and Saurabh [6] apply decision tree classification algorithm to extract knowledge that describes students’ performance in end semester examination. The result helps in identifying the dropouts and students who need special attention and allow the teacher to provide appropriate advising/counselling.

Paulo and Alice [7] approach student achievement in secondary education using data mining techniques. Real-world data (e.g. student grades, demographic, social and school related features) was collected by using school reports and questionnaires. The two core classes (i.e. Mathematics and Portuguese) were modeled under binary/five-level classification and regression tasks. Also, four DM models (i.e. Decision Trees, Random Forest, Neural Networks and Support Vector Machines) and three input selections (e.g. with and without previous grades) were tested. The results show that a good predictive accuracy can be achieved, provided that the first and/or second school period grades are available. Although student achievement is highly influenced by past evaluations, an explanatory analysis has shown that there are also other relevant features (e.g. number of absences, parent’s job and education, alcohol consumption).

Romero, Ventura and Garcia [1, 8] described how different data mining techniques can be used in order to improve the course and the students’ learning. Tissera, Athauda and Fernando [1, 9] also described the use of data mining techniques to predict the strongly related subject in courses’ curricula. The information provided in their works can further be used to improve the syllabi of any course in any educational institute.

Bhardwaj and Pal [6, 10] conducted study on the student performance based by selecting 300 students from 5 different degree college conducting BCA (Bachelor of Computer Application) course of Dr. R. M. L. Awadh University, Faizabad, India. By means of Bayesian classification method on 17 attribute, it was found that the factors like students‟ grade in senior secondary exam, living location, medium of teaching, mother‟s qualification, students other habit, family annual income and student‟s family status were highly correlated with the student academic performance. Al-Radaideh, et al [11] applied a decision tree model to predict the final grade of students who studied the C++ course in Yarmouk University, Jordan in the year 2005. Three different classification methods namely ID3, C4.5, and the Naïve Bayes were used. The outcome of their results indicated that Decision Tree model had better prediction than other models.


C:\Users\Alo Tosin Segun\Desktop\data2

Here you can view the first rows of the file you have loaded. Push the Display All Data button is available; press the Help button on the right for details. When you are satisfied that you have th

Show data

Path name of file: C:\Users\Alo Tosin Segun\Desktop\data2.csv
Initial label row detected: yes
Number of columns read: 10
Number of data rows read: 193

	Matric_No	Math	Math 111	Math 121	Phy	Phy 114	Phy 115	Chem	Pr -score	J
1	111207	6	0	1	2	-	-	4	57	p
2	111535	4	0	0	1	-	-	4	65	p
3	111725	1	1	-	3	-	-	2	57	p
4	111737	4	3	1	4	1	1	5	51	p
5	111744	1	0	0	1	-	-	3	60	p
6	112027	6	1	0	1	-	3	2	42	p
7	116618	1	0	0	4	3	2	3	50	p
8	116619	4	1	0	4	2	1	2	37	f
9	116621	3	0	0	2	0	3	4	56	p
10	116623	6	2	0	6	0	0	6	35	f
11	116624	5	7	7	6	7	7	5	76	p
12	116625	6	7	7	4	7	7	3	68	p
13	116627	5	7	7	6	7	7	3	85	p
14	116628	4	1	0	1	0	0	4	24	f
15	116630	2	0	0	1	0	1	1	25	f
16	116631	6	6	7	4	7	7	5	71	p
17	116635	3	1	0	2	0	1	3	37	f
18	116636	2	0	0	5	0	1	4	48	p
19	116637	3	0	1	4	0	0	3	45	p
20	116638	4	2	5	2	5	3	2	73	p
21	116641	4	0	2	4	0	3	4	33	f
22	116644	4	0	0	3	0	1	1	50	p
23	116645	5	3	4	4	0	1	4	60	p
24	116648	2	0	0	4	1	4	1	55	p
25	116649	5	1	0	1	0	1	3	50	p
26	116650	1	0	2	5	2	3	4	41	p
27	116651	4	0	3	1	2	5	3	29	f
28	116652	4	0	3	4	3	0	1	40	p

Total data rows: 193 Selected rows: 193 (from 1 to 19

Yahoo! Messenger ... CHAPTER FOUR.do... C:\User

Figure 2. Sample Data used

Al-Radaideh, et al [11] applied a decision tree model to predict the final grade of students who studied the C++ course in Yarmouk University, Jordan in the year 2005. Three different classification methods namely ID3, C4.5, and the Naïve Bayes were used. The outcome of their results indicated that Decision Tree model had better prediction than other models.

Table 1: Variables

Variable	Description	Possible Coding values
Matric No.	The matriculation numbers of the students.	Not used in the data mining process.
Math	Ordinary Level result obtained in Mathematics by the students	6 = A1 (Distinction) 5 = B2 4 = B3 3 = C4 2 = C5 1 = C6 (Average)
Math 111	Grade Point Result obtained in MAT 111 (Algebra) at first year (100 Level) by the students	7 = 70 – 100% 6 = 65 – 69% 5 = 60 – 64% 4 = 55 – 59% 3 = 50 – 54% 2 = 45 – 49% 1 = 40 – 44% 0 = 0 – 39%
Math 121	Grade Point Result obtained in MAT 121 (Calculus and Trigonometry) at first year (100 Level) by the students	7 = 70 – 100% 6 = 65 – 69% 5 = 60 – 64% 4 = 55 – 59% 3 = 50 – 54% 2 = 45 – 49% 1 = 40 – 44% 0 = 0 – 39%
Phy	Ordinary Level result obtained in Physics by the students	6 = A1 (Distinction) 5 = B2 4 = B3 3 = C4 2 = C5 1 = C6 (Average)
Phy 114	Grade Point Result obtained in PHY 114 (Mechanics & Properties of Matter) at first year (100 Level) by the students	7 = 70 – 100% 6 = 65 – 69% 5 = 60 – 64% 4 = 55 – 59% 3 = 50 – 54% 2 = 45 – 49% 1 = 40 – 44% 0 = 0 – 39%
Phy 115	Grade Point Result obtained in PHY 115 (Heat and Thermodynamics) at first year (100 Level) by the students	7 = 70 – 100% 6 = 65 – 69% 5 = 60 – 64% 4 = 55 – 59% 3 = 50 – 54% 2 = 45 – 49% 1 = 40 – 44% 0 = 0 – 39%
Chem	Ordinary Level result obtained in Chemistry by the students	6 = A1 (Distinction) 5 = B2 4 = B3 3 = C4 2 = C5 1 = C6 (Average)
Pr Score	Score obtained in CSC 232 (Structured Programming with Java) at 200 Level	Absolute values of the scores used
J	The predictor variable	P = passed CSC 232 F = Failed CSC 232

Ghaleb and Qeethara [13] predict the factors affecting the University of Jordan students' performance using Artificial Neural Networks (ANN) model. Various factors that may likely influence the performance of a student were identified. Results from their study showed that secondary school performance which is measured by scores in secondary school certificate examination, measured in a percentage form having the largest regression value.

3. Methodology / Data Mining Process

3.1 Data Preparations

The data set used in this study was obtained from records of Computer Science students kept in the Department of Computer Science, University of Ibadan, Nigeria. 200 data sets were collected from 2003/04/ 2004/05, 2005/06, 2007/08 and 2008/09 sessions in the department. 2006/07 was cancelled by the University authority due to incessant strike present in the session. Ordinary Level basic entry qualifications of the students, their scores in two Physics and Mathematics each as well as their scores in Structured Programming Course (CSC232) at their 200 Levels were collected for the data mining study. Figure 2 shows the structure of the data collected and given to Neuro Shell Classifier for analysis.

3.2 Data selection and transformation

In this step only those records and fields were selected which were required for data mining. A few derived variables were selected. All the predictor and response variables which were derived for the data mining activity are presented in Table 1

3.3 The ANN Back Propagation Algorithm:

A Multi-Layer Perceptron Feed-Forward Back Propagation Neural Network was employed in this work. This type of ANN was chosen because of its ease of use and capabilities for supervised learning. Multi-layer means that the network has three layers: input, hidden and output layers. The term, "feed foreword" describes how this neural network processes the pattern and recalls patterns. When using a "feed forward neural network" neurons are only connected foreword.

Each layer of the neural network contains connections to the next layer (for example from the input to the hidden layer), but there are no connections back. The term back propagation describes how this type of neural network is trained. Back propagation is a form of supervised training [15]. Backpropagation, or propagation of error, is a common method of teaching artificial neural networks how to perform a given task. The back propagation algorithm is used in layered feedforward ANNs.

This means that the artificial neurons are organized in layers, and send their signals “forward”, and then the errors are propagated backwards. The back propagation algorithm uses supervised learning, which means that we provide the algorithm with examples of the inputs and outputs we want the network to compute, and then the error (difference between actual and expected results) is calculated. The idea of the back propagation algorithm is to reduce this error, until the ANN *learns* the training data. Yashpal and Alok [14] summarized the ANN technique thus:

1. Present a training sample to the neural network.
2. Compare the network's output to the desired output from that sample. Calculate the error in each output neuron.
3. For each neuron, calculate what the output should have been, and a *scaling factor*, how much lower or higher the output must be adjusted to match the desired output. This is the local error.
4. Adjust the weights of each neuron to lower the local error.

The Actual algorithm can be found in [14, 15 and 16]. The Neuro Shell Classifier by Ward System Group, Executive Park West, Maryland Inc. US, was finally used for the prediction data mining process.

3.4 Data Sets

The data were divided into three sets: Training, Verification and Test, in the ratio 5: 3: 2 respectively as suggested in the literatures [15, 16].

The training data set is used to update the neural weight parameter during learning while the validation data set is used to crosscheck or monitor the quality of the neural network model during training while the test data is used to examine the generalization capability and quality of the developed model using some performance measures [16]. The training of the network started with all the inputs and one hidden layer in order to have a good network topology and to avoid the problems of over-learning, under-learning or local minimum [16]. The number of hidden neurons was progressively adjusted, while network layers and weights were randomly generated by the tool.

4. RESULTS

4.1 Relative Importance of Inputs

Table 2 gives the result obtained from the network pertaining to the relative importance of the input variables to the network. The table shows that the students' scores in PHY 114 (Mechanics and Properties of Matter) contributes most significantly to the prediction (29.6%) followed by the Ordinary Level Physics results of the Students (27.1%) while MAT 121 (Calculus and Trigonometry, 23.1%) takes the next lead. Figure 2 shows the chart produced by the tool to illustrate further these contributions from the inputs.

Table 2: Importance of Inputs

Subjects		% Contribution
Phy 114	0.296	29.6
Phy (OL)	0.271	27.1
Math 121	0.231	23.1
Math 111	0.095	9.5
Math (OL)	0.055	5.5
Phy 115	0.029	2.9
Chem. (OL)	0.023	2.3
TOTAL	100.00	

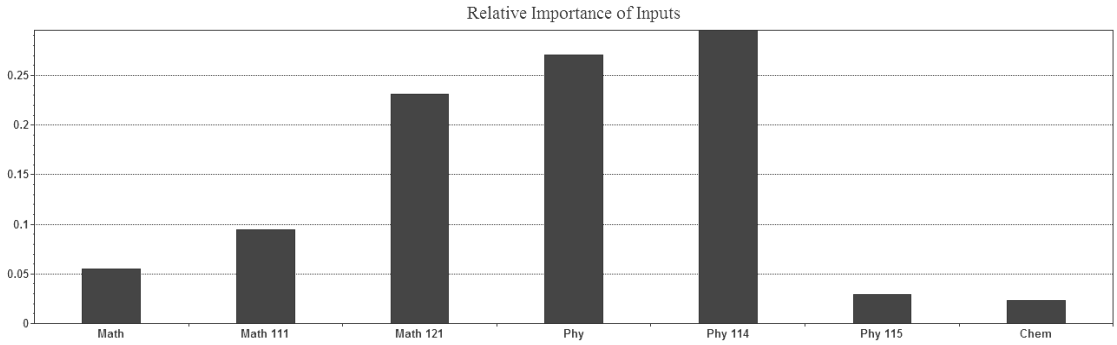


Figure 2: Relative Importance of inputs

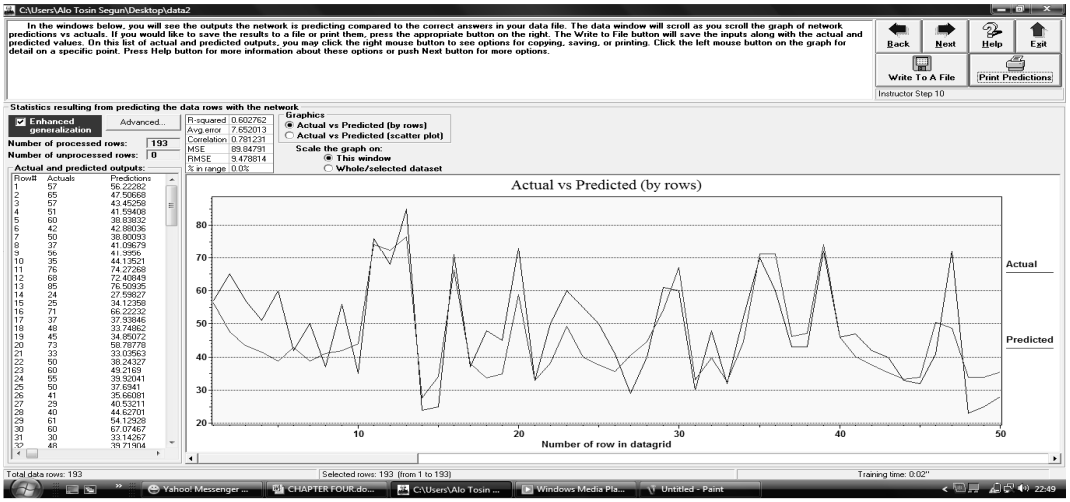


Figure 3: Actual versus Predicted Values

4.2 Best Net Statistics (BNS)

The best net statistics shows the measure of performance of the network with regard to the input variables presented to it. The BNS values obtained from the network is presented in Table 3.

Table 3: Best net Statistics

R-squared	0.602762
Avg.error	7.652013
Correlation	0.781231
MSE	89.84791
RMSE	9.478814

The Correlation between the actual versus predicted output values is very good (0.781) meaning that the correlation is very positively strong. And with the coefficient of determination R-squared value of 0.602 (60%), the correlation is somehow acceptable. The Mean Squared Error (MSE) is used to determine how well the network output fits the desired output [16]. RMSE means Root Mean Squared Error.

4.2 Discussion of Results

Data Mining can be used in educational field to enhance our understanding of learning process to focus on identifying, extracting and evaluating variables related to the learning process of students.

The fact that many Universities and even Polytechnic demand that a prospective Computer Science candidate should have a sound science background calls for inquiry on the effect of the science subjects on their performance in their course of study. Physics, Chemistry and Mathematics are perhaps the basic science subjects that must be passed at credit level by the prospective candidates for Computer Science. Of course, everybody will agree to the fact that Physics and Mathematics are brain demanding subjects. Computer programming might also be regarded as a brain tasking area of Computer Science.

