IMAGE: A MAP OF THE STARS OF THE ORION CONSTELLATION

Print ISSN: 2514-863X Online ISSN: 2514-8648

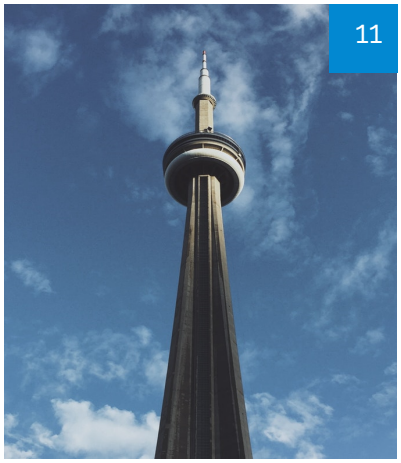
JournalPreview

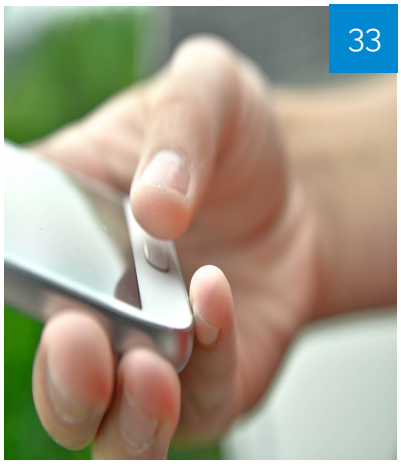
London Journal of Research in Computer Science and Technology
Volume 18 | Issue 1 | Compilation 1.0

London Journals Press

Journal Content
In this Issue

London Journals Press





Risk Assessment: Harnessing Positive Risks in ICT Systems

Nwagu, Chikezie Kenneth, Omankwu, Obinnaya Chinecherem & Okonkwo, Obikwelu Raphael

Nnamdi Azikiwe University

ABSTRACT

Every Organization continually assesses her network – data and communication devices of risks, to ensure risks are proactively averted. And in situations where they cannot be avoided, mitigation plans are put in place to ensure very minimal business impact if triggered. This research work is a product of practical investigations of ICT systems with the immediate focus on data and communication devices. This unravels direct utilization/application of positive risks and further emphasizes that risks are not entirely negative. It also ensures protection against negative risks and conscious activation and utilization of positive risks. Risk Assessment is every stakeholder`s responsibility and should be carried out as many times as possible. This assists stakeholders to proactively monitor their data and communication devices against invasion and unforeseen vulnerabilities. The general and long standing perception of risks is that risks and their impacts are entirely negative.

Keywords: positive risks(opportunities), negative risks(threats), risk assessment; vulnerabilities, virtualization, cloud computing.

Classification: H.0

Language: English

London Journals Press

LJP Copyright ID: 975711
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology

Volume 18 | Issue 1 | Compilation 1.0

Risk Assessment: Harnessing Positive Risks in ICT Systems

Nwagu, Chikezie Kenneth[α], Omankwu, Obinnaya Chinecherem[σ] & Okonkwo, Obikwelu Raphael[ρ]

I. ABSTRACT

Every Organization continually assesses her network – data and communication devices of risks, to ensure risks are proactively averted. And in situations where they cannot be avoided, mitigation plans are put in place to ensure very minimal business impact if triggered. This research work is a product of practical investigations of ICT systems with the immediate focus on data and communication devices. This unravels direct utilization/application of positive risks and further emphasizes that risks are not entirely negative. It also ensures protection against negative risks and conscious activation and utilization of positive risks. Risk Assessment is every stakeholder`s responsibility and should be carried out as many times as possible. This assists stakeholders to proactively monitor their data and communication devices against invasion and unforeseen vulnerabilities. The general and long standing perception of risks is that risks and their impacts are entirely negative. This further bares salient opportunities (positive Risks) associated with data and communication devices. And examined their inherent dormant features from direct industry and practical application points of view, which could be activated to further enhance their usage and maximize their benefits to organizations and stakeholders. Specific positive risks associated with specific data and communication devices were identified and their applications/ utilizations were also discussed The Methodologies used were highlighted and risk assessment data gathering techniques explicitly x-rayed. Strategies for mitigating negative risks and harnessing positive risk were explicitly discussed too. Of course, principal advantages of positive risks are enumerated. Finally showed how positive risks have immensely contributed to changing trend of computing today, contributed immensely to advent, deployment and use of Cloud computing.

Keywords: positive risks(opportunities), negative risks(threats), risk assessment; vulnerabilities, virtualization, cloud computing.

Author α ρ : Computer Science Department, Nnamdi Azikiwe University Awka Anambra State, Nigeria.
σ : Computer Science Department, Michael Okpara University of Agriculture Umudike Umuahia, Abia State, Nigeria.

II. METHODOLOGY

This work explored and used Structured System Analysis and Design Method (SSADM), Dynamic System Development Method (DSDM) and Spiral Methodologies.

Reasons for the adoption are their direct applicability and features among which are respectively as follow:

For SSADM,

- Intensive users involvement
- Clear and easily understandable documentation
- Procedural Process

For DSDM

- Focuses on Business need and delivers on time
- Communicate continuously and clearly without compromising quality
- For Spiral

London Journal of Research in Computer Science and Technology

- Risk driven and keeps track of risk patterns in a project.
- Iterative and incremental

III. RISKS AND RISK ASSESSMENT

Risks, contrary to the general notion, are both positive and negative. Therefore Risk, as adapted from Stoneburner, G., Goguen, A. & Feringa, A. (2002, July), is net negative or positive effect of exercise of vulnerabilities or opportunities which can be exploited, enhanced, shared, transferred, or even accepted.

Risk Assessment is a continuous process of identifying, analyzing, prioritizing and evaluating risks. For negative risks, it is done along with thorough evaluation of available controls with intention of recommending more robust and effective controls or enhancing existing controls in order to reduce or eliminate vulnerabilities. Vulnerabilities if exercised, be it intentionally exploited or accidentally triggered could lead to loss of integrity, availability and confidentiality. Hence one of the principal reasons for carrying out risk assessment against negative risks is to assess and ascertain the degree of potency and resilience of the available controls in order to make appropriate control recommendations. This further ensures full protection of the organizations` huge investments and raise awareness of risks trends. It practically assists organizations to be proactively and strategically positioned against any unforeseen threats. For Positive risks, the aim is targeted at maximizing and optimizing the use of additional features of the IT systems to the organization and stakeholders. Therefore, Risk assessment assists in discovering unique features (positive risks) of IT systems which might have been dormant or underutilized.