At the University of Ibadan, Nigeria, prospective candidates are also expected to pass and even offer Chemistry at the University Matriculation Examination conducted by JAMB. And those who mistakenly do other subjects like Economics or Biology *in lieu* of Chemistry at the UME are denied being considered for admission.

Our motivation for this study stems out from these premises. We pose to answer the question on whether Chemistry is actually needed to be passed by prospective Computer Science candidates and secondly, which of the basic science subjects will enhance the programming skill of the candidates. Programming is very important for Computer Science students in that it forms the basis of their course of study - Computer Science. More than half of the courses the students will pass through are programming based. Take for example, data structures, databases, information systems, operating systems, algorithms analyses, etc. need a sound programming background to teach and learn them.

This study adopts a data mining approach to answer the questions by subjecting data to an Artificial Neural Network (ANN) data mining technique. Results from this study show that PHY 114 (Mechanics) input, contributes mostly to the performance of students in computer programming followed by their Ordinary Level Physics (PHY) and MAT 121 in that order. PHY 114 (Mechanics) and MAT 121 (Calculus and Trigonometry), are two of the courses that must be offered and possibly passed by Computer Science students in their first year of their study in the University of Ibadan. Students of Computer Science at the University of Ibadan do not offer Chemistry at all throughout their course of study. Results from the study also show that Ordinary Level Chemistry demanded from the candidates contributes least to their efficiency in programming.

This result implies that candidates with good background in Physics and Mathematics (especially Further Mathematics) will perform efficiently in Computer Programming and would eventually be good programmers beyond the school. These two subjects are calculation-intensive and they demand sharp brains that can think fast and precisely.

Computer programming involves developing efficient algorithms and being able to turn these algorithms to efficient working programs. In some cases, these algorithms are mathematically based. Results obtained in this study therefore justify this fact, although, many other variables like social, interest, etc. are also there to determine the proficiency of somebody in a vocation.

Our result is in line with existing works that embrace the fact that pre-higher institution qualifications would contribute immensely to the performance of students in their chosen course of studies. For instance, the work of Bhardwaj and Pal [10] shows that students' grades in senior secondary examination are one of the factors that contributed to the academic performance of 300 candidates used in their study at India. Kabakchieva *et al* [12] and Ghaleb and Qeethara [13] also work on data mining task to predict the student university performance based on the student personal and pre-university characteristics. Other related works come from Brijesh and Saurabh [6] and Al-Radaideh, *et al.,* [11].

5. Conclusion

This study employed the use of Artificial Neural Network data mining tool to predict the performance of students in Computer Programming. The study reveals that background knowledge of mathematics and Physics is very much essential to becoming a good programmer at school and beyond.

Result from this study will help the students to get fully prepared for the programming course especially if they are deficient in calculation intensive pre-qualification subjects. This study will also help programming tutors to identify those students that will need special attention to reduce fail rate and taking appropriate steps to imparting programming course to students.

References

1. Varun Kumar and Anupama Chadha (2011). An Empirical Study of the Applications of Data Mining Techniques in Higher Education, (IJACSA) International Journal of Advanced Computer Science and Applications,Vol. 2, No.3, pp. 80 -84.
2. Jing Luan (2006). Data Mining Applications in Higher Education, SPSS, www.spss.com/ Downloaded in October 2011.
3. Statisoft.com. Data Mining Techniques, www.statisoft.com, downloaded in October, 2011.
4. Usama Fayyad and Ramasamy Uthurusamy (2002), Evolving Data Mining into Solutions for Insights, Communications of the ACM, vol. 45, No. 8, pp. 28 – 31.
5. Tukey, J. (1977).Exploratory Data Analysis. Addison-Wesley, Reading, MA.
6. Brijesh Kumar Baradwaj and Saurabh Pal (2011). Mining Educational Data to Analyze Students' Performance, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, pp. 63 – 69.
7. Paulo Cortez and Alice Silva (2011). Using Data Mining to Predict Secondary School Student Performance, 15th Portuguese Conference on Artificial Intelligence, EPIA 2011, Lisbon, Portugal, October 10-13, 2011, pp. 491 – 505, http://www3.dsi.uminho.pt/pcortez/student.pdf, visited in Oct. 2011.
8. Romero, C., Ventura, S. and Garcia, E. (2008). "Data mining in course management systems: Moodle case study and tutorial", Computers & Education, Vol. 51, No. 1, pp. 368-384.
9. Tissera, W.M.R., Athauda, R.I.,and Fernando, H. C. (2006). "Discovery of Strongly Related Subjects in the Undergraduate Syllabi using Data Mining", IEEE International Conference on Information Acquisition.
10. Bharadwaj, B.K. and Pal, S. (2011). "Data Mining: A prediction for performance improvement using classification", International Journal of Computer Science and Information Security (IJCSIS), Vol. 9, No. 4, pp. 136-140.
11. Al-Radaideh, Q. A., Al-Shawakfa, E. W. and Al-Najjar, M. I. (2006) "Mining student data using decision trees", International Arab Conference on Information Technology (ACIT'2006), Yarmouk University, Jordan.
12. Kabakchieva, D., Stefanova, K. and Kisimov, V. (2011). Analyzing University Data for Determining Student Profiles and Predicting Performance, http://educationaldatamining.org/EDM2011/wp-content/uploads/proc/edm2011_poster15_Kabakchieva.pdf, visited in October 2011.

13. Ghaleb A. El-Refae and Qeethara Kadhim Al-Shayea (2010). Predicting Students' Academic Performance Using Artificial Neural Networks: A Case Study, International Journal of Computer Science and Information Security, Vol. 8 No. 5 pp. 97 – 100.

14. Yashpal Singh & Alok S. Chauhan (2009) Neural Networks In Data Mining. Journal of Theoretical & Applied Information Technology. 2005-2009, pp 37-42.
15. Jeff Heaton (2008). Introduction to Neural Networks with Java, 2nd Edition, ISBN-13: 9781604390087, http://www.heatonresearch.com/articles/5/page6.html,
16. Osofisan, A. O., Akomolafe, O. P and Akinola, S. O. (2005). Discovering Knowledge in Road Accident Database, Publications of the ICMCS, Vol. 2, pp. 31 – 43.
17. Yashpal Singh and Alok Singh Chauhan (2009), Neural Networks In Data Mining, Journal of Theoretical and Applied Information Technology, 2005 – 2009, pp. 37 – 42.

Olalekan S. Akinola is currently a lecturer of Computer Science at the University of Ibadan, Nigeria. He had his PhD Degree in Software Engineering from the same University in Nigeria. His research interests include Data Mining and Software Engineering.

Babatunde Opeoluwa Akinkunmi is a member of the academic staff at the Dept of Computer Science University of Ibadan. He has authored over twenty five research articles in computer science. His research interests include Knowledge Representation, Formal Ontologies and Software Engineering.

Alo Tosin had BSc and MSc degrees in Computer Science from the University of Ibadan, Nigeria. He is currently engaged as a banker in Nigeria.

Comparative Analysis of Similarity Check Mechanism for Motif Extraction

A. Makolo & A.O. Osofisan
Department of Computer Science
University of Ibadan
Ibadan, Nigeria
amakolo@ui.edu.ng, aosofisan@ui.edu.ng

E. Adebiyi
Department of Computer and Information Sciences
Covenant University
Ota, Nigeria
ezekiel.adebiyi@covenantuniversity.edu

ABSTRACT
In this work, a comparative analysis of the similarity check mechanism used in the most effective algorithm for mining simple motifs GEMS (Gene Enrichment Motif Searching) and that used in a popular multi-objective genetic algorithm, MOGAMOD (Multi-Objective Genetic Algorithm for Motif Discovery) was done. In our previous work, we had reported the implementation of GEMS on suffix tree –Suffix Tree Gene Enrichment Motif Searching (STGEMS) and shown the linear asymptotic runtime achieved. Here, we attempt to empirically proof the high sensitivity of the resulting algorithm, STGEMS in mining motifs from challenging sequences like we have in *Plasmodium falciparum*. The results obtained validates the high sensitivity of the similarity check mechanism employed in GEMS and also shows that a careful deployment of this mechanism in the multi-objective genetic algorithm, improved the sensitivity level of the resulting algorithm. The end results gave us room to exhaustively mine structured motifs.

Keywords: Motifs, GEMS, MOGAMOD, STGEMS, Suffix tree.

1. INTRODUCTION

Sequence motif discovery is one of the fundamental concerns in Bioinformatics and it has important applications in locating regulatory sites and drug target identification. The extraction of structured motifs (i.e. several words with well defined gaps) is particularly interesting because of its application to the detection of binding sites.

African Journal of Computing & ICT Reference Format:

A. Makolo, A.O. Osifisan & E. Adebiyi (2012) - Comparative Analysis of Similarity Check Mechanism for Motif Extraction. Afr J. of Comp & ICTs. Vol 5, No.1 pp 53-58

- ISSN 2006-1781

This binding sites respect distance constraints. In this paper we consider the extraction of structured motifs from the deadly organism, the malaria parasite, *P. falciparum*. The highly repetitive and specific alphabet (in this case AT) bias sequences of *P. falciparum* makes the task of extracting structured motifs a very challenging one. The GEMS algorithm by [8] has been shown to successfully mine simple motifs in the *P. falciparum* kind of sequences. This success is attributed mainly to the unique approach employed in GEMS similarity check mechanism. It involves the use of hypergeometric-based scoring function to compute the p-value of the candidate motifs, ranking them according to this value and computing the position weight matrix as a neighborhood in the sequence space.

Long distance metric can also be used to which considered positional information to merge non-

unique motif candidates. Position-Specific Scoring Matrices (PSSMs) and their derivatives (i.e. position frequency matrix, position weight matrix) have become the standard representation of a transcription factor's DNA-binding preference. For example, experimentally derived DNA-binding preferences for a growing number of transcription factors are stored as frequency matrices in databases such as JASPAR [6] and TRANSFAC [5]. In addition, most de novo motif-finding software tools report statistically over-represented degenerate sequence features in the form of frequency matrices or consensus sequences. [7]. PSSMs have been used in this work as a representative model for the extracted candidate motifs.

Similarity check is a measure of the degree of closeness of two strings. This is useful in computational biology where two slightly different patterns can represent the same motif due to the presence of a number of mismatches allowed. For instance motifs AAAATGC, AACATGC, AAATTGC are similar motifs with one mismatch.

MOGAMOD algorithm by [3] used a well-known high-performance multi-objective genetic algorithm called NSGA II.[1] to find a large number of tradeoff motifs with respect to conflicting objectives of similarity, motif length and support maximization. MOGAMOD's similarity check involves performing a measure of similarity among all motif instances defining a candidate motif. This is achieved by first generating a position weight matrix from the motif patterns found in every sequence. Then, the dominance value of the dominant nucleotide is computed from the position weight matrix which forms the basis of the similarity function.

In [4], we introduced the STGEMS algorithm which implemented the GEMS algorithm on the suffix tree and also demonstrated the improved run time of STGEMS over GEMS algorithm. This present work is an extension of that work; here we present the result of the comparative analysis of the similarity check mechanism used in GEMS and MOGAMOD after implementing them in C programming language on Linux platform.

The structure of this paper is as follows. In section 2, we discuss the technical details of GEMS and MOGAMOD's similarity check. In section 3, we show the experimental experience of running the algorithm on the data of interest and a discussion of the result and we conclude the paper in section 4.

2.0 GEMS AND MOGAMOD SIMILARITY CHECK MECHANISM

In [8] MOGAMOD algorithm was introduced using multi-objective genetic algorithm and it was used to discover optimal motifs in sequential data. Multi-objective optimization involves having a solution which is a family of pareto-optimal set or nondominated solutions . He converted the optimal motif discovery problem into three conflicting optimization problem which is to maximize similarity, motif length and support for candidate motifs, thus obtaining a large number of optimal motifs by a single run of the algorithm. The implementation of MOGAMOD which was based on a well known high performance multi-objective genetic algorithm called NSGA II [1], which is a global multi-objective optimization problem solver can be applied to any field of optimization with conflicting objectives. NSGA II is unique in that unlike other optimization solvers which convert multiple objectives into a single one by using some subjective preference information, NSGA II is capable of finding a well distributed set of trade-off optimal solutions for two or more conflicting objectives of design. MOGAMOD was compared with three well-known motif discovery methods AlignACE, MEME and Weeder using yeast data from TRANSFAC. The result showed that MOGAMOD outperformed them in terms of accuracy and runtime.

The similarity check used in MOGAMOD measures similarity among all motif instances defining an individual solution. To compute the similarity, it first generates a position weight matrix from the motif patterns found in every sequence. Then, the dominance value (dv) of the dominant nucleotide in each column is found using the formula:

$$dv(i) = \max_b\{f(b,i)\} \text{ , } i = 1,\ldots., l$$

Where *f(b, i)* is the score of the nucleotide *b* on column *i* in the position weight matrix,
dv(i) is the dominance value of the dominant nucleotide on column *i*,
and *l* is motif length.

The similarity objective function of motif *M* is the average of the dominance values of all columns in the position weight matrix.

i. e $Similarity(M) = \sum_{i=1}^{i} \mathrm{dv(i)}/l$

The likelihood of the candidate motif been discovered as a real motif depends on the value of the similarity score. In other words, the closer the value of the similarity *M* is to one, the greater the probability that the candidate motif *M* will be discovered as an optimal motif.

Figure 1 below depicts the steps involved in the similarity check of MOGAMOD

The motif discovery tool by [8] Gene Enrichment Motif Searching (GEMS) used hyper geometric-based scoring function (to calculate p-values for position weight matrices (PWMs)) and a position weight matrix optimization routine to identify with a high degree of accuracy simple motifs in the nucleotide-biased and repeat sequence rich genome of *P. falciparum* The PWMs were built from seeds with the most enriched candidates i.e. those with lowest p-values, while identifying all sequences with one mismatch from the seed words, then a *p*-value enrichment score is computed using a hypergeometric formula below.

$$P(X,x,Y,y) = \sum_{i=y}^{min(x,Y)} \frac{\binom{x}{i}\binom{X-x}{Y-i}}{\binom{X}{Y}} \quad \text{.....(1)}$$

Where X is the total set of genes, i.e. positive and negative set, *x* a subset of the gene of interest,

Y is the total promoter sequence that matches the genes, *y* is the subset of the promoters which fall within the cluster of interest. The smaller the p-value scores for a motif, the higher the likelihood of it being an optimal motif.

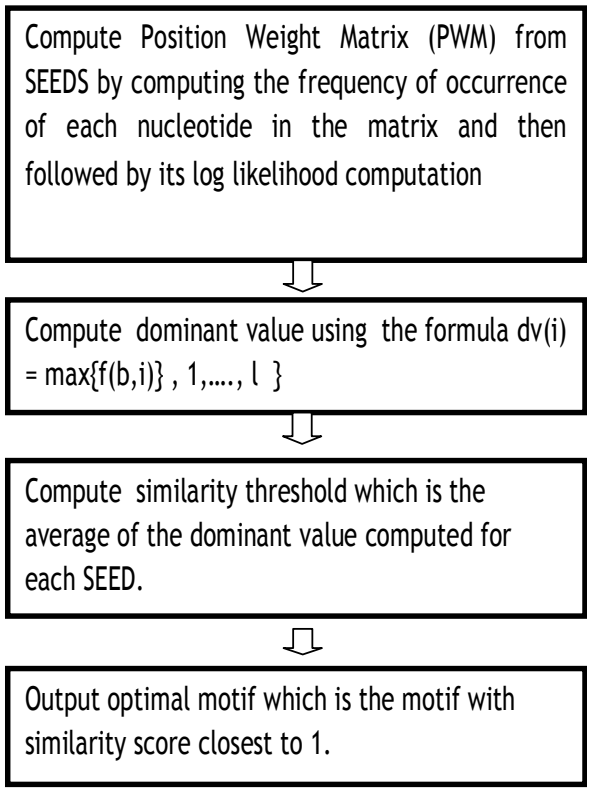


Fig 1. MOGAMOD's Similarity Check Implementation

The success of GEMS in extracting significant motifs for *P. falciparum* was based on using this hypergeometric scoring function i.e. using geometric mean to calculate similarity function and setting a threshold based on the p-value of the position weight matrix. The threshold setting was achieved by utilizing an exhaustive parameter optimization routine similar to the probability minimization protocol used in the OPI clustering algorithm of [10]. This threshold helps to determine how similar any given sequence in a promoter region must be to the PWM to be considered an actual motif. In addition, GEMS merged non-unique motif candidates using a distance metric. This approach is better for *P. falciparum* which has highly repetitive sequences present.

The approach canceled out these repetitions where others motif discovery tools especially those using background modeling approach

would identify the same repetitive sequence as potential motif because they vary significantly from the background estimation.

The details of the algorithm are encapsulated in figure 2.

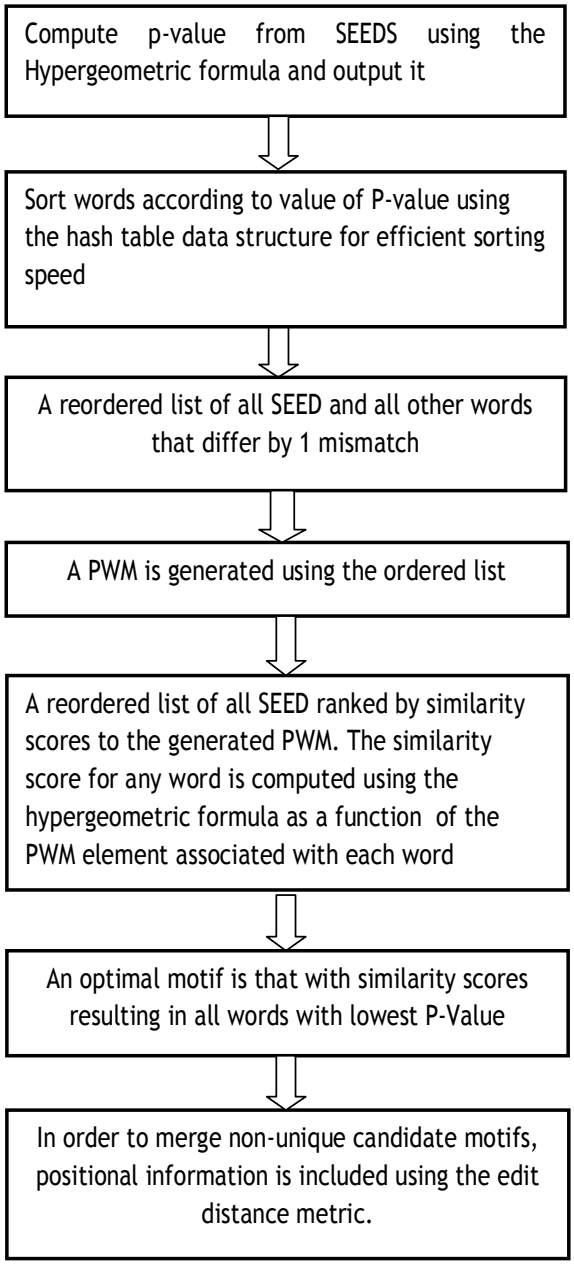


Figure 2: Similarity Check Implementation of GEMS

3.0 EXPERIMENTAL EXPERIENCE AND DISCUSSION OF RESULTS

In this section, we show the effectiveness (in mining binding sites (motifs) of our approach. To this effect, two sets of sample genes in *P.falciparum*, which have been experimentally proven to co-regulate via structured motifs were used for testing. The implementation and testing of algorithm was done in C.

In our implementation of gene enrichment searching, we used the hash table to store the extracted SEED while sorting it according to the p-values computed. The hash table data structure is a desirable choice because of the speed advantage in sorting large data sizes.

The first experiment used the set of genes in the work of [2] which experimentally extracted regulatory elements for P.*falciparum*. i.e 100 base pairs upstream of gene start codons as shown in table 1. Table 2 shows the result obtained running the data set on the two implementations, i.e. the first used GEMS's similarity check while the second used the similarity check used in MOGAMOD. The second experiment used the set of genes used by [9] which identified transcription factors in the mosquito-invasive stage of malaria parasite shown in table 3. The resulting output using the two implementations are depicted in table 4.

From [2], experimentally, the set of genes in table 1 co-regulate using the following motif: N(C/G/A)TGCA-4to5-(A/G/C)GTGC(A/G). 'N' indicates any of the four nucleotides A/C/G/T can occur at this position, while four to five gaps are between the two boxes. In [9], it was also experimentally shown that TAGCTA-100to1500-TAGCCA and TAGCTA-100to1500-TGGCTA are the structured motifs used in their co-regulation.

Table1. Set of genes from Plasmodium falciparium Intraerythrocytic stage

Accession Number	Description
PFF0645c	*Plasmodium falciparum* 3D7 , integral membrane protein, putative
PFI0265c	*Plasmodium falciparum* 3D7, high molecular weight rhoptry protein
PFE0075c	*Plasmodium falciparum* 3D7, rhoptry-associated protein 3
PFE0080c	*Plasmodium falciparum* 3D7, rhoptry-associated protein 2
PFC0120w	*Plasmodium falciparum* 3D7 , cytoadherence linked asexual protein 3.1
MAL7P1.208	*Plasmodium falciparum* 3D7, rhoptry-associated membrane antigen
PFI1730w	*Plasmodium falciparum* 3D7, cytoadherence linked asexual protein 9
PF14_0102	*Plasmodium falciparum* 3D7, rhoptry-associated protein 1
PFD0295c	*Plasmodium falciparum* 3D7 , apical sushi protein
MAL7P1.119	*Plasmodium falciparum* 3D7 , rhoptry-associated leucine zipper-like protein 1
PFI1445w	*Plasmodium falciparum* 3D7, high molecular weight rhoptry protein 2

Table 2. Output from running the two algorithms on the DNA sequences of the above genes

	IDENTIFIED BY	
Consensus	**GEM's Similarity Check**	**MOGAMOD's similarity Check**
GGTGCG	NO	NO
CGTGCG	NO	NO
CTGCA	YES	NO
GTGCA	YES	NO
ATGCA	YES	NO
AGTGCG	YES	NO

Table 3. Set of genes from the Mosquito invasive stage of malaria parasite.

Accession Number	Description
PF08_0136b	*Plasmodium falciparum* 3D7 , von Willebrand factor A-domain related protein
PFC0905c	*Plasmodium falciparum* 3D7, oocyst capsule protein
PFL0550w	*Plasmodium falciparum* 3D7, HSP20-like chaperone
PFC0640w	*Plasmodium falciparum* 3D7,CSP and TRAP-related protein
PFD0425w	*Plasmodium falciparum* 3D7 , sporozoite invasion-associated protein 1, putative
PF08_0030	*Plasmodium falciparum* 3D7, conserved Plasmodium protein, unknown function
PFL2135c	*Plasmodium falciparum* 3D7, conserved Plasmodium protein, unknown function
MAL13P1.203	*Plasmodium falciparum* 3D7 , secreted ookinete protein, putative
PF10_0027	Plasmodium falciparum 3D7,conserved Plasmodium protein, unknown function
PFL2510w	*Plasmodium falciparum* 3D7, chitinase
PF13_0355	*Plasmodium falciparum* 3D7, secreted ookinete protein
PFD0435c	*Plasmodium falciparum* 3D7, conserved Plasmodium protein
PFE0360c	*Plasmodium falciparum* 3D7, conserved Plasmodium protein
PF14_0040	*Plasmodium falciparum* 3D7, secreted ookinete adhesive protein
PFF0975c	*Plasmodium falciparum* 3D7, conserved Plasmodium protein
PF10_0302	*Plasmodium falciparum* 3D7, 28 kDa ookinete surface protein
PF10_0303	*Plasmodium falciparum* 3D7, 25 kDa ookinete surface antigen precursor
PFC0420w	*Plasmodium falciparum* 3D7, calcium dependent protein kinase 3
PFI1145w	*Plasmodium falciparum* 3D7, perforin like protein 3

Table 4. Output from running the two algorithms on the DNA sequences of the above genes

	IDENTIFIED BY	
Consensus	**GEM's Similarity Check**	**MOGAMOD's similarity Check**
TAGCTA	NO	NO
TGGCTA	NO	NO
TAGCCA	NO	NO

From table 2 above, we observed that the experimentally extracted motif was also mined by GEMS's similarity check but not by MOGAMOD's. However, in table 4, none of the experimentally extracted motifs was found by GEMS's similarity or by MOGAMOD's. This inability to mine the motifs in table 4 makes it obvious that a number of fine tunings, which are not necessarily algorithmic, are needed to effectively mine the desired structured motifs in the set of table 3.

A novel algorithm, STGEMS which incorporated the similarity mechanism used in GEMS has been shown to be more effective when compare to MOGAMOD's similarity check. This is because of the hypergeometric scoring function used in GEMS's similarity check mechanism which made it successful in discovering motifs from the challenging highly repetitive elements and base bias sequence of the malaria parasite, *Plasmodium falciparum*.

4.0 CONCLUSION AND FUTURE LEADS

A comparative analysis of the similarity mechanism of two popular motif discovery algorithms was achieved in this work. The result shows the superiority of the similarity check of GEMS over that of MOGAMOD especially when discovering motifs from organisms with some peculiarities in their genomic sequences such as the malaria parasite, *Plasmodium falciparum*. A possible future work would be to formalize the fine tunings required to effectively extract biologically motivated motifs as indicated in the observation made in session 3 above.

References

[1] Deb, K et al., "A fast and elitist multi-objective genetic algorithm II." IEEE Transactions on Evolutionary Computation, 6, 182–197, 2002

[2] Flueck C, Bartfai, R, Niederwieser, I, Witmer, K, Alako, B, Moes, S, Bozdech, Z, Jenoe,P, Stunnenberg, H, Voss T., "A major Role for the Witmer, K, Alako, B, Moes, S, Bozdech, Z, Jenoe,P, Stunnenberg, H, Voss T., "A major Role for the Plasmodium *falciparum* ApiAP2 Protein PfSIP2 in Chromosome End Biology", PLoS Pathog 6(2): e1000784, 2010.

[3] Kaya, M., "MOGAMOD: Multi-Objective Genetic Algorithm for Motif Discovery", Expert Systems with Applications, 36 (2): 1039-1947, 2009.

[4] Makolo, A, Adebiyi E and Osofisan A., "STGEMS: Mining Structured Motifs with Gene Enrichment Motif Searching on Suffix tree", Journal of Computer Science and its Applications 18(1) : 79-91, 2011.

[5] Matys V., Fricke,E., Geffers,R., Gossling,E., Haubrock,M., Hehl,R., Hornischer,K., Karas,D., Kel,A.E. et al.,"TRANSFAC: transcriptional regulation, from patterns to profiles", Nucleic Acids Res., 32,D81–D96, 2003.

[6] Sandelin, A., Alkema,W., Engstrom ,P., Wasserman,W.W. and Lenhard,B. "JASPAR: an open-access database for eukaryotic transcription factor binding profiles". Nucleic Acids Res., 32,D91–D94, 2003

[7] Tompa. M., "An Exact Method for Finding Short Motifs in Sequences with Application to the Ribosome Binding Site Problem", 7th Intl. Conf. Intelligent Systems for Molecular Biology, Heidelberg, Germany, Aug 1999 10-12, 2003.

[8] Young J, Johnson, J, Benner, C, Yan, F, Chen, K, Roch, K, Zhou, Y, Winzeler, E., "In silico discovery of transcription regulatory elements in Plasmodium falciparum", BMC Genomics ,9:70, 2008.

[9] Yuda, M, Iwanaga, S, Shigenubu, S, Mair, G, Janse, C, Waters, A, Kato, T, Kaneko, I., "Identification of a transcription factor in the mosquito-invasive stage of malaria parasite. Molecular Microbilogy, 71, 1402-1414, 2009.

[10] Zhou Y, Young JA, Santrosyan A, Chen K, Yan SF, Winzeler EA. "In silico gene function prediction using ontology-based pattern identification". *Bioinformatics*, 21(7):1237-1245,2005

Combining Optical Character Recognition (OCR) and Edge Detection Techniques to Filter Image-Based Spam

B. Fadiora
Department of Computer Science
The Polytechnic Ibadan
Ibadan, Nigeria
tundefadiora@yahoo.com

[2]F. Wada & [3]O.B. Longe
[2]Nelson Mandela School of Public Policy
[3]Fulbright Fellow and Research Scholar
International Centre for Information Technology & Development
Southern University, Baton Rouge, Louisiana, USA.
friwada@yahoo.com longeolumide@fulbrightmail.org

ABSTRACT

Early spam filtering systems were designed for text-based spam format. However, the emergence and prevalence of image-based spam mails has resulted into newer challenges that renders mechanisms for filtering text-based spam mails inefficient for image-based spam. This research work employed pattern recognition (Grey scaling, edge detection and binarization) in combination with Object Character Recognition (OCR) mechanisms to analyze and classify image-based spam e-mails. The resulting system is an enhancement of existing OCR systems that suffers from the problem of noise generated around spam messages embedded in images. Findings from research show that the hybridized system outperforms OCR-based systems in terms of classification and efficiency.

Keywords: **Image, Bayesian classification, spam, Optical Character Recognition.**

1. INTRODUCTION

Spam mails are unsolicited, unwanted large amount of messages in the mailboxes of users. Another version of this is also called the Unsolicited Commercial E-mails. Most spam take the form of advertising or promotional materials ranging from debt reduction plans to getting-rich-quick schemes.