Good example is in the deployment of intelligent Ethernet switches to work as OSI layer 3 devices in addition to their primary known OSI layer 2 function. This is further explored in section IV according to the research.

IV. RISK ASSESSMENT DATA GATHERING TECHNIQUES

Risk assessment data gathering happens principally in the initial steps of risk assessment process of ICT systems and throughout their life span – before deployment, during deployment and after deployment. However, it is advisable to initiate risk assessment as soon as when the need for the ICT system is established.

The techniques used are:

- Prompt List: This is a predetermined list of risk categories that might give rise to individual IT devices risks type (negative or positive). This was used as a framework which aided in idea generation during risks identification.
- Assumption and Constraint Analysis: Every ICT system is conceived and developed based on a set of assumptions and within a series of constraints. Assumptions and constraint analysis is used to explore the validity of the assumptions and constraints to determine which constitute a risk to ensure full utilization of the systems. Vulnerability maybe identified from the inaccuracy, inconsistency or incompleteness of the assumptions. Constraints may give rise to opportunities (positive risks) through removing or relaxing the limiting factor in the design of the systems as detected in the intelligent Ethernet switches such as cisco intelligent catalyst Ethernet switch 3550, 3580, 4948 etc.
- Root Cause Analysis: This was basically used to discover the underlying causes that lead to problem statement and develop preventive action. It was, therefore, used to identify threats and vulnerabilities with a clear problem statement and of course, exploring which threat or vulnerability might result which made that problem occurred. It was also used to find opportunities by starting with a benefit statement and exploring which opportunities might result in that benefit being realized such as in the virtualization

technology (VT) which is used in the popular cloud computing, high availability etc.

- SWOT Analysis: This examined data and communication devices from strengths, weaknesses, opportunities and threats (SWOT) perspectives. SWOT analysis identifies any opportunities in the devices that may be utilized as the strengths and any threats resulting from weaknesses that may be avoided or reduced. The analysis was also used to examine the degree to which the strengths may offset the threats and determines if weaknesses might hinder opportunities.
- Brainstorming: This was used to obtain a comprehensive list of IT devices risks. IT teams in my professional networks and fora were engaged and they performed brainstorming, often with a multidisciplinary set of experts who are not part of the team. As the facilitator, Ideas about ICT devices risk were generated in a traditional free-form brainstorm session. Categories of risk, such as in a risk breakdown structure (RBS), from high level to finer risks level; for example ,type of risks, likely sources, probability of occurrence, motivation etc.
- Delphi Technique: The Delphi technique, as a technique that seeks means of reaching subject Matter Experts (SME) consensus, was used too. And as the facilitator, a simple questionnaire was used to solicit ideas about the important risks – main opportunities and vulnerabilities. The responses from round 1 were summarized and were then recirculated to the experts for further final comment to reinforce earlier responses based on the questions. Consensus was reached in a few rounds of this process. The Delphi technique helped reduce bias in the data and prevented any person from having undue influence on the outcome as responses were submitted anonymously. See result output as attached Appendix A
- Expert Judgement: Risks identified were further validated directly by consulting experts with relevant experience of similar projects or business areas. Such experts were identified and invited via online fora and considered all aspects of ICT devices and suggested possible risks based on their previous experience and areas of expertise. The experts` biases were also taken into account and OEMs` websites/portals and device documentations were checked for confirmation and updates
- Documents Analysis: Risks were identified from structured review of systems/devices documents (technical, administrative etc.) Uncertainty or ambiguity in the documents as well as inconsistencies within a document or between different documents were indicators of risks which propelled further investigations to ascertain clarity.
- Checklist Analysis: Risk identification checklists were developed based on historical information and knowledge that has been accumulated from previous similar systems and from other sources of information. The lowest level of the RBS was used as a risk checklist. These made clear some common IT System vulnerabilities encounter in live environment. However it is advisable that the checklist be prune from time to time to remove or archive related or outdated items. In fact the exercise should incorporate new lessons learned and improve it for use in future IT systems.
- On-site Interviews. Oral Interviews with IT system support and management personnel. (Case study: Mantrac Nigeria Limited – Caterpillar Nigeria) were conducted in order to collect useful information about the IT system (e.g., how the system is operated and managed). On-site visits also allowed for direct observation and gathering of information about the physical, environmental, and operational security of the IT devices. For devices in the design phase, on- site visit which is face-to-face data gathering exercises was used. And it provided the opportunity to evaluate the physical environment in which devices will operate in.

- Use of Automated Scanning Tool. The use of different tools for different platforms (windows, Linux and other open-source OSs) for detecting IT systems vulnerabilities and other Proactive technical methods were used to collect system information efficiently. For example, Software such as MBSA(Microsoft Baseline Security Analyzer), Advisor etc. were used to identify the services that run on a large group of hosts and provided a quick way of building individual profiles of the target IT device(s) which immensely aided in gathering common security vulnerabilities on a stand-alone and networked workstations

V. HARNESSING POSITIVE RISKS

Risks are not entirely negatives as general notion tends to hold. It could, therefore, be positive which opportunity is. And as opportunity, it can either be enhanced, exploited, shared or accepted. These bare most of the latent capabilities of IT devices in addition to their known primary roles. Therefore, these latent capabilities can be harnessed and fully utilized for maximum benefit of these devices. These further offer assurance and basis for justification for the huge capital investment on the devices. Hence Management and organizations could cut down on excess costs associated with purchasing devices which can be effectively and efficiently substituted with devices whose primary roles are needed and also have the required positive risks (opportunities) which are core roles and features of the devices which are being substituted.

An outstanding example is in the use of intelligent Ethernet switches such as Cisco catalyst 3550, 3580 and 4948 intelligent Ethernet switches. The earlier known assertion of routers as layer 3(network layer) device are now the strengths and opportunities (positive risks) which these switches offer.

Layer 3 switching include Layer 3 routing capabilities. Many of the current-generation Catalyst Layer 3 switches can use routing protocols such as BGP, RIP, OSPF, and EIGRP to make optimal forwarding decisions. With these, the switches operate in layer 3 in addition to layer 2(data-link layer) of the OSI model. There is evidence also that cisco catalyst 4948 10 Gigabit Ethernet switch operates in layers 2, 3 and 4 of the OSI layer.

Hence the principal network services offered by the switches, which used to be core functions of routers and routers only, are:

Security - Access Control List (ACL) which is use of Access Control Entries (ACE) to identify and grant access to trustees or bona fide entities seeking access.

The switch Port ACLs and that of the router shared common features. There seems few differences and /or advantages which are mainly due to its occurrence in layer 2.

The switch Port ACLs are similar to Router ACLs.But as you may know, they are supported on physical interfaces and configured on Layer 2 interfaces on a switch. All the three access list types are configurable on the switches such as standard, extended and MAC-extended. Unlike routers, Switch port ACL supports only inbound traffic filtering.

Processing of the Port ACL is similar to that of the Router. The switch examines ACLs associated with features configured on a given interface and permits or denies packet forwarding based on packet-matching criteria in the ACL.

If applied to a trunk port, the ACL filters traffic on all VLANs present on the trunk port. When applied to a port with voice VLAN, the ACL filters traffic on both data and voice VLANs.

The main benefit with Port ACL is that it can filter IP traffic (using IP access lists) and non-IP traffic (using MAC access list). Both types of filtering can be achieved.That is, a Layer 2 interface can have both an IP access list and a MAC access list applied to it at the same time.

Rate-limiting is basically used to limit and control rate of traffics sent or received over a network

interface. This offers protection against Denial-of-Service (DoS) attack. This actually limits upload speed on LAN ports and download speed on WAN port.

Advanced Quality of Service (QoS) which enables packets to be queued, classified, prioritized and policed to ensure packets delivery optimization and efficiency. In addition, congestion is avoided by all means possible. The good news here is that configuration of QoS is so simplified through automatic QoS (auto QoS) which is also a feature that detects devices mainly the IP phones and automatically configures the switch for appropriate packet classification and queuing.

Security Enhancement: Using the switches as router does not in any way trade-off security as there are many security features that make them fit-for-purpose – protection of network and administration of traffics, prevention of unauthorized users and granting of granular access to network and its tracking.

The security features are:

Secure Shell (SSH) - as a protocol provides a secure remote access connection to network devices and Communication between the client and server. It is encrypted using enhanced security algorithm in both SSH version 1 and SSH version 2.

Kerberos, according to Cisco press, is a secret-key network authentication protocol. It uses the Data Encryption Standard (DES) cryptographic algorithm for encryption and authentication as it authenticates requests for network resources. Kerberos uses the concept of a trusted third party to perform secure verification of users and services. This trusted third party is called the key distribution center (KDC).Kerberos verifies that users are who they claim to be and the network services that they use are what the services claim to be. To do this, a KDC or trusted Kerberos server issues tickets to users. These tickets, which have a limited life span, are stored in user credential caches. The Kerberos server uses the tickets instead of user names and passwords to authenticate users and network services.

Simple Network Management Protocol version 3 (SNMPv3) protects and encrypt administrative and network management information, from tampering or eavesdropping.

Terminal Access Controller Access Control System (TACACS+) or Remote Access Dial-In User Service (RADIUS) authentication which centralized access control of switches and restricts unauthorized users from altering the configurations.

There is also another option of configuring a local username and password database on the switch itself. In addition, there is Fifteen levels of authorization on the switch console and two levels on the web-based management interface makes different level of access possible which offer administrator with different granular configuration capabilities.

High-performance IP routing is basically routing IP packets between different IP network intelligently. It is a highly advanced form of traditional forwarding of frames via the ports using the forwarding table based on Media Access Control (MAC) Address, instead of IP routing table is used.

These switches have some proprietary architecture such as Cisco Express Forwarding (CEF)-based routing architecture for the cisco switches and this allows for increased scalability and performance. The switches have primarily hardware-based IP routing which also ensures high performance dynamic IP routing. These architectures also allows for very high-speed lookups while ensuring the stability and scalability necessary to meet the dynamic needs of present ICT demands. With these features, these switches can improve network performance when used as a stackable wiring closet switches or as a top-of- the-stack wiring closet aggregator switch.

VLAN and Inter VLAN connection: This is basically creation of logical boundary based on

device types, users and functions, for example, there could be finance users VLAN, Servers VLAN or Control systems VLAN which ensures secure and seamless communication of a device type or users etc. within a VLAN. This happens in layer 2 and it`s one of the traditional and primary roles of the switches while interVLAN connection ensures that one VLAN communicates with the other which happens in Layer 3 as communication from one VLAN is routed to others using VLAN Trunks.

Ethernet Switching between OSI layers: 2, 3 and 4: Switches basically carry out ethernet switching function within lay 2 of the OSI model. In addition to this, Ethernet catalyst switches (3500 and 4800 series) extend this function to layers 2 and 3 while 4800 series extends this switching further to layer 4 of the OSI. Based on these, what happens at each layer is briefly explained so that this positive risk is understood and be appreciated. According to Cisco Press (by Sivasubramanian B. et al) the switching process for each layer is as follows:

For Layer 2 switching:

Switching is based on MAC address
Restricts scalability to a few switches in a domain
May support Layer 3 features for QoS or access-control

For Layer 3 switching:

Switching is based on IP address Interoperates with Layer 2 features Enables highly scalable designs

Layer 4 switching:

Switching is based on protocol sessions. In other words, Layer 4 switching uses not only source and destination IP addresses in switching decisions, but also IP session information contained in the TCP and User Datagram Protocol (UDP) portions of the packet. Most common method of distinguishing traffic with Layer 4 switching is to use the TCP and UDP port numbers.

With all these, the switches still keep to their traditional LAN switching property in addition to the positive risks identified thereof as depicted in figures 1 and 2.