African Journal of Computing & ICT Reference Format:

B. Fadiora, O.B. Longe & F. Wada (2012) - Combining Optical Character Recognition and Edge Detection to OCR and Edge Detection to Filter Image-based Spam E-Mails. Afr J. of Comp & ICTs. Vol 5, No.1 pp 59– 68

- ISSN 2006-1781

One third of spam mails are porn-based, health-related, and the rest of them cover a variety of topics in the way of promotion of products from companies or get rich quick schemes [5][10]. Several modules are embedded inside spam filters which are meant to devotedly analyze different components of e-mails (sender's address, header, body, attachments) and to detect specific characteristics of spam e-mails [8]. The degree of spam of such an e-mail is then calculated from the outputs of the different modules when they are combined. For example, consider an e-mail user who receives an unsolicited mail from another user, with some of the content of the mail repeatedly changed for other words (e.g. in for, replacing "o" with "0"), could be an indication of spam.

This example describes how content based anti spam filtering works. Some approaches that have been adopted in fighting spam include:

- White listing
- Blacklisting
- Keyword-based or Bayesian filters
- Fingerprinting (for Image spam)
- Content based spam filters

When the spammers discovered that their approach and techniques are being countered by different anti spam filtering procedures and techniques, they changed their methods of operation by stepping up their efforts by resorting to image spam. Image spam "is a kind of spam in which the text message is embedded into attached images to defeat spam filtering techniques based on the analysis of e-mail's body text" [2][3][4][7]. Conventional text based spam filtering cannot handle the image spam, thus they pass through the filters successfully.

Spammers do meet brick walls in the different approaches until they manoeuvre their tactics using image spam. This is created by embedding the spam inside a JPEG or any other image format and then attaches and sends it as attachment. There is no way any text based filtering technique can stop this. What was employed is using the Optical Character Recognition, which helped a great deal. Image-based spam contains scanned or computer generated words that make reader gets angry.

There are features that are common to all images that spammers take advantage of. These includes: File Format, image metadata, average color, color saturation, random pixel test.

After many image-based anti spam procedure were developed, spammers again resorted to generate noise around the images created to fool and foil the Optical Character Recognition (OCR). This is called obfuscation. However, the drawback of this process is that it takes processor time (when it is handling multitudes of images every few seconds [3][9].

"Image spam is really succeeding because spammers have made image spam stubborn by multiple filter-thwarting techniques. Some of the techniques adopted are randomizing image borders, animated GIF with embedded spam image, image segmentation etc. Averagely, a single image spam message is twice the size of an HTML or text-based spam. All this creates a drain on the server, storage and bandwidth resources" [9].

2. RELATED WORKS

In [3] it was mentioned that image spam filtering is a pattern recognition task in adversarial environment. It was shown by experiments that filtering of adversarial obscured images is very challenging, thus the actions of spammers for avoiding classifiers have to be taken into account explicitly. The results achieved showed it is possible that spammers can evade OCR tools quite easily using obscured text images without compromising human readability.

Thus, an approach was brought forward to filter obscured spam images which were based on the detection of obfuscated text (an approach which takes into account explicitly the adversarial environment). Three features were devised, aimed at detecting defects in some kinds of text image (e.g., character fragmentation or fusion, presence of background noise around text, etc) that are caused by obfuscation techniques used by spammers to fool OCR effectiveness.

Two of their features are based on the *perimetric complexity measure*, which is able to discriminate images of clean characters from fragmented or fused characters and background noise components in the binarized image. In "Improved Spam Filtering by Extraction of Information from Text Embedded Image E-mail" [11] an approach based on the two level of ontology spam filters; which are first level global ontology filter and user customized ontology filter was introduced.

It was believed that adding the image filtering dimension is able to make filtering efficient and robust. As mentioned earlier, OCR (Optical Character Recognition) was used, which was responsible for translation of images (e.g. scanned documents), into actual text characters. OCR has the ability of editing and reusing the text that is normally locked inside scanned images. In order to get best result from the OCR, accuracy of 95% is maintained.

From the training data set, we get the image e-mails which are fed into the OCR and then text information is retrieved from text embedded image e-mail. Training data set is a collection of text-oriented e-mail data. A technique called TF-IDF (term frequency-inverse document frequency) is used to extract features from the data set. A *Weka* input file is created based on the selected features and the data set. With this, classification results are generated. The classified results are converted to RDF file.

The converted RDF file is fed into *Jena*, which is a Java framework for building Semantic Web applications. Ontologies are created using Jena, and query can be given to Jena. Jena will give an output for the query using ontologies created in *Jena*. Through these procedures, global and user-customized ontology filters are created. Incorrectly classified e-mails through global ontology filter are inserted into the user customized ontology filter.

The training data set is the set of e-mail that gives us a classification result. It is composed of both text e-mail and image e-mail. The test data is actually the e-mail will run through our system which we test to see if classified correctly as spam or not.

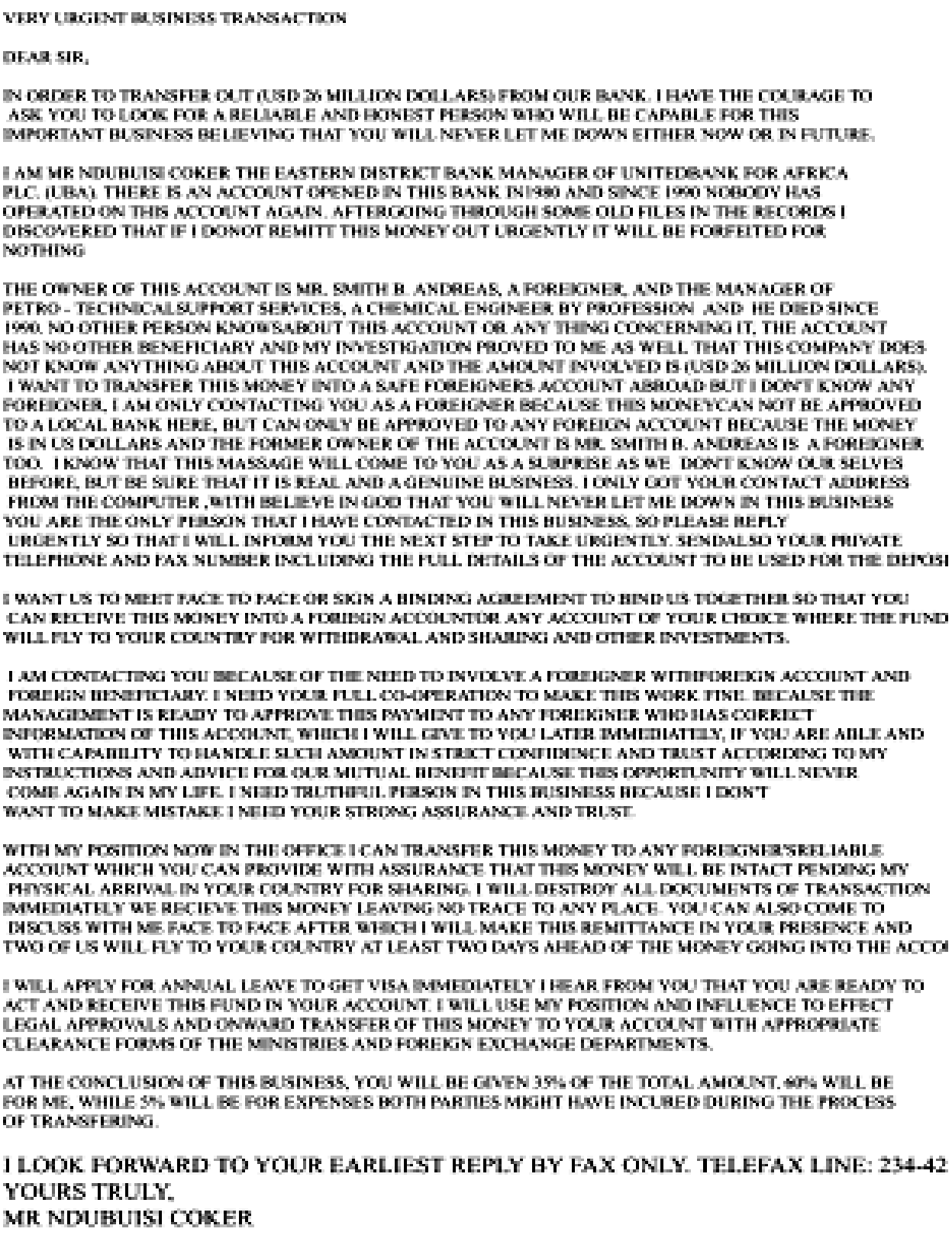
In [1] and [2] an attempt was made to improve the filtering of image based spam by using image text features. In this journal, it was argued that though OCR modules could be used to filter image based spam, the drawback of the OCR approach takes much of the time of the processing time of the CPU and secondly it cannot be effective where the images are not clean therefore it is easy for spammers to obfuscate the text into images.

Thus, a spam filter equipped with an OCR based module as the unique countermeasure against image spam is vulnerable to image spam with obfuscated text.

In [4] the authors described an anti-spam filtering framework that combines text-based and image-based anti-spam filters. It is an incremental framework that starts by reducing mismatches between training and test datasets to resolve the problem of a lack of training data for legitimate e-mails that contain both text and images. Thereafter, the outputs of text-based and image-based filters are combined with the weights determined by a Bayesian framework.

3. OUR APPROACH

Our approach to the image-based spam detection employ edge detection which looked for the probable edge in the particular image we are dealing with, greyscaling whose function is to remove the remaining white and black element in the remaining image and the binarization techniques which made the foreground to be black while having a white background. Colour Channel splitting is another tool that came in handy. It eliminated the three basic colours in images (Red, Green and Blue). We followed it by the binarisation method which makes the foreground to be black while having a white background. The remaining texts were extracted by the Optical Character Recognition and were passed to the classifier, Bayesian classifier which determined from the result whether the image is spam or not.



VERY URGENT BUSINESS TRANSACTION

DEAR SIR,

IN ORDER TO TRANSFER OUT (USD 26 MILLION DOLLARS) FROM OUR BANK. I HAVE THE COURAGE TO ASK YOU TO LOOK FOR A RELIABLE AND HONEST PERSON WHO WILL BE CAPABLE FOR THIS IMPORTANT BUSINESS BELIEVING THAT YOU WILL NEVER LET ME DOWN EITHER NOW OR IN FUTURE.

I AM MR NDUBUISI COKER THE EASTERN DISTRICT BANK MANAGER OF UNITEDBANK FOR AFRICA PLC. (UBA). THERE IS AN ACCOUNT OPENED IN THIS BANK IN1980 AND SINCE 1990 NOBODY HAS OPERATED ON THIS ACCOUNT AGAIN. AFTERGOING THROUGH SOME OLD FILES IN THE RECORDS I DISCOVERED THAT IF I DONOT REMITT THIS MONEY OUT URGENTLY IT WILL BE FORFEITED FOR NOTHING

THE OWNER OF THIS ACCOUNT IS MR. SMITH B. ANDREAS, A FOREIGNER, AND THE MANAGER OF PETRO - TECHNICALSUPPORT SERVICES, A CHEMICAL ENGINEER BY PROFESSION AND HE DIED SINCE 1990. NO OTHER PERSON KNOWSABOUT THIS ACCOUNT OR ANY THING CONCERNING IT. THE ACCOUNT HAS NO OTHER BENEFICIARY AND MY INVESTIGATION PROVED TO ME AS WELL THAT THIS COMPANY DOES NOT KNOW ANYTHING ABOUT THIS ACCOUNT AND THE AMOUNT INVOLVED IS (USD 26 MILLION DOLLARS). I WANT TO TRANSFER THIS MONEY INTO A SAFE FOREIGNERS ACCOUNT ABROAD BUT I DON'T KNOW ANY FOREIGNER, I AM ONLY CONTACTING YOU AS A FOREIGNER BECAUSE THIS MONEYCAN NOT BE APPROVED TO A LOCAL BANK HERE, BUT CAN ONLY BE APPROVED TO ANY FOREIGN ACCOUNT BECAUSE THE MONEY IS IN US DOLLARS AND THE FORMER OWNER OF THE ACCOUNT IS MR. SMITH B. ANDREAS IS A FOREIGNER TOO. I KNOW THAT THIS MASSAGE WILL COME TO YOU AS A SURPRISE AS WE DON'T KNOW OUR SELVES BEFORE, BUT BE SURE THAT IT IS REAL AND A GENUINE BUSINESS. I ONLY GOT YOUR CONTACT ADDRESS FROM THE COMPUTER ,WITH BELIEVE IN GOD THAT YOU WILL NEVER LET ME DOWN IN THIS BUSINESS YOU ARE THE ONLY PERSON THAT I HAVE CONTACTED IN THIS BUSINESS, SO PLEASE REPLY URGENTLY SO THAT I WILL INFORM YOU THE NEXT STEP TO TAKE URGENTLY. SENDALSO YOUR PRIVATE TELEPHONE AND FAX NUMBER INCLUDING THE FULL DETAILS OF THE ACCOUNT TO BE USED FOR THE DEPOSI

I WANT US TO MEET FACE TO FACE OR SIGN A BINDING AGREEMENT TO BIND US TOGETHER SO THAT YOU CAN RECEIVE THIS MONEY INTO A FOREIGN ACCOUNTOR ANY ACCOUNT OF YOUR CHOICE WHERE THE FUND WILL FLY TO YOUR COUNTRY FOR WITHDRAWAL AND SHARING AND OTHER INVESTMENTS.

I AM CONTACTING YOU BECAUSE OF THE NEED TO INVOLVE A FOREIGNER WITHFOREIGN ACCOUNT AND FOREIGN BENEFICIARY. I NEED YOUR FULL CO-OPERATION TO MAKE THIS WORK FINE. BECAUSE THE MANAGEMENT IS READY TO APPROVE THIS PAYMENT TO ANY FOREIGNER WHO HAS CORRECT INFORMATION OF THIS ACCOUNT, WHICH I WILL GIVE TO YOU LATER IMMEDIATELY, IF YOU ARE ABLE AND WITH CAPABILITY TO HANDLE SUCH AMOUNT IN STRICT CONFIDENCE AND TRUST ACCORDING TO MY INSTRUCTIONS AND ADVICE FOR OUR MUTUAL BENEFIT BECAUSE THIS OPPORTUNITY WILL NEVER COME AGAIN IN MY LIFE. I NEED TRUTHFUL PERSON IN THIS BUSINESS BECAUSE I DON'T WANT TO MAKE MISTAKE I NEED YOUR STRONG ASSURANCE AND TRUST.