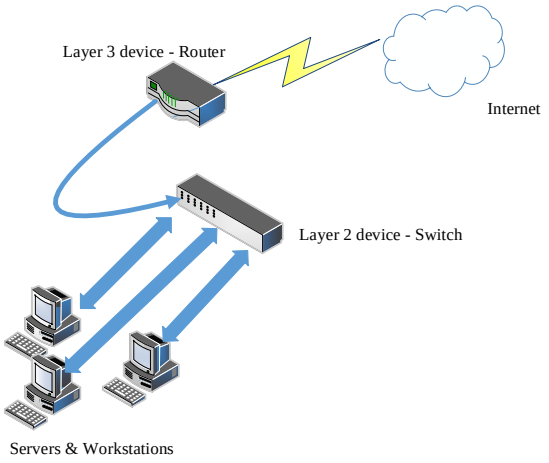


Figure 1: Traditional LAN setup without activating and harnessing positive risks associated with intelligent switches.

Traditional LAN setup must have switch and router in order for the packets to be routed in and out of the network.

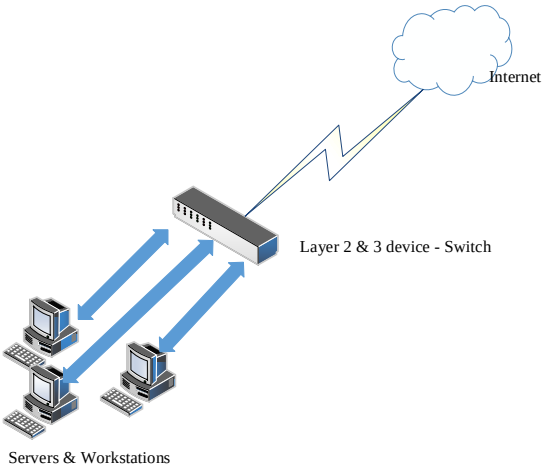


Figure 2: Utilizing Positive risk of the intelligent switch as both router and switch.

Utilizing positive risks establishes that you only need intelligent Ethernet switches in order to route packets in and out of a network without compromising any security. In fact security is further enhanced and assured.

Another outstanding example is in the Virtualization Technology where a physical device(server or workstation) has capability and capacity to house virtually similar devices(servers or workstations) of same or better specifications and functionality than the compared physical devices(servers or workstations). This is palpable in Microsoft Hypervisor (Hyper-V), VMware EXSI, Oracle virtual box etc.

Historically, Virtualization Technology (VT) has been in existence since 1960s but the Opportunity (Positive Risk) became fully activated and come to the lime light in the 1990s. This is obvious in the Dell Latitude D series in the early 2000s which is VT-enabled but must be activated in the BIOS (Basic Input/output System), just as most other positive risks.

However, nowadays, due to market demands and dynamic technology trend, most Original Equipment Manufacturers (OEMs) make this positive risk principal and default feature of their respective products. With VT approach, instead of having several physical workstations, servers etc. ,one physical server with virtualization capability is purchased and several virtual servers/ workstations are created in it with same or enhanced capabilities (memory (RAM), CPU, Storage, Operating System etc.) just same as or better than the usual physical server/workstation or other relevant data and communication devices. They (virtual Machines VMs) are presented to the user communities and they (users) access them as if they are physical devices as the functionalities and capabilities are exactly the same or even better than respective individual physical devices, in most cases better, depending on the specifications/configurations of the virtual devices. See Depiction in figures 3 and 4.

With Virtualization technology, organizations gain significant capital and operational efficiencies. This is as a result of improved workstation/server utilization and consolidation, dynamic resource allocation and management, workload isolation, security and automation. Virtualization makes on-demand self-provisioning of services and software-defined orchestration of resources easy and possible.

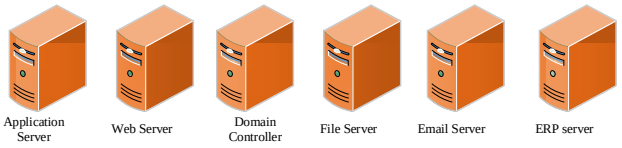


Figure 3: Non-virtualized servers (physical servers) – Acquiring one physical server for each server function

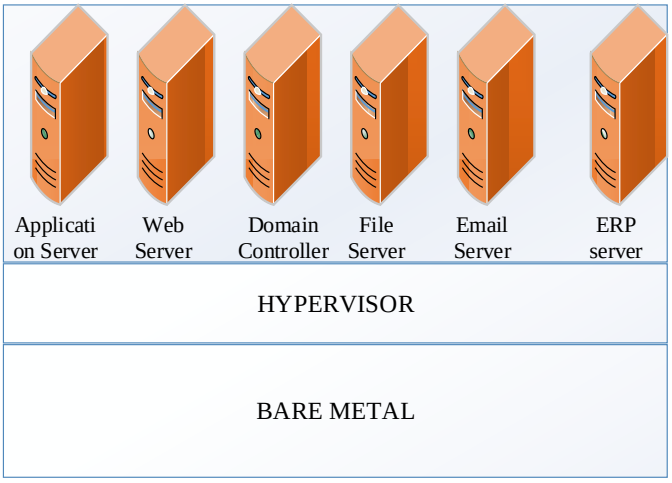


Figure 4: Virtualized Servers – Acquiring just one Physical server on which many virtualized servers are built and deployed.

The VMs seat on the Hypervisor Abstraction layer which rests on the bare metal of the physical Machine (IT devices). Comparing traditional infrastructure and virtualized infrastructure as discussed below, further justifies the industry importance of this positive risk to organizations and entire IT world:

A. Traditional Infrastructure (non-virtualized):

a. A physical machine has single OS image with non- flexible and scalable specifications such as Processor, Memory etc.
b. Most often highly under-utilized, hence costly infrastructure
c. Attempts to run many applications, comes with some bottlenecks such as interrupts conflicts, freezing of the processor and the entire machine over time.
d. In case of disaster resolution and recovery, it is difficult and far more time-consuming

B. Virtualized Infrastructure:

a. Virtual Machines (MV) `s OS and applications are independent of hardware.
b. VM carries out specific role/function and it's optimally utilized.
c. Easy, flexible and scalable to provision, recover and relatively faster to resolute issues.
d. Faster and quick to setup

Furthermore, Virtualization could take any of the forms such as:

Full Virtualization: the VM and its operating systems is not aware that it is residing in a virtualized environment. The simulated hardware are virtualized and created by the host. Hence the VMs run and operate as if they are independent physical machines both in capability and capacity. Partial Virtualization: The VMs can run many applications. However its entire operating Systems cannot run wholly in the VM, it could stimulate instance of underlying hardware of the environment. Simply put, the VM consists of independent address space so it is address space virtualization.

Para virtualization: Here the VM is aware that it is residing in a virtualized environment, hence with appropriate driver installed, can issue commands to the host operating system etc. There is explicit and direct communication between the VMs and the Hypervisor to share activity such as in interrupt handling, thread and memory management.

Virtualization as a positive risk, is the cornerstone of cloud computing. Without virtualization technology, cloud computing and its benefits may not have progressed the way it is today. In fact, it may not have seen the light of the day. This is to simply say that, cloud computing was born out of virtualization technology. And it has eliminated most of the bottlenecks associated with traditional computing.

VI. ADVANTAGES OF POSITIVE RISKS

Among numerous advantages of these positive risks to organizations are:

- Huge amount of money (Capital Expenditure - CAPEX) which would have been used to acquire those physical IT devices and their associated power consumptions are saved.
- Management, deployment, control and inspection of the virtual environment are made simple, much easier than the traditional physical servers/ workstations or other IT devices.
- Offers great flexibility and scalability for different environments – production, tests or simulations.
- Substantially made cloud computing reality, simple and flexible to orchestrate and manage.

VII. STRATEGIES FOR HARNESSING POSITIVE RISKS

Exploit: This strategy is used for risks with positive impacts on the devices where the stakeholders wish to ensure the opportunity is realized. It seeks to eliminate the uncertainty associated with a particular upside risk by ensuring the opportunity definitely happens. For example, engaging a vast expert to configure and administer ICT devices who ensures that all the devices` full potential are utilized and also embraces trends of new technologies including their upgrades in order to proactively minimize any vulnerability and negative risks.

Enhance: This is used to increase probability of occurrence and/or positive impacts of an opportunity. Identifying and Maximizing key drivers of this positive-impact risk may increase the probability of their occurrence. For example, Changing/upgrading the software (Operating systems, application etc.) and hardware of a device will definitely increase the throughput and security.

Share: Prior to sharing, Delphi technique could be used among time-tested experts with relevant experience across different platforms to explore

thoroughly ICT systems for positive risks thereof as was fully used during risk identification and risk information gathering. This allows organizations to be in the know of their inherent positive risks associated with these devices and have full conviction and justification of the needs of a particular device before purchase. Sharing a positive risk involves allocating some or all of the ownership of the opportunity to a third party who is best able to capture the opportunity for the benefit of the stakeholders. For example, forming a risk-sharing partnership, teams or joint ventures can be established with express purpose of taking advantage of the opportunity so that all stakeholders gain from their actions.

Accept: Accepting an opportunity is being willing to take advantage of the opportunity if it arises but not practically pursuing it.

VIII. STRATEGIES FOR HANDLING NEGATIVE RISKS

There are three main strategies used to deal with threats that may lead to compromise of data/ information integrity, availability and confidentiality by exploiting the vulnerabilities in the devices; if they occur are:

Risk Avoidance: This strategy is used where the risk impact is high. The stakeholders act to eliminate the threats. The most radical avoidance strategy is to shut down the devices or disconnect them from network. This may prompt the stakeholders to consult the manufacturers for immediate solution, if there is no other alternative.

Risk Transfer: Here, the stakeholders shift the impact of the threat to a third party and ownership of the responsibility by use of insurance, warranties, guarantees etc.

Risk Mitigate: In this strategy, stakeholders act early to reduce the probability of occurrence or impact of a risk. Thereby making the risk to be within acceptable threshold.

Risk Acceptance: This is used for Negative and Positive risks. In this scenario, stakeholders decide to acknowledge the risks and take no action unless the risk occurs. However this strategy provides room for periodic reviews of the threats to ensure that the risk does not change significantly. This also happens for the risks under close monitoring such as those in the risk registers. For the positive risks, the stakeholders having been made aware of the inherent opportunities of the devices may decide not to exploit and utilize them, thereby sticking to traditional use of the devices.

IX. CONCLUSION

Every ICT device has one or more positive risks but needs to be discovered and activated. Once discovered, it is also advisable to seek expert judgement and Original Equipment Manufacturers (OEMs) confirmation. With these, it is crucial also to seek services and /or advice of Subject Matter Experts (SMEs) with relevant hands-on experience and conversant with trends of technologies/security. This ensures all associated positive risks are detected and optimally utilized. Furthermore, since attacks assume dynamic forms, it is advisable to charge network Engineers , systems administrators and Experts to make thorough risk assessment a daily routine and carry out aggressive end-users awareness campaign on what to do once they sense any vulnerability that could lead to devices compromise. This is important as Risk assessment is the duty of all stakeholders hence the All stakeholders are enjoin to ensure consistent two-way communication and consultation. The overall objective is to minimize (if not totally eliminate) likelihood of occurrence and/or exploitation of negative risks and maximize likelihood of exploitation of positive risks inherent in the devices. This assures full utilization of the devices which would at the end ensure fair return on investment (ROI) and justification of their purchase.