WITH MY POSITION NOW IN THE OFFICE I CAN TRANSFER THIS MONEY TO ANY FOREIGNER'SRELIABLE ACCOUNT WHICH YOU CAN PROVIDE WITH ASSURANCE THAT THIS MONEY WILL BE INTACT PENDING MY PHYSICAL ARRIVAL IN YOUR COUNTRY FOR SHARING. I WILL DESTROY ALL DOCUMENTS OF TRANSACTION IMMEDIATELY WE RECIEVE THIS MONEY LEAVING NO TRACE TO ANY PLACE. YOU CAN ALSO COME TO DISCUSS WITH ME FACE TO FACE AFTER WHICH I WILL MAKE THIS REMITTANCE IN YOUR PRESENCE AND TWO OF US WILL FLY TO YOUR COUNTRY AT LEAST TWO DAYS AHEAD OF THE MONEY GOING INTO THE ACCO

I WILL APPLY FOR ANNUAL LEAVE TO GET VISA IMMEDIATELY I HEAR FROM YOU THAT YOU ARE READY TO ACT AND RECEIVE THIS FUND IN YOUR ACCOUNT. I WILL USE MY POSITION AND INFLUENCE TO EFFECT LEGAL APPROVALS AND ONWARD TRANSFER OF THIS MONEY TO YOUR ACCOUNT WITH APPROPRIATE CLEARANCE FORMS OF THE MINISTRIES AND FOREIGN EXCHANGE DEPARTMENTS.

AT THE CONCLUSION OF THIS BUSINESS, YOU WILL BE GIVEN 35% OF THE TOTAL AMOUNT, 60% WILL BE FOR ME, WHILE 5% WILL BE FOR EXPENSES BOTH PARTIES MIGHT HAVE INCURED DURING THE PROCESS OF TRANSFERING.

I LOOK FORWARD TO YOUR EARLIEST REPLY BY FAX ONLY. TELEFAX LINE: 234-42
YOURS TRULY,
MR NDUBUISI COKER

Fig. 1 – Original Image Spam

Fig 2. The result of our approach

Dear Sir,

I am contacting you on behalf of my family, and I strongly solicite your assistance in the name of the Almighty Allah. I presume my dad did business with your company, because I found your contact address in one of his personal files.

I am Abba Abacha, the son of the late General Sani Abacha a past Head of State of the Federal Republic of Nigeria. Since my father's death, my entire family has been under immense pressure from the present Government of Nigeria, to make available the funds in my father's possession. As a result of this, my family has lost over US$75,000,000.00, which my father left behind for his family made numerous attempts to travel out of the country with the funds my father kept in his country home, but all her efforts were frustrated by the government of the day. This gave my mother a lot of psychological problems.

As a result of these failed attempts, I decided to confidentially move some of the funds to a neighboring country, this I succeeded in doing without the knowledge of my country's Government. The funds are in a total amount of US$12,500,000.00, and sealed in trunk boxes. I have made arrangements with a Security Company where I lodged the trunk boxes as photographic printing materials, to help airlift the trunk boxes to my would be destination (your country). At the moment, I cannot travel out of the country, because my country's Government is holding on to my International Passport. It is in view of this, I have decided to solicit your assistance in the dispersing of these funds in a destination location assigned to you. Remuneration is open to negotiation.

I appeal to you in the name of almighty Allah, the beneficent and the merciful, to sincerely and honestly render us this assistance. I know you would wonder why I am not soliciting for assistance from well placed Nigerians, it's really sad. You see, all my late father's friends whom we have asked for

Fig 3. Output after Grayscale operation

Then, we send the image to the Optical Character Recognition (OCR) engine to extract the text. The traditional document processing steps which involve tokenization, indexing and classification) have been extended by including in the tokenization phase the plain text extraction by OCR from attached images, besides plain text extraction from the subject and body fields.

The OCR engine extract characters (letters, numbers and barcodes) from the images and from there the characters are compared with the database of words suspected to be used by spammers. Finally, we use Bayesian filtering techniques to ascertain whether the image is a spam or a ham.

$$Ham\ Probability = \frac{Token\ frequency\ in\ Ham\ messages}{Number\ of\ Ham\ messages\ in\ the\ training\ set}$$

.....(1)

$$Spam\ Probability = \frac{Token\ frequency\ in\ Spam\ messages}{Number\ of\ Spam\ messages\ in\ the\ training\ set}$$

.....(2)

$$Spamicity = \frac{Spam\ probability}{Ham\ probability + Spam\ probability}$$

.....(3)

Mathematically, the Bayesian general formula is represented as:

$$P(A,B) = \frac{ab}{ab + (1-a)(1-b)}$$

.....(4)

4. CLASSIFICATION OF IMAGE-BASED SPAM E- MAILS

Image-based e-mails are e-mails in which the spam mails or text are embedded inside images. In order to filter, classify these images, we have to do some image processing where we have to alter the images in order to extract the text and then classify the texts so that they are filtered.

E-mails have to be classified depending on the content of the e-mail. After the analysis and extraction of the edge image features of each e-mail, what follows is to classify that e-mail as to whether it is a spam or ham mail. The classification can be achieved in various ways. We want to use a machine learning approach which is unsupervised such as the Bayesian approach for the training.

6. 5AYESIAN FILTERING APPROACH

Bayesian filter is a content based filter named after Thomas Bayes. A formula was developed which allowed him to determine the probability of an event occurring based on the probabilities of two or more independent evidentiary events. In order to be very efficient, the filter has to be trained in order to recognize good from bad e-mails. A database is created where extracted tokens are stored. When analyzing a new message, the message is split into tokens and each token is given a value according to the following criteria:

1. The frequency of the token in good messages that the filter has been trained on
2. The frequency of the token in spam messages that the filter has been trained on
3. The number of good messages the filter has been trained on
4. The number of spam messages the filter has been trained on

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum P(B|A_i)\ P(A_i)}$$

.....(5)

A Bayesian filter splits the e-mail message into tokens. From this, a table of dictionary is built for all the tokens that the program will use in making its decision. Tokens can be defined as a character or a group of characters separated by a space or special characters.

Once the Bayesian filter has the list of tokens in the message, it searches the spam and non-spam token databases for these tokens, calculates each tokens spamicity. Depending on what the resulting figure or a particular threshold value, the Baye‘s filter can then determine if the e-mail massage containing of the tokens are spam messages. However, they can be used to filter only text based e-mail messages.

We are therefore interested in filtering the text inside the images, hence the reason for using the Bayesian techniques. Words that occur disproportionately rarely in spam (like "though" or "tonight" or "apparently") contribute as much to decreasing the probability as bad words like "unsubscribe" and "opt-in" do to increasing it. The Bayesian method takes the whole message into account.

It recognizes keywords that identify spam, but it also recognizes words that denote valid mail. For example: not every e-mail that contains the word “free” and “cash” is spam. The advantage of the Bayesian method is that it considers the most interesting words (as defined by their deviation from the mean) and comes up with a probability that a message is spam. Bayesian filtering is a much more intelligent approach because it examines all aspects of a message, as opposed to keyword checking that classifies a mail as spam on the basis of a single word.

Bayesian filtering learns easily [1], when it learns from new spam and new valid outbound mails, the Bayesian filter evolves and adapts to new spam techniques. Citing an example, spammers started using “f-re-e“ instead of “free“ they also succeeded in evading keyword checking until “f-r-e-e“ was also included in the keyword database.

On the other hand, the Bayesian filter automatically notices such tactics; in fact if the word “f-r-e-e“ is found, it is an even better spam indicator, since it’s unlikely to occur in a ham mail. Another example would be using the word “5ex” instead of “Sex“. One might not have 5ex in a ham mail, and therefore the likelihood that it is spam increases. Furthermore, as strong as Bayesian filtering is, it is very difficult to fool. This is very unlike a keyword filter. An advanced spammer who wants to trick a Bayesian filter can either use fewer words that usually indicate spam.

Lastly, Bayesian method for filtering of mail is multi-lingual and international. Because the Bayesian anti-spam filter is adaptive, it can be used for any language required. Most keyword lists are available in English only and are therefore quite useless in non English-speaking regions. The Bayesian filter also takes into account certain languages deviations or the diverse usage of certain words in different areas, even if the same language is spoken. This intelligence enables such a filter to catch more spam” [6].

7. Creating Bayesian word database

Sequel to filtering of spam mail using the Bayesian filtering method, a database needs to be created with the tokens of words collected from a sample of spam mail and valid mail (referred to as ‘ham’). A *token* is an instance of a sequence of characters in some particular document that are grouped together as a useful semantic unit for processing. Next is the assignment of probability value to each word or token; this probability is based on calculations that take into account how often that word occurs in spam as opposed to legitimate mail (ham).

This is done by analyzing the users' outbound mail and analyzing known spam: All the words and tokens in both pools of mail are analyzed to generate the probability that a particular word points to the mail being spam. Here is an example of how probability is calculated. If the word "mortgage" occurs in 400 of 3,000 spam mails and in 5 out of 300 legitimate e-mails, then its spam probability would be 0.8889 (that is, [400/3000] divided by [5/300 + 400/3000]).

After the two databases have been created, the word probabilities can then be calculated and the filter is ready for use. At the incoming of a new mail, it is broken down into words and the most relevant words i.e., those that are most significant in identifying whether the mail is spam or not are singled out. From these words, the Bayesian filter calculates the probability of the new message being spam or not. If the probability is greater than a threshold, say 0.9, then the message is classified as spam. With Bayesian method, spam detection rates of over 99.7% can be achieved with a very low number of false positives. We chose Bayesian filters because if given appropriate time and training data, Bayesian filters can achieve a combination of extremely high accuracy rates with a low percentage of false positives (Longe *et al*, 2007).

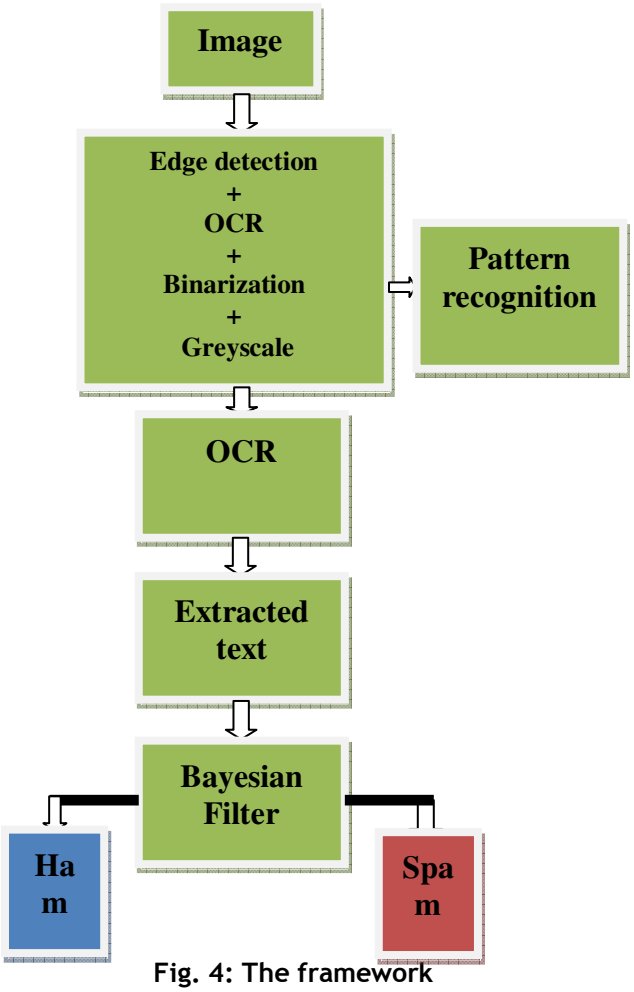


Fig. 4: The framework

7. CONCLUDING REMARKS

Early spam filtering systems were designed for text-based spam format. However, the emergence and prevalence of image-based spam mails has resulted into newer challenges that renders mechanisms for filtering text-based spam mails inefficient for image-based spam.

In this work, we employed pattern recognition (Grey scaling, edge detection and binarization) in combination with Object Character Recognition (OCR) mechanisms to analyze and classify image-based spam e-mails.

This efforts have been able to provide an enhancement for existing OCR systems that suffers from the problem of noise generated around spam messages embedded in images. Our system also outperforms OCR-based systems in terms of classification and efficiency.

8. REFERENCES

1. Chiemeke, S.C., Longe, O.B., Onifade, O.F.W, Longe, F.A. (2007): Text Manipulations and Spamicity Measures: *Implications for Designing Effective Filtering Systems for Fraudulent 419 Scam Mails*. Paper presented at the international conference of Adaptive Science and Technology, Accra.
2. Battista Biggio, Giorgio Fumera, Ignazio Pillai, Fabio Roli, (2007): *Image Spam Filtering by Content Obscuring Detection*, Fourth Conference on E-mail and AntiSpam, Mountain View, California USA
3. Biggio, B., Fumera, G., Pillai, I., Roli, F., (2008), *Improving Image spam filtering Using Image Text Features,* Fifth conference on E-mail and Antispam, CEAS 2008
4. Byungki Byun, Chin-Hui Lee, Steve Webb, Danesh Irani, and Calton Pu, (2009): *An Anti-spam Filter Combination Framework for Text-and-Image E-mails through Incremental Learning,* Conference on E-mail and Spam.

5. Danquah, P & Longe, O.B and P.Danquah & O.B. Longe (2011). An Empirical Test Of The Space Transition Theory of Cyber Criminality: The Case of Ghana and Beyond. Afr J. of Comp & ICTs. Vol 4, No. 2. pp 37-48
6. GFI White Paper on Image Spam Filtering. (2009) www.gfi.com.imagespam/...
7. Longe, O.B.(2011). On the use of Image-based Spam Mails as Carriers for Covert Data Transmission. Computer & Information Systems Journal. Vol. 15. No.1 http://cis.uws.ac.uk/research/journal/vol15.htm
8. Longe, O.B. & Osofisan, O.A. (2011). On the Origins of Advance Fee Fraud Electronic Mails: A Technical Investigation Using Internet Protocol Address Tracers," The African Journal of Information Systems: Vol. 3: Iss. 1, Article 2. http://digitalcommons.kennesaw.edu/ajis/vol3/iss1/2
9. Proof point MLX Review: *Next generation solution for spam a predictive approach*. Proof point MLX Whitepaper, 2008. http://whitepapers.techrepublic.com.com/whitepaper.aspx?docid=291451
10. Sahami, M., Dumais, S., Heckerman, D. and Horvitz, E., 1998: *A Bayesian approach to filtering junk E-mail*. AAAI Technical Report WS-98-05, Madison, Wisconsin.
11. Youn Seongwook and McLeod Dennis, 2009, *Improved Spam Filtering by Extraction of Information from Text Embedded Image E-mail,* SAC'09, Honolulu, Hawaii, U.S.A.

.

Electronic Banking and Cyber Crime In Nigeria - A Theoretical Policy Perspective on Causation

F. Wada
Nelson Mandela School of Public Policy
Southern University
Baton Rouge, USA
friwada@yahoo.com

G.O. Odulaja
Department of Computer Science
Tai Solarin University of Education
Ijagun-Ijebu Ode, Nigeria
goddyseyi@yahoo.com

ABSTRACT
The information communication technology (ICT) revolution has had impacts in almost every area of human endeavor. From business, industry, government to not-for-profit organizations, ICT has simplified business processes such as sorting, summarizing, coding, editing, customized and generic report generation in a real-time processing mode. However, ICT has also brought unintended consequences such as criminal activities, spamming, credit card frauds, ATM frauds, phishing, identity theft and other related cyber crimes. This study sought to assess cyber crime and its impact on the banking institutions in Nigeria. It also examined the existing policy framework and assessed the success of the institutional countermeasures in combating cyber crime in the banking industry. This papers X-rays cybercrime policy issues and provide insight into how cybercrime impacts on E-banking from a Nigerian perspective. Social theories were then used to explain causation with a view of guiding policy makers on behavioural issues that should be considered when formulating policies to address cyber criminal activities in Nigeria.