REFERENCE

1. Stoneburner G., Goguen, A. & Feringa, A. (2002). Risk Management Guide for Information Systems. Retrieved January 4, 2015 From http://csrc.nist.gov/publications/nistpubs/800-30/sp800- 30.pdf.
2. Alshboul, A. (2010). Information Systems Security measures and countermeasures: Protecting Organizational Assets from malicious Attacks, 3, Article 486878. Journals on Communications of the IBIMA, 2010, 1-9. Retrieved March 15 2015 from http://www.ibimapublishing.com/journals/CIBIMA/2010/486 878/486878.pdfJ.
3. Manes C. (2014). The 21 most common misconfigurations that will come back to haunt you!Retrieved March 20 2015 from http://www.gfi.com/blog/the-21-most-common-misconfigurations-that-will-come-back-to-haunt-you/
4. Pascucci M. (2012). Network Security Horror Stories: Router Misconfigurations. Retrieved March 22 2015 from http://blog.algosec.com/2012/09/network-security-horror-stories-router-misconfigurations.html
5. IRS Office of Safeguards Technical Assistance Memorandum Protecting Federal Tax Information (FTI) Through Network Defense-in-Depth. Retrived April 20, 2015 from https://www.irs.gov/pub/irs-utl/protecting-fti-throughnetworkdefense-in-depth.doc.
6. Cisco Press(2014). Cisco Networking Academy's Introduction to Basic Switching Concepts and Configuration. Retrieved June 20,2015 from http://www.ciscopress.com/articles/article.asp?p=2181836&s eqNum=7
7. Valsamakis, A. C., (2003). Risk Management. Heinemann Higher and Further Education (Pty) Ltd. Sandton.
8. PMI. (2012). A Guide to the Project Management Body of Knowledge (PMBOK) Fifth Edition
9. Sabrina, M. (2014). Dell Model years. https://kb.wisc.edu/education/page.php?id=44855
10. Young, P.V. (2013). Observation Technique Definition, Principles and Validity. http://www.studylecturenotes.com/social-research-methodology/observation-technique-definition-principles-validity.
11. Cisco press(2017). Security Configuration Guide, Cisco IOS XE Release 3SE (Catalyst 3850 Switches)
12. Intel Virtualization Technology (Intel VT) https://www.intel.com/content/www/us/en/virtualization/virtualization-technology/intel-virtualization-technology.html
13. Sivasubramanian, B.,Frahim E., & Froom R.(2010). Analyzing the Cisco Enterprise Campus Architecture. Cisco Press http://www.ciscopress.com/articles/article.asp?p=1608131
14. Bhaiji,Y.(2008). Cisco Press – Security Features on switches. http://www.ciscopress.com/articles/article.asp?p=1181682&seqNum=4

Encryption & Decryption Audio Communications in Mobile Networks based on a New Hyperchaotic System

S. N. Lagmiri, J. Elalami σ & N. Elalami

Mohammed V University

ABSTRACT

With the significant development of digital communications and networking technologies, there is need to protect the sensitive data (such as digital audio signals, images, and videos) from unauthorized access, getting leak or misused. Cryptography plays a major role within the field of network security. There are many encryption techniques available currently to secure the data. Traditional encryption such as DES and many others perform poorly for multimedia data because of the large data size and high redundancy. Due to the random-like property and high sensitivity for initial values and control parameters, hyperchaotic systems are usually proposed as a solution to data encryption. In this paper, a study on security of audio data encryption based on a new six dimensional hyperchaotic system is presented. The experimental results on audio data encryption / decryption over open networks, key sensitivity tests, and statistical analysis show that the proposed cryptosystem have excellent encryption performance, high sensitivity to the security keys and can be applied for secure real time encryption.

Keywords: NA

Classification: K.6.5

Language: English

LJP Copyright ID: 975711
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology

Volume 18 | Issue 1 | Compilation 1.0

Encryption & Decryption Audio Communications in Mobile Networks based on a New Hyperchaotic System

S. N. Lagmiri[α], J. Elalami[σ] & N. Elalami[ρ]

I. ABSTRACT

With the significant development of digital communications and networking technologies, there is need to protect the sensitive data (such as digital audio signals, images, and videos) from unauthorized access, getting leak or misused. Cryptography plays a major role within the field of network security. There are many encryption techniques available currently to secure the data. Traditional encryption such as DES and many others perform poorly for multimedia data because of the large data size and high redundancy. Due to the random-like property and high sensitivity for initial values and control parameters, hyperchaotic systems are usually proposed as a solution to data encryption. In this paper, a study on security of audio data encryption based on a new six dimensional hyperchaotic system is presented. The experimental results on audio data encryption / decryption over open networks, key sensitivity tests, and statistical analysis show that the proposed cryptosystem have excellent encryption performance, high sensitivity to the security keys and can be applied for secure real time encryption.

Author α: SIP, Mohammadia School Engineering Mohammed V University, Rabat, Morocco.

σ: LASTIMI, Higher School of Technology of Sale Mohamed V University, Rabat, Morocco.

ρ: LAII, Mohammadia School Engineering Mohamed V University, Rabat, Morocco.

II. INTRODUCTION

A secure communication is one of the most important of our needs in digital world. Many studies on hiding data types like text, image, audio and video have been accomplish in order to meet such need. Speech cryptography can be defined as the art or science of altering information, so that the real information is hard to extract during transfer over any unsecured channel. The strength of the Encryption technique comes from the fact that no one can read or steal the information without altering its content [1].

In general, there are two types of encryption schemes namely symmetric encryption and asymmetric encryption. Symmetric key otherwise known as secret key or shared key or private key is one of the encryption methods [5] which use one key for encryption as they do for decryption process. Asymmetric cryptography [2, 3] uses different encryption keys for encryption and decryption.

Therefore, the efficient voice security design will has new challenges. Can the proposed system provide high security to the voice signal? To realize this, a number of voice encryption techniques have been studied [3-16]. Some of these included directly hiding audio files while others included methods of hiding the information by embedding some other data in the audio files. The general objective of all these studies is to prevent the possession of data by undesired people far [6, 13, 15, 17].

During the last decades, chaotic systems have received great attention from mathematicians,

London Journal of Research in Computer Science and Technology

physicists, biologists, control engineers, etc. see e.g. [7, 9]. This interest has been greatly motivated by the possibility of encrypted information transmission by using a chaotic carrier; see e.g. [4, 8-14].

From these researches, the chaotic system was regarded as an efficient technique for voice data. They provide high secure techniques. This is because of that the chaotic techniques have a high sensitivity to any change in its initial conditions, in addition to the other properties such as random behavior, ergodicity, and the long periodicity.

In this paper, we discuss an alternative symmetric-key encryption algorithm for securing audio message. One level of security is used to encrypt the input signal which this level is digital chaotic system in order to increase the key space.

The organization of this paper is as follows. Section 2 describes the proposed six hyperchaotic system. In section 3, proposed encryption algorithm is described. Section 4 presents the experimental part, and discusses the corresponding results. The last section concludes the paper.