Keywords: E-banking, Nigeria, policy, social theories, frauds, banking.

1. INTRODUCTION

There are few innovations that have changed the dynamics of banking as much as the e-banking revolution. Throughout the world, banks are reorganizing their business strategies to take advantage of new business opportunities offered by e-banking. Electronic banking is believed to have started in the early 1980s [68]. It has since then been growing in an unprecedented dimension in line with the growth in ICT development. E-banking has enabled banks to overcome borders, adopt strategic outlook, and bring in new possibilities. According to Nitsure 55], information communication technology has reduced the cost of processing and facilitating the transmission of information leading to drastic changes in the banking business. It is worth noting that e-banking has not been limited to advanced countries, but is found even in countries with underdeveloped e-banking systems, as a result of the many new business opportunities offered by e-banking.

African Journal of Computing & ICT Reference Format:

F. Wada & G.O. Odulaja (2012) . E-Banking and Cyber Crime in Nigeria – A Theoretical Policy Perspective on Causation . Afr J. of Comp & ICTs. Vol 5. No. 1. pp 69-82.

- ISSN 2006-1781

Although no official definition of e-banking has been established, it generally implies a service that allows customers to use some form of computer to access account-specific information and possibly conduct transactions from a remote location like home or workplace. Additionally, e-banking has obvious advantages to the customer in terms of convenience where customers conduct routine banking transactions from the comfort and security of any location from which they wish to transact [43]. The emerging concept of e-banking has drawn the attention of the business fraternity as well as of scholars and researchers to the effects of such dynamics on the banking industry. For instance, Liao and Wong [43] in their study of the determinants of customer interactions with Internet-enabled e-banking found that factors such as perceived usefulness, ease of use, security, convenience, and responsiveness to service requests to be a strong measure of the variation in customer interactions. Based on this finding, they suggested that stringent security control is critical to e-banking operations. Such arguments do not only have managerial implications for enhancing Internet banking operations and developing viable electronic banking services, but also form the basis upon which this study is based.

Anguelov et al. [4] hold the same view which suggests that consumer acceptance and use of e-banking technologies are related to the characteristics of both the individual consumer and the specific technology. They further argued that acceptance of such technology is linked to a consumer's socioeconomic, technological, and personal characteristics and preferences. Key investments in e-banking can be traced back to the early 1980s, when the home computer was still rare [70]. Since then, major innovations in electronically-enabled bank service delivery like automatic teller machines, touch-tone telephone banking, voice response units, and centralized technology-intensive telephone call centers have emerged. In 2000, for instance, e-banking was utilized by approximately 10% of all retail banking customers in the United States [51]. According to an August 2009 survey of U.S. consumers by the American Bankers Association (ABA), Internet banking now ranks first among the methods of banking.

Evidence from the existing literature indicates that many customers have adopted the electronic banking services. ATM, for instance, has turned out to be the most popular service delivery channel 18]. It is estimated that the world is now home to 1.65 million ATM's and that number was expected to grow by 1.7 million by 2009 and beyond. Although the USA and other parts of the developed world were the first to experience ATM proliferation, the trend subsequently has stretched to other developing countries [34]. Banks in Malaysia, for instance, have introduced automated teller machines (ATMs) in order to ease the constraints on banking as far as time and geographical location are concerned.

E-banking technology created a revolution by extending banking hours beyond office hours and beyond national boundaries (Balachandran & Balachandher, 2000). In Nigeria, several studies on e-banking have been done. Chiemeke, Evwiekpaefe, & Chete [19], for instance, conducted a pragmatic study on adoption of e-banking where major hindering factors to Internet banking adoption such as insecurity and inadequate operational facilities, including telecommunications facilities and electricity supply, were identified.

In another study on Nigeria, it was revealed that e-banking is still at the infant level in the country with most of the banks having mainly information sites and providing little Internet transactional services. However, most studies in these areas revealed that there has been a very steady move away from cash as transactions are now being automated [2].

While the findings revealed that e-banking is able to enlarge customer relationships, and loyalty and to give banks a competitive advantage as far as market share is concerned, the problems of ineffectiveness of telecommunications services, supply of power, high cost, fear of fraudulent practices, and lack of facilities necessary for their operation still remain. Despite this, ongoing research on the impact of e-banking is inconclusive, especially in developing economies, and serves as an open ground for more research in the area of e-banking.

2. E-BANKING CRIMES

Crime and corruption represent a major concern for business executives not only in Nigeria but also in other parts of Africa. In Nigeria, for instance, the most serious impediments to economic activities and business are crime and corruption which averages 75% and 71% respectively. Theft and fraud are the second most popular crimes after burglary [28][29]. By definition, cyber crime may be referred to as any form of misconduct in cyber space. It is simply defined as the criminal use of the Internet. Cyber crime is believed to have started in the 1960's in the form of hacking. This was followed by privacy violations, telephone tapping, trespassing and distribution of illegal materials in the 1970s. The 1980s witnessed the introduction of viruses 56]. The fast pace of development of ICT from the 1990s till today has added to the list of criminal exploits in cyber space. Today, the Internet is used for espionage and as a medium to commit terrorism and transnational crimes. With e-banking gaining ground in Nigeria and other parts of SSA, customers and online buyers are facing great risk of unknowingly passing on their information to fraudsters. "Hackers" get information of those who have made purchases through websites and then make fake cards, which they use with less detection. Absence of a law specifically dealing with card-related crimes in Nigeria may be giving thieves a loophole to operate freely. Police treat card-related crimes like any other case of fraud.

2.1 Types of Cyber Crime

This study presents the types of cyber crimes that have economic impact either directly or indirectly on the financial system of a nation or having cross border ripple effects. Longe & Chiemeke (2008) simplified the list of unintended consequences of ICT to include acts such as Phishing, cyber terrorism, electronic spam mails, cyber-stalking, and fake copy -cat websites. While some types of cyber crimes are specific to Nigeria, other types, such as identity theft and false statements, cut across all countries.

2.1.1 PHISHING

According to Roger [62], phishing is simply a high-tech identity theft that does not only steal personal information and identity from unsuspecting consumers, but also an act of fraud against the legitimate businesses and financial institutions that are victimized by phishing.

Phishing is usually a social engineering crime pervasive in attacking organisations' or individuals' (customers') information systems (IS) in order to gather private information to be used against organisations to extract some benefit for the perpetrator through the anonymity of identity theft or identity deception acts (Rodger, 2008). According to recent estimates from the Anti-Phishing Working group [6,7], phishing scams remain a relatively small percentage of spam sent worldwide today. Phishing attempts to pose significant dangers for unsuspecting victims. It has become one of the fastest-growing worldwide threats on the Internet. This rapid growth has made combating it a huge priority for electronic mail service providers, since phishing impacts every aspect of the Internet and computing and there is no single action from any one company or organization to solve the problem. The remedy can only come in a holistic fashion involving collaboration between technology innovation, industry, government, and user education as prescriptive guidance.

To build systems shielding users from fraudulent websites, designers need to know which attack strategies work and why. What makes a web site credible? This question has been addressed extensively by researchers in computer-human interaction. Successful phishing must not only present a high credibility web presence to its victims; it must create a presence that is so impressive that it causes the victim to fail to recognize security measures installed in web browsers [61]. Data suggest that some phishing attacks have convinced up to 5% of their recipients to provide sensitive information to spoofed websites [45]. About two million users gave information to spoofed websites resulting in direct losses of $1.2 billion for U.S. banks and card issuers in 2003 [44].

If we hope to design web browsers, websites, and other tools to shield users from such attacks, we need to understand which attack strategies are successful and what proportion of users they fool. In an analysis of phishing attacks carried out in 2006, Rachna, Tygar, & Hearst [61] found that good phishing websites fooled 90% of participants. Existing anti-phishing browsing clues are ineffective and 23% of participants in the study did not look at the address bar, status bar, or the security indicators.

Perpetrators target both document categories to secure personal identifying information. Often they obtain a 'set' of point of information documents in order to present themselves as 'legitimate customers' to deceive the target organisation's authentication and verification processes to commit identity fraud [41]. Increasingly, the mode of attack for the fraud, especially the identity fraud perpetrator, is tending to rely on electronic commerce or mechanical/digital devices to initiate the identity theft or identity deception act. This is to some extent enabled by Internet adoption. For example, 77% of United States (US) adults were online in May 2006, up from 74% in 2005, 66% in 2002, 64% in 2001, and 57% in 2000, according to e-Marketer [31].

In phishing e-mail messages, the senders must gain the trust of the recipients to convince them to divulge their personal information. To gain this trust, fraudsters "spoof," or mimic, a reputable company. The companies spoofed most often are financial services- Internet organizations such as the Bank of America, Citibank, eBay, PayPal, etc. Retailers and Internet service providers are also targeted [7,8, 44].

These phishing e-mails are usually mass mailed (Warner, 2004). Many of the recipients are not customers of the spoofed companies and may quickly realize that the e-mail is fraudulent, or may believe that the e-mail was mistakenly sent to them and ignore the e-mail. Fraudsters rely on the responses from the few recipients who are customers of the spoofed company and who fall victim to the scam. According to Longe, Mbarika, Korouma, Wada, & Isabalija [49], the scammers claim to be from reputable companies and go to great lengths to emulate the company's visible branding.

Their fraudulent e-mails often contain the company's logo and use similar fonts and color schemes as those used on the company's web site. Some of the fraudulent e-mails simply reference images from the legitimate company's site. The main link in a fraudulent e-mail sends the recipient to the fraudulent phishing web site, but many fraudulent e-mails include other links that send the recipient to sections of the real company's web site.

To further convince the recipient that the e-mail originated from the reputable company, the scammers use a "from" e-mail address that appears to be from the company by using the company's domain name (e.g., @ebay.com, @paypal.com) [49]. Phishing e-mails also try to assure the recipient that the transaction is secure in hopes of gaining the recipient's trust. The following are assurances that were included in fraudulent e-mails:

"Remember: eBay will not ask you for sensitive personal information (such as your password, credit card, bank account numbers, social security number, etc.) in an e-mail."

This e-mail then sends users to a fraudulent web site that asks for personal and account information while promising that the information is submitted via a secure server. The phishing perpetrators could then notify the victim of a "security threat." Such a message may be welcomed or expected by the victim, who would then be easily induced into disclosing personal information [37]. The number of unique phishing websites detected by APWG during the second half of 2008 saw a constant increase from July to October with a high of 27,739 [7,8]).

In Nigeria, the most recent phishing attacks were on the customers of Inter-switch, which remains the organization with the highest customer base in electronic transactions. The Nigeria Deposit Insurance Corporation (NDIC) disclosed in its 2007 annual report and statement of account that underhand deals by bank staff, among others, resulted in attempted fraud cases totalling over N10.01 billion (over 65 million USD) and actual losses of N2.76 billion (13 million USD) in 2007 [3].

With the present situation in the world economy and the appropriate technology, fraudulent action is most likely to increase and phishing remains one of the main means of performing "fraud without borders." The extent of readiness to stem phishing in Nigeria needs to be determined because fraudulent activities emanating from these nations have far-reaching consequences beyond her borders.

2.1.2 Cyber Terrorism

According to the U.S. Federal Bureau of Investigation, cyber terrorism is any "premeditated, politically motivated attack against information, computer systems, computer programs, and data which results in violence against non-combatant targets by sub-national groups or clandestine agents" [67].

Unlike a nuisance virus or computer attack that result in a denial of service, a cyber terrorist attack is designed to cause physical violence or extreme financial harm. According to the U.S. Commission of Critical Infrastructure Protection, possible cyber terrorist targets include the banking industry, military installations, power plants, air traffic control centers, and water systems. Apart from that, there is another dimension to cyber terrorism – the use of cyber infrastructure to launder money for financing physical terrorism. In 2005, FBI officials reported that Al Qaeda terrorist cells in Spain used stolen credit card information to make numerous purchases [73].

According to Wilson (2008), cyber terrorism is said to have taken place when the effects of a widespread computer network attack is unpredictable and might cause enough economic disruption, fear, and civilian deaths, to qualify as terrorism. At least two views exist for defining the term cyber terrorism [20, 27]. These are (1) Cyber terrorism exists when computer attacks result in effects that are disruptive enough to generate fear comparable to a traditional act of terrorism, even if done by criminals. (2) Cyber terrorism exists when unlawful or politically motivated computer attacks are done to intimidate or coerce a government or people to further a political objective, or to cause grave harm or severe economic damage.

The terrorist's use of the Internet and other telecommunications devices is growing both in terms of reliance for supporting organizational activities and for gaining expertise to achieve operational goals. Tighter physical and border security may also encourage terrorists and extremists to try to use other types of weapons to attack the United States. Persistent Internet and computer security vulnerabilities, which have been widely publicized, may gradually encourage terrorists to continue to enhance their computer skills, or develop alliances with criminal organizations.

They will also probably consider attempting a cyber attack against the U.S. critical infrastructure [47]. Cybercrime has increased dramatically in past years, and several recent terrorists events appear to have been funded partially through online credit card fraud. Reports indicate that terrorists and extremists in the Middle East and South Asia may be increasingly collaborating with cybercriminals for the international movement of money and for the smuggling of arms and illegal drugs [27]. These links with hackers and cybercriminals may be examples of the terrorists' desire to continue to refine their computer skills, and the relationships forged through collaborative drug trafficking efforts may also provide terrorists with access to highly skilled computer programmers.

2.1.3 Electronic Spam Mails

These are unsolicited bulk e-mail to multiple recipients. They can be commercial, political, or religious. While the most widely recognized form of spam is e-mail spam, the term is applied to similar abuses in other media: instant messaging spam, web search engines, and blogs. Spamming is popular because the advertisers have no operating costs beyond the management of their mailing lists and it is difficult to hold senders accountable for their mass mailings. As a result, costs such as lost productivity and fraud are borne by the public and by Internet service providers that have been forced to add extra capacity to cope with the deluge [48].

A good example is 419 mails or the Nigerian advance fee frauds which in 1996 was estimated to have cost unsuspecting clientele over five billion dollars [70]. These mails emanate in a triangle called the "The Nigerian Connection" mostly in Europe and in some parts of Africa, "The 419 Coalition, 2005." The Nigerian Scam, according to published reports, is the third to fifth largest industry in Nigeria [79]

It is the 419 Coalition view that, in effect, the elites from which successive governments of Nigeria have been drawn are the scammers. Therefore, victims have little recourse in this matter. Monies stolen by 419 operations are almost never recovered from Nigeria. Most 419 letters and e-mails originate from or are traceable back to Nigeria. However, some originate from other nations, mostly also West African nations such as Ghana, Cameroon, Togo, Liberia, Sierra Leone, Ivory Coast (Cote D'Ivoire), etc.

The effects of such scams have immense effects with confirmed losses of millions of dollars annually(Herald Tribute, 2007). According to Longe and Longe [46], governments have tried to come up with policies to try to curtail this menace. Nigeria, through the EFCC, banned night browsing. This is because most fraudulent activities are perpetrated at cyber cafés at nights. For now, there are no quantitative data to measure the effect of this action on the reduction or otherwise of cybercrime in Nigeria. Apart from the availability and usage of Internet facilities in cyber cafes for pornography and other cybercrimes, the evolution of fixed wireless facilities in Nigeria, for instance, has added another dimension to the cybercrime problem. Nigeria therefore enjoys a dubious distinction of being the source of what is now generally referred to as '419' mails, named after Section 419 of the Nigerian Criminal Code (Capp 777 of 1990) that prohibits advance fee fraud.

These crimes are similar to theft and the likes that have existed for century's offline even before the development of high-tech equipment. Progress in the fight against Internet pornography has been moving at a very slow pace in Africa. A majority of public internet access point deals with the problem in unorthodox manners such as placing notices on cyber cafe walls warning against browsing pornographic sites and other spamming activities. Those with some technical expertise resort to the use of content filters which are downloaded and installed to filter unwanted Internet content [46].

2.1.4 Cyber Stalking

Stalking in the physical sense generally involves harassing or threatening behavior in which an individual engages repeatedly, such as following a person, appearing at a person's home or place of business, making harassing telephone calls, leaving written messages or objects, or vandalizing a person's property. According to Ellison and Akdeniz [30] cyber stalking refers to the use of the Internet, e-mail, or other electronic communications devices to stalk another person. This term is used interchangeably with online harassment and online abuse. A cyber stalker does not present a direct physical threat to a victim, but follows the victim's online activity to gather information and make threats or other forms of verbal intimidation.

The anonymity of online interaction, they argued, reduces the chance of identification and makes cyber stalking more common than physical stalking. Although cyber stalking might seem relatively harmless, it can cause victims psychological and emotional harm, and occasionally leads to actual stalking. Cyber stalking is becoming a common tactic in racism and other expressions of hate. Cyber stalkers target and harass their victims via websites, chat rooms, discussion forums, open publishing website (e.g., blogs) and e-mail. The availability of free e-mail and website space, as well as the anonymity provided by these chat rooms and forums, has contributed to the increase of cyber stalking as a form of harassment [30].

Most stalking laws require that the perpetrator make a credible threat of violence against the victim; others include threats against the victim's immediate family; and still others require only that the alleged stalker's course of conduct constitute an implied threat [74]. While some conduct involving annoying or menacing behavior might fall short of illegal stalking, such behavior may be a prelude to stalking and violence and should be treated seriously. The nature and extent of the cyber stalking problem is difficult to quantify. Indeed, current trends and evidence suggest that cyber stalking is a serious problem that will grow in scope and complexity as more people take advantage of the internet and other telecommunications technologies [17].

Important advances can only be made if industry, law enforcement, victims, service providers, support groups, and others work together to develop a more comprehensive and effective response to this problem. Ultimately, however, the first line of defense will involve industry efforts that educate and empower individuals to protect themselves against cyber stalking and other online threats, along with prompt reporting to law enforcement agencies trained and equipped to respond to cyber stalking.

Physical stalking, online harassment, and threats may be a prelude to more serious behavior, including physical violence. For example, the first U.S. cyber stalking law went into effect in 1999 in California. Other states include prohibition against cyber stalking in their harassment or stalking legislation. In Florida, HB 479 was introduced in 2003 to ban cyber stalking. This was signed into law in October 2003.

The crime of cyber stalking is defined in Florida Statutes 784.048(1) (d) which is one of the strictest such laws in the United States [69].

2.1.5 Fake Copy-Cat Web Sites

One recent trend in on-line fraud is the emergence of fake 'copy-cat' web sites that take advantage of consumers what are unfamiliar with the Internet or who do not know the exact web address of the legitimate company that they wish to visit. The consumer, believing that they are entering credit details in order to purchase goods from the intended company, is instead unwittingly entering details into a fraudster's personal database. The fraudster is then able to make use of this information at a later stage, either for his own purposes or to sell on to others interested in perpetrating credit card fraud [78].

3. EFFECTS OF CYBER CRIME ON BANKING

According to Reuter's media briefs from Cameroon [16], British prime minister, cyber crime costs the British economy some 27 billion pounds a year. On the other hand, the Economic and Financial Crimes Commission Report [28, 29] ranks Nigeria as third among the top ten sources of cyber crime in the world. It is estimated that after the United States with 65 per cent of cyber-criminal activities and the United Kingdom with 9.9 per cent, Nigeria is the next hub of cyber criminals in the world with 8 per cent. The growth of online banking further presents enhanced opportunities for perpetrators of cyber crime. Funds can be embezzled using wire transfer or account takeover. Criminals may submit fraudulent online applications for bank loans; disrupt e-commerce by engaging in denial of service attacks, and by compromising online banking payment systems [8]. Identity takeover can also affect online banking, as new accounts can be taken over by identity thieves, thus raising concerns regarding the safety and soundness of financial institutions.

Therefore unless crime detection and prevention are confronted collectively, Nigeria like any other country will remain warm breeding grounds for cartels of such criminal activity. A global effort to combat this crime is of essence. Financial fraud is one of America's largest growth industries, creating annual losses of $189 billion [39, 49]. The cost of application fraud alone, they argued, is more than $35 billion a year.

This is by far more damaging than delinquent or bankrupt accounts, fraud losses which are generally three times higher than normal charge-off rates. This situation poses a real and constant threat to profitability and may raise the price of goods and services for consumers. They further argued that by far, the greatest threats is from e-commerce fraud, identity theft and international criminal organizations, all of which are becoming more widespread and sophisticated every day.

As e-commerce continues to grow, it will become an even bigger attraction for criminals. The report indicated that identity theft is escalating at 40% a year and is particularly problematic compared with more traditional forms of financial fraud. Greater access to credit, an abundance of information, faster electronic communications, and intense competition among financial institutions make it easier than ever for perpetrators to steal identities and falsify information. The existence of cyber crime and its effects require the formulation of appropriate policies to address them. The next section presents existing policies on cyber- related crime in Nigeria.

4. CYBER CRIME POLICY IN NIGERIA

There is presently no law that is specific to cyber crime in Nigeria. However, this is not to say that cyber criminals are free to operate in the country. There are general laws that are not specifically related to cyber crime but are being enforced to deal with the crime. Some of these laws are: the Nigeria criminal code, Economic and Financial Crimes Commission (EFCC) (Establishment) Act 2004, and the Advance Fee Fraud and other Related Offences Act 2006 [32].

The Nigeria Criminal Code Act 1990

The Criminal Code Act of 1990 (Laws of the Federation of Nigeria, 1990) criminalizes any type of stealing of funds in whatever form, an offence punishable under the Act. Although cyber crime is not mentioned in the Act, it is a type of stealing punishable under the criminal code. The most renowned provision of the Act is Chapter 38, which deals with "obtaining Property by false pretences-Cheating." The specific provisions relating to cyber crime is section 419, while section 418 gave a definition of what constitutes an offence under the Act.

(418) "Any representation made by words, writing, or conduct, of a matter of fact, either past or present, which representation is false in fact, and which the person making it knows to be false or does not believe to be true, is a false pretence."

(419)"Any person who by any false pretence, and with intent to defraud, obtains from any other person anything capable of being stolen, or induces any other person to deliver to any person anything capable of being stolen, is guilty of a felony, and is liable to imprisonment for three years." (Part 6, chapters 34 &38, Laws of the Federation of Nigeria Act, 1990)

The Economic and Financial Crime Commission Act, 2004
(Source: National Assembly of Nigeria, 2004)

The Economic and Financial Crime Commission Act (Laws of the Federation of Nigeria, 2004, as amended) provide the legal framework for the establishment of the Commission. Some of the major responsibilities of the Commission, according to part 2 of the Act, include:

- the investigation of all financial crimes, including advance fee fraud, money laundering, counterfeiting, illegal charge transfers, futures market fraud, fraudulent encashment of negotiable instruments, computer credit card fraud, contract scam, etc.;
- the coordination and enforcement of all laws against economic and financial crimes laws and enforcement functions conferred on any other person or authority;
- the examination and investigation *of* al1 reported cases *of* economic and financial crimes with a view to identifying individuals, corporate bodies, or groups involved;
- undertaking research and similar works with a view to determining the manifestation, extent, magnitude, and effects *of* economic and financial crimes and advising government on appropriate intervention measures for combating same;
- Taking charge of, supervising, controlling, coordinating all the responsibilities, functions, and activities relating to the current investigation and prosecution of all offences connected with or relating to economic and financial crimes, in consultation with the Attorney-General *of* the Federation;
- the coordination of all investigating units for existing economic and financial crimes, in Nigeria;
- The Commission is further charged with the responsibility of enforcing the provisions of
  the Money Laundering Act 1995 ; the Advance Fee Fraud and Other Fraud- Related Offences Act 1995 ; the Failed Banks (Recovery of Debts) and Financial Malpractices in Banks Act 1994, as amended; the Banks and other Financial Institutions Act 1991, as amended; and Miscellaneous Offences Act (EFCC, 2004) .

Advance Fee Fraud and Related Offences Act 2006 (Source: National Assembly of Nigeria, 2006)

According to Section 23 of the advance fee fraud Act (Laws of the Federation of Nigeria, 2006): 'False pretence means a representation, whether deliberate or reckless, made by word, in writing or by conduct, of a matter of fact or law, either past or present, which representation is false in fact or law, and which the person making it knows to be false or does not believe to be true.' Section 383 sub-section 1 of the Nigerian Criminal Code states: 'A person who fraudulently takes anything capable of being stolen, or fraudulently converts to his own use or to the use of any other person anything capable of being stolen, is said to steal that thing.' (Advance Fee Fraud Act, Laws of the Federation of Nigeria, 2006) [1]

Economic crime is defined by the Act as " the non-violent criminal and illicit activity committed with the objectives of earning wealth illegally, either individually or in a group or organized manner thereby violating existing legislation governing the economic activities of government and its administration to include any form of fraud, narcotic drug trafficking, money laundering, embezzlement, bribery, looting, and any form of corrupt malpractices, illegal arms deal, smuggling, human trafficking and child labor, oil bunkering and illegal mining, tax evasion, foreign exchange malpractices including counterfeiting of currency, theft of intellectual property and policy, open market abuse, dumping of toxic wastes and prohibited goods."

Advance Fee Fraud and Other Fraud Related Offences Act 2006 is currently the only law in Nigeria that deals with internet crime issues, and it only covers the regulation of internet service providers and cybercafés, it does not deal with the broad spectrum of computer misuse and cyber crimes [32]. There are presently six bills on cyber crime being considered by the National Assembly (legislative arm) of Nigeria. These are: the Computer Security and Critical Information Infrastructure Protection Bill 2005 (sponsored by the Executive), the Cyber Security and Data Protection Agency (Establishment, etc.) Bill 2008 (sponsored by Hon. Bassey Etim), the Electronic Fraud Prohibition Bill 2008 (sponsored by Senator Ayo Arise), the Nigeria Computer Security and Protection Agency Bill 2009 (another executive bill), the Computer Misuse Bill 2009 (sponsored by Senator Wilson Ake) and the Economic and Financial Crimes Commission Act (Amendment) Bill 2010, sponsored by Hon. Abubakar Shehu Bunu.

5. THEORETICAL UNDERPINNINGS

This section discusses some theories relating to the electronic media and security issues. Electronic media have been emphasized by various theoretical traditions. Sociologists, for instance, argued that point-to-point communication media- for instance, telephones- support shared aims which demonstrate a powerful collective representation. Some, especially the Marxists, look at communication media as an exploitative tool by the elitist groups for socioeconomic and political control 25]. In their own contribution to the digital communication.

Bell, Garland and Platt [14] argued that the invention of mini-electronic and optical circuits capable of speeding the rate of information flow through networks would have a big impact on society. Despite the positive impact of technology on society, it has on the other hand led to the unintended use in criminal activities like cybercrime. He concluded by saying, it is easier to steal a penny from millions of bank account owners using the internet than using physical robbery.

5.1 Routine Activity Theory

This theory proposes that three situations facilitate the occurrence of crime. Proponents argue that such events must happen at the same time and in the same space. The three situations are the existence of a suitable target, lack of security, and a motivated offender for the crime to occur [22]. The assessment of the situation determines whether or not a crime takes place.

5.2 Opportunity Theory

This theory does not focus on the events that contribute to the crime but on the opportunities that emerge as a result of preventive measures to curb the crime. Proponents of this theory argue that crimes transverse between location, time, target, direction, and method of committing the crime [33]. They further assert that Opportunity to commit a crime is a root cause of crime. Also, they posit that no crime can occur without the physical opportunity and therefore opportunity plays a role in all crimes, not just those involving physical property thereby reducing opportunity of crime.

5.3 Technology Theory

The response of technology to the cyber crime problems centre on the use of computer security theories to design and evolve solutions that provides authentication, verification, non repudiation and validation. These theories and models rely on the use of cryptography, steganography, network protocols, and the use of software engineering process/models to develop systems that offer some form of protection for users and the information infrastructure. Cybercrime thrives on the web today because the internet did not inculcate in its protocols from the onset a mechanism that allows a host to selectively refuse messages [24].

This implication is that a benign host that desires to receive some particular messages must read all messages addressed to it. In essence, a malfunctioning or malicious host has the capacity to send many unwanted messages. This problem is exacerbated by the ubiquitous nature of the web and remains the Achilles heel of the issue of web security today. Although all the theories discussed above are related to cyber crime, we are inclined to adapt routine activity theory to this study because the theory captured the philosophical assumptions upon which this study is based.

5.4 Social Theories

From a social scientific point of view, security theories on providing and implementing protection against breaches and information system misuse have evolved. They focus on user security awareness, motivation, deterrents, technology and training [39, 36, 11, 23, 59, 60]. Researchers have theorized that user perception of risks and choices based on those perceptions can influence system security [9]. The situational characteristics theory proponents argued that situations within a system usage domain can impact on ethics and user behaviour [58, 52, 12]. Wood [76, 77] proposed the Human Firewall theory stating that those user actions can undo technical security measures. He advocated that organizations must sensitize and educate users and evaluate their compliance with security policies and procedures.

The *theory of least possible privilege* as proposed by Beatson [13] suggests psychological profiling of potential new users, while Bray [15] argues that new users are more vulnerable to security breaches when using information systems (IS). Denning [27] theorizes about *defensive information warfare* and proposes that security policy training and awareness will better equip users against threats. Forcht, Pierson and Bauman [35] *theorized about ethical awareness and culture* as factors that influence IT security. Kabay (2002) theorized about using *social psychology* as a tool to improve user security conduct. The importance of the interest of senior management and integrating security issues as part of the *corporate asset protection model* was highlighted by Katsikas [40], Kovacich and Halibozek [42] and Perry [58]. Vroom and von Solms [75] also modelled an Information System security awareness program to address end-users, IT personnel and management executives.