III. HYPERCHAOTIC PROPOSED SYSTEM

2.1 Novel hyperchaotic system

The novel six-dimensional hyperchaotic, that exhibit hyperchaotic behavior for a selective set of its parameter, is defined by:

$$\begin{cases} \dot{x}_1 = -\mathrm{a}x_1 + \mathrm{a}x_2 \\ \dot{x}_2 = -x_1x_3 + \mathrm{b}x_4 \\ \dot{x}_3 = -\mathrm{c}x_3 + \mathrm{h}x_1x_2 \\ \dot{x}_4 = \mathrm{a}x_2 - \mathrm{a}x_4 \\ \dot{x}_5 = -x_3x_6 + \mathrm{b}x_4 + 10x_2 - 10x_5 \\ \dot{x}_6 = -\mathrm{c}x_6 + \mathrm{h}x_4x_5 \end{cases} \quad (1)$$

Where:

- x_i are the state variables and a, b, c and h are positive constants.

When $a = 5$, $b = 20$ $c = 1$ and $h = 3.5$, the system (1) is hyperchaotic.

By using the initial conditions $x_0 = [1, 0, 3, -1, 2, 4]$. Figure 1 shows the attractor of our new hyperchaotic system.

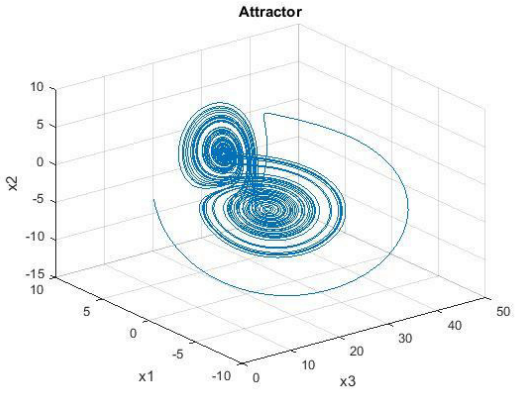


Fig. 1: Hyperchaotic attractor

One of the fundamental principles of chaotic functions is sensitive dependence, or sensitivity to initial conditions and highly complex random-like nonlinear behaviors. The performance of the system must be studied in this important feature.

2.2 Sensitivity to initial conditions

Sensitivity to initial conditions means that each point in a chaotic system is arbitrarily closely approximated by other points with significantly different future paths, or trajectories. Thus, an arbitrarily small change, or perturbation, of the current trajectory may lead to significantly different future behavior. The figure 2 compares the time series for two litely different initial conditions for the six states. The two time series stay close together, but after that, they are pretty much on their own. [10].

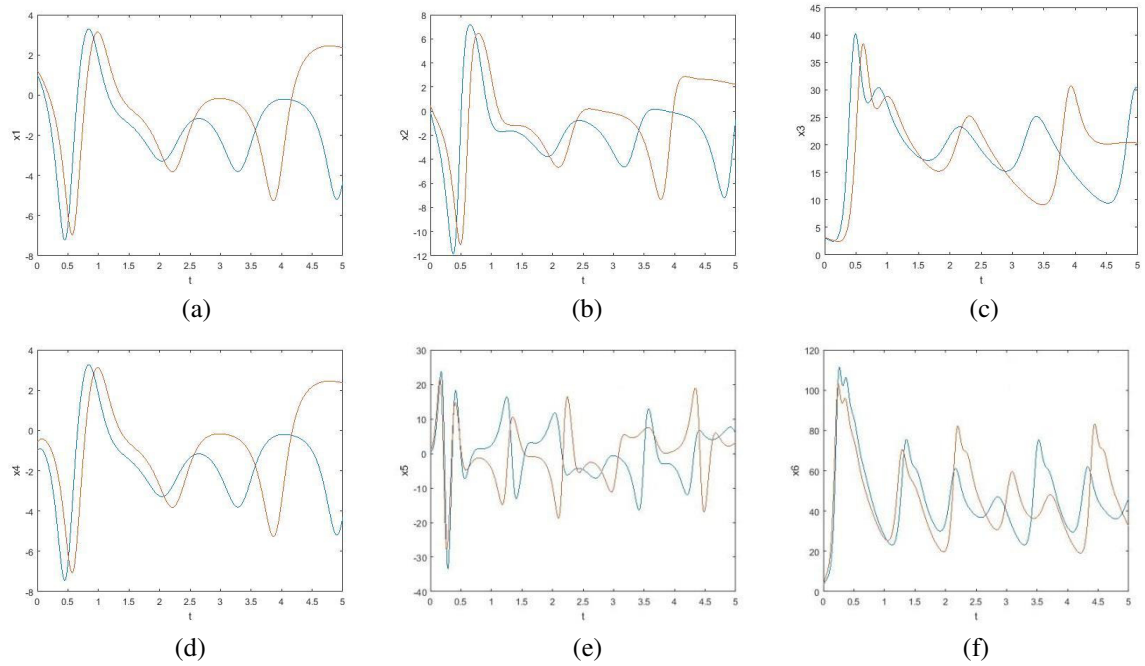


Fig. 2: Sensitivity to two initial conditions [1 0 3 -1 2 4] and [1.2 0.5 3.1 -0.6 2.7 4.2]
(a): x_1 (b): x_2 (c): x_3 (d): x_4 (e): x_5 (f): x_6

III. AUDIO ENCRYPTION IN MOBILE NETWORKS COMMUNICATIONS

With rapid advances in circuit design and prime focus on miniaturization, mobile phones have kept shrinking in size with each passing day. Hence power consumption and charge storage assume particular importance in mobile technology. Any design of a mobile communication block must take this into full account.

Enlargement of the mobile community has increased the call for secure data transmission. A computationally simple technique can be implemented easily using few components and hence consumes less power, but has limitations in the amount of security it can provide. The task of this paper is to choose an efficient and simple chaos-based encryption [12, 19, 20] strategy to meet the requirements of users.

3.1 Proposed audio encryption scheme

In this section, a cryptosystem based on synchronized chaotic systems is described. The aim is to transmit encrypted audio messages from transmitter A to remote receiver B as is depicted in Figure 3. An audio message m is to be transmitted over an insecure communication channel [18].