McLean [53] theorized about using *values, perceptions and behaviour* to change user attitude about security, while Murray [54] argues that ignorance and incompetence about the consequence of security policy abuse is a serious problem among users. Parker [57] proposed a theory that uses *rewards and penalties* to influence attitudes toward security in information systems.

Sasse, Brostoff, and Weirich [63] theorized that the *nature of the technology* with respect to the user's goals and intentions significantly influence security features and usage in IS systems. They went further to propose the use of training, punishment, and reporting security as a motivation for creating security awareness among users. Schlienger and Teufel [64] adopted a *socio cultural approach* to information security and posited that the *cultural theory* can be used to enhance security at different cultural layers-namely, corporate policies, top management, and individuals. Siponen [66] used *human morality* as a force that can impact on security. Straub and Welke [72] theorized about using *strong deterrents* to convince potential violators of those organizations means in business about protecting information infrastructures. Tudor [[73] argued for a theory that uses a *holistic IS security architecture* to incorporate infrastructure, policies, standards, awareness and compliance. He however, concentrated on awareness training at the expense of all the other components.

5.5 The Peel Theory

The *Peel theory of community policing* as highlighted by Longe et al [50] assumes that violators or criminals and victims are usually proximate and used spatial distribution as a basis for apprehending criminals and solving crimes. This theory subsumes the role of the citizens in responding to partial and completed crime, crime control, and internal order and makes the police responsible for all crime control and law enforcement activities. Although some consensus exists among nations on how to combat and deal with crimes across borders using international policing such as the Interpol, the underlying theory still relates to the Peel model and it is therefore inadequate to face the cyber crime problem. We cannot say as a matter of fact that there is any theory in existence from the criminal justice and policing angle that specifically addresses the problem of cyber crime.

5.6 Space Transition Theory:

Proponents of space transition theory argue that behavior of people in cyber space tends to bring out their compliance and noncompliance behavior both in the physical and in cyber space. This theory does not explain physical crime but cyber crime and how people move and behave from one space to the other (Schmalleger & Pittaro, 2009). This entails persons with repressed criminal behavior (in the physical space) having a propensity to commit crime in cyberspace, which they would not otherwise commit in physical space, due to their status and position. It also implies that the status of persons in physical space does not transit to cyber space. Jaishankar [37], for instance, argues that the individual behavior repressed in physical space is not repressed in cyber space.

6. CONCLUDING REMARKS

This paper examined the impact of the information communication technology (ICT) revolution on business, industry and government in the light of the unintended consequences such as criminal activities, spamming, credit card frauds, ATM frauds, phishing, identity theft and other related cyber crimes. We specifically assessed cyber crime and its impact on the banking institutions in Nigeria. Existing policy framework were examined and their success as institutional countermeasures in combating cyber crime in the banking industry were assessed.

Finally, we provide insight into how cybercrime impacts on E-banking from a Nigerian perspective using social theories to explain causation with a view of guiding policy makers on behavioural issues that should be considered when formulating policies to address cyber criminal activities in Nigeria.

REFERENCES

1. Advance Fee Fraud and Other Fraud Related Offences Act 2006, Laws of the Federation of Nigeria
2. Agboola, A. A. (2006). Electronic Payment Systems and Tele-banking Services in Nigeria, Journal of Internet Banking and Commerce, Vol. 11, No. 3, online source: http://www.arraydev.com/commerce/jibc
3. Ahmed, I.(2008) Nigeria: N10 Billion Lost to Bank Fraud in 2007 – NDIC, Daily Trust, 28 October 2008
4. Anguelov, C. E. et al.(2004). U.S. Consumers and Electronic Banking,1995-2003. Federal Reserve Bulletin
5. Alexander, H., (1988) National collective action and economic performance: International Studies Quarterly,
6. APWG (Anti-Phishing Working Group) (2004). Phishing Activity Trends Report. Available: http://www.antiphishing.org
7. APWG (Anti-Phishing Working Group) (2008). Phishing Activity Trends Report. Available: http://www.antiphishing.org
8. Atherton, M. (2010) Criminals switch attention from cheques and plastic to internet transactions. The Sunday Times of March 10, 2010
9. Aytes, Computer Security and Risky Computing Practices: A Rational Choice Perspective
10. Balachandran and Balachandher K. G. (2000), "E-Banking Development in Malaysia: Prospects and Problems", 10 JIBL, 250.
11. Barman S (2002) Writing IS security Policies. New Riders Publishing, Indianapolis
12. Banerjee D, Cronan TP & Jones TW (1998) Modeling IT Ethics: A Study in Situational Ethics.MIS Quarterly 22(1): 31-60.
13. Beatson JG (1991) Security - a personnel issue. The importance of personnel attitudes and securityeducation. Proceedings of the Sixth IFIP International Conference on ComputerSecurity.
14. Bell, R,. Garland, & Platt, R.B (1997) Bridging and signalling subsystems and methods for private and hybrid
15. Bray TJ (2002) Security actions during reduction in workforce efforts: what to do when downsizing. Information system security 11(1): 11-15.
16. Cameroon, D. (2011) Cyber crime costs UK 27 Billion pounds, Reuters media briefs.

17. CCIPS, (1999) Cyber stalking: A New Challenge for Law Enforcement and Industry. Workshop on the economics of information security
18. Centeno, C. (2003). Adoption of Internet Services in the Enlarged European Union: Lessons from the Internet Banking Case. European Commission Joint Research Center Report http://fiste.jrc.es/download/eur20822en.pdf
19. Chiemeke, S. C., Evwiekpaefe, A. and Chete, F.(2006) The Adoption of Internet Banking in Nigeria: An Empirical Investigation, Journal of Internet Banking and Commerce, Vol. 11, No.3,
20. Collin, B.C.(1996) The Future of Cyber Terrorism: Where the Physical and Virtual Worlds Converge, 11th Annul International Symposium on Criminal Justice Issues
21. Criminal Code Act Chapter 77, Laws of the Federation of Nigeria 1990.
22. Cohen, L. and Felson M. (1979)« Social Change and Crime Rate Trends : A Routine Activity Approach », American Sociological Review, 44 (4), 1979, pp. 588-608
23. Cox A, Connolly S & Currall J (2001) Raising IS security awareness in the academic setting. VINE, Issue 123: 11-16
24. Crocker, D.(1982) Standard for the format of ARPA Internet text messages
25. Davis, R.& Hutchison S. (1997) Computer Crime in Canada, Toronto: Thompson Canada Limited.
26. Desman, M. (2002). Building an information security awareness program, Auerbach Pub
27. Denning DE (1999) Information Warfare and Security. ACM Press, USA
28. EFCC/ NBS/ (2009) Business Survey on Crime & Corruption and Awareness of EFCC in Nigeria, Summary Report
29. EFCC/ NBS/ (2010) Business Survey on Crime & Corruption and Awareness of EFCC in Nigeria, Summary Report
30. Ellison, L., & Akdeniz, Y.,(1998) "Cyber-stalking: the Regulation of Harassment on the Internet," [1998] Criminal Law Review, December Special Edition: Crime, Criminal Justice and the Internet, pp 29-48.
31. E-Marketer(2006) estimate and projections. http://www.emarketer.com/docs/emar_whitepaper.pdf
32. Ewelukwa, N.(2011). This Day Newspaper, Nigeria, March 31.2011
33. Felson, M. & Clarke, R. V. (1998). Opportunity Makes the Thief. Police Research Series, Paper 98. Policing and Reducing Crime Unit, Research, Development and Statistics Directorate.. London: Home Office. [www.homeoffice.gov.uk/rds/prgpdfs/fprs98.pdf]
34. Finextra(2005) Deutsche Bank signs global license for TT dealing system www.finextra.com/news/fullstory.aspx?newsitemid=14314
35. Forcht KA, Pierson JK & Bauman BM (1988), Developing awareness of computer ethics. Proceedings of the ACM SIGCPR conference on management of information systems personnel: 142-143.
36. Gaunt, N.(1998), Installing an appropriate IS security policy in hospitals. International Journal of Medical Informatics, 131-134.
37. Jaishankar K., (2008), Space Transition Theory of Cyber Crimes, Crimes of the Internet, Pearson, ISBN-13:978-0-13-231886-0 pp.283-299
38. Kabay ME (2002) Using Social Psychology to Implement Security Policies. In: Bosworth S & Kabay ME (eds) Computer Security Handbook, 4th edition. John Wiley & Sons, Inc., USA,32.1-32.16.
39. Kajava, J Siponen, M.T. (1997): Effectively Implemented Information Security Awareness - An Example from University Environment. Proceedings of IFIP-TC 11 (Sec'97/WG 11.1). 13th International Conference on Information Security: Information Security Management - The Future. 13th May 1997, Copenhagen, Denmark.
40. Katsikas, S.K. (2000) Health care management and information system security: awareness, training or education?. International Journal of Medical Informatics 60(2): 129-135.
41. Kochems, A. & Keith, L.(2006) Successfully Securing Identity Documents: A Primer on Preventive Technologies and ID Theft
42. Kovacich GL & Halibozek EP (2003) The Manager's Handbook for Corporate Security: Establishing and Managing a Successful Assets Protection Program. Butterworth- Heinemann,USA.

43. Liao Z. & Wong W. K., (2008).The Determinants of Customer Interactions with Internet-Enabled e-Banking Services. The Journal of the Operational Research Society, Vol. 59, No. 9 (Sep., 2008), pp. 1201-1210.
44. Litan, A. (2004). Phising attack victims likely targets for identity theft. Available: http://www.gartner.com/DisplayDocument?doc_cd=120804
45. Loftness, S. (2004) Responding to "Phishing" Attacks. Glenbrook Partners
46. Longe, O.B & Longe, F.A. (2005). The Nigerian Web Content: Combating the Pornographic Malaise Using Web Filters. Journal of Information Technology Impact. Vol. 5, No. 2 Loyola University, United States of America. www.jiti.net
47. Longe, O.B & Chiemeke, S.C.(2007). Beyond Web Intermediaries: Framework for Protecting Web Contents on Clients Systems. Paper Presented at the International Conference of the International Association of Engineers (IAENG) Imperial
48. Longe, O.B.& Chiemeke, S.C. (2008): Cybercrime and Criminality in Nigeria-What roles are internet Access Points in Playing. European Journal of Social Sciences, Volume 6 No 4.
49. Longe,O.B, Mbarika, V, Kourouma, M, Wada, F & Isabalija, R. (2009). Seeing Beyond the Surface: Understanding and Tracking Fraudulent Cyber Activities. International Journal of Computer Science and Information Security. Vol. 6 (3) (pp. 124-135
50. Longe,o.B., Wada, F. , Anadi, A, Jones, C. & Mbarika, V (2010). A Critical Appraisal of the Peel Theory of Community Policing in the Age of Cybercrime. 84th Annual Meeting Louisiana State University at Alexandria. http://www.laacademy.org/docs/Schedule.pdf
51. Lorin M. H. & Frei,F., X. (2002) Do Better Customers Utilize Electronic Distribution Channels? The Case of PC Banking. Management Science, Vol. 48, No. 6 (Jun., 2002), pp. 732-748.
52. Martins, A. Eloff, J.H.P(2002): Information Security Culture. SEC 2002: 203-214
53. McLean K (1992) IS security awareness - selling the cause. Proceedings of the IFIP TC11, 8th International Conference on IS security, IFIP/Sec '92.
54. Murray, B. (1991). Running corporate and national security awareness programmes.
55. Nitsure, R. R. (2003) E-Banking: Challenges and Opportunities. Economic and Political Weekly, Vol. 38, No. 51/52 pp. 5377-5381
56. Olasanmi, O. O (2010). Computer Crimes and Counter Measures in the Nigerian Banking Sector. Journal of Internet Banking & Commerce, 15(1), 1-10 (http://www.arraydev.com/commerce/jibc/)
57. Parker DB (1999) Security motivation, the mother of all controls, must precede awareness. Computer Security Journal 15(4): 15-23.
58. Perry, W. (1985). Management strategies for computer security, Butterworth-Heinemann Newton, MA, USA.
59. Pipkin DL (2000) IS security: Protecting the Global Enterprise, Hewlett-Packard Professional Books. Prentice Hall PTR, Upper Saddle River, USA.
60. Proctor PE & Byrnes FC (2002) The Secured Enterprise: Protecting Your Information Assets. Prentice Hall, Upper Saddle River, USA
61. Rachna D., Tygar, J.& Hearst, M. (2006): "Why Phishing Works" in the Proceedings of the Conference on Human Factors in Computing Systems (CHI2006)
62. Roger, E.S. (2008) Rogers Communications Inc, 2008 Annual Report
63. Sasse A, Brostoff S & Weirich D (2001) Transforming the 'weakest link' a human computer interaction approach to usable and effective security. BT technology Journal 19(3): 122- 131.
64. Schlienger T & Teufel S (2002) IS security Culture: The Socio-Cultural Dimension in IS security Management. Proceedings of IFIP TC 11.
65. Schmalleger, F. & Pittaro M. (2009) Crimes of the Internet, Pearson Prentice hall.
66. Siponen MT (2000) On the Role of Human Morality in Information System Security: The Problems of Descriptivism and Non-descriptive Foundations. Proceedings of IS security for Global Information Infrastructures, IFIP TC11 Fifteenth Annual Working Conference on IS security: 401-410.

67. Search security (2009), Information Security magazine
68. Shandilya, A. (2011) Online Banking: Security Issues for Online payment Services. www.buzzle.com/articles.
69. Smith, M. (2004) Cyberstalking and the Law
70. Smith, R.G., Holmes, M.N. & Kaufmann, P. (1999): Nigerian advance fee fraud. Trends and Issues in Crime and Criminal Justice, No. 121. Australian Institute of Criminology, Canberra.Available online t http://www.aic.gov.au
71. Steiner,T .,D . Teixeira.1990.Technologyin Banking. Irwin,New York
72. Straub, D. and R. Welke (1998). "Coping with systems risk: security planning models for management decision making." MIS Quarterly 22(4): 441-469.
73. Tudor JK (2001) IS security Architecture, An Integrated Approach to Security in the Organization. Auerbach Publications, USA
74. US Attorney General (1999) Cyberstalking: A New Challenge for Law Enforcement and Industry.A Report from the Attorney General to the Vice President August 1999
75. Vroom, C. and R. v. Solms (2002). A Practical Approach to Information Security Awareness in the Organization. Proceedings of the IFIP TC11 17th International Conference on Information Security: Visions and Perspectives, Kluwer, B.V.: 19- 38
76. Wood CC (1995) IS security awareness raising methods. Computer Fraud & Security Bulletin (June): 13-15.
77. Wood CC (2002) The Human Firewall Manifesto. Computer Security Journal 18(1): 15- 18.
78. www.bbc.co.uk