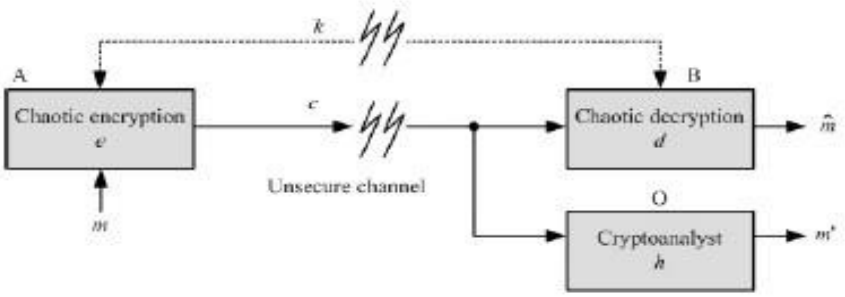


Fig. 3: Chaotic cryptosystem for audio communication

To avoid any unauthorized receiver located at the mentioned channel; m is encrypted prior to transmission to generate an encrypted message c:

$$c = e(m, k) \quad (2)$$

by using a chaotic system e on transmitter A. The encrypted message c is sent to receiver B, where m is recovered as m^ from the chaotic decryption d, as:

$$\hat{m} = d(c, k) \quad (3)$$

If e and d have used the same key , then at receiver end B it is possible to obtain m^ = m. A secure channel is used for transmission of the keys, *k*. Generally, this secure communication channel is a courier and is too slow for the transmission of m. Our chaotic cryptosystem is reliable, if it preserves the security of m, i.e. if m^ ≠ m for even the best cryptanalytic function ħ, given by

$$m' = h(c)$$

To achieve the proposed chaotic encryption scheme, we appeal to an hyperchaotic system for encryption/ decryption purposes (c and d, respectively).

The novel six dimensional hyperchaotic system have a number of parameters determining their dynamics; such parameters and initial conditions are the coding "key", k.

IV. SIMULATION RESULTS AND SECURITY ANALYSIS

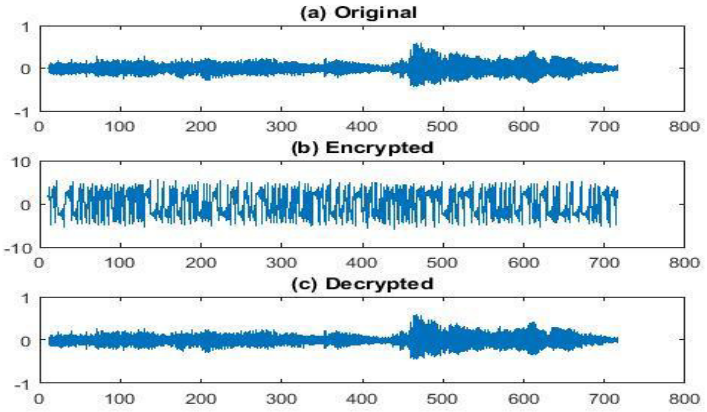
In this part, via numerical simulations, we illustrate the encrypted audio transmission. We use as transmitter and receiver the hyperchaotic system given in (1) for initial conditions $_{O1}$ = [1, 0, 3, -1, 2, 4].

The properties of the three audio signals used in this paper are presented in Table 1.

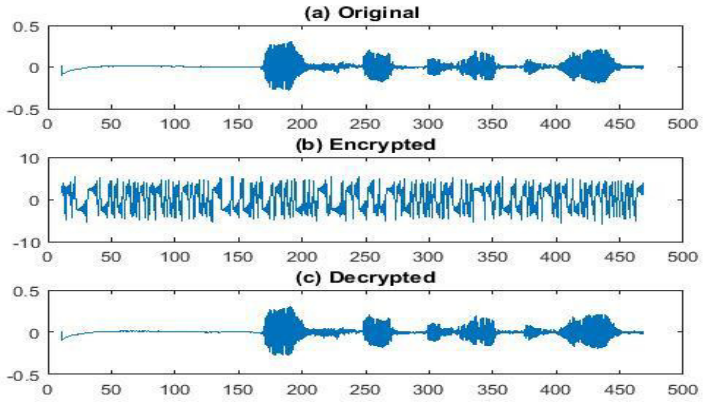
Table 1: Audio Signals Properties

	Number Channels	Frequency (KHz)	Duration (sec)
Music	1	22.05	3.2503
Speech 1	1	22.05	2.1246
Speech 2	1	48	8.9634

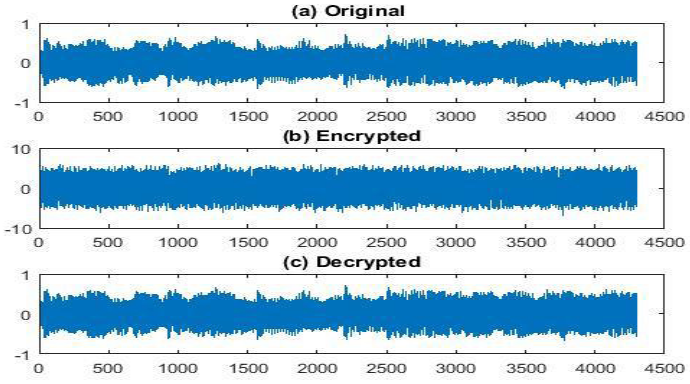
Figure 4 shows audio communication via the hyperchaotic system given in (1). Original audio message m(t) to be encrypted and transmitted (top of figure), transmitted hyperchaotic signal c(t) (middle of figure), and recovered audio message m^(t) (bottom of figure). For the three audio data, we observe that the encrypted message is very different of the original one and looks like a white noise.



(a)



(b)



(c)

Fig. 4: Encryption results of the audio communication
(a) Music - (b) Speech 1 - (c) Speech 2

4.1 Security Analyses of Encryption Applications

Encryption processes may have been performed successfully. Yet, security analyses must be carried out in order to assess the reliability of encryption processes. Encrypted data with disappointing results in security analyses will not be preferred as they are so vulnerable to be decrypted. Key space analysis, key sensitivity analysis, chaos effect and histogram were performed in order to compare the chaotic systems utilized in this study.

4.2 Histogram analysis

Distributions of data values in a system comprise the histogram. Histogram analyses can be made by examining data distributions in many different fields. In encryption practices, if the distributions of numbers that represent encrypted data are close, this means encryption has been performed well. The closer the data distributions are, the more difficult it will be to decrypt the encrypted data [25].

Exploring the histogram of audio data in Figure 5. a), b) and (c), one can see that the distribution in Middle is totally different of the one in Left Therefore, it can be concluded that encryption with our new six hyperchaotic system is successful, and that's confirmed by the decrypted data in Right that have the same histogram as the original message.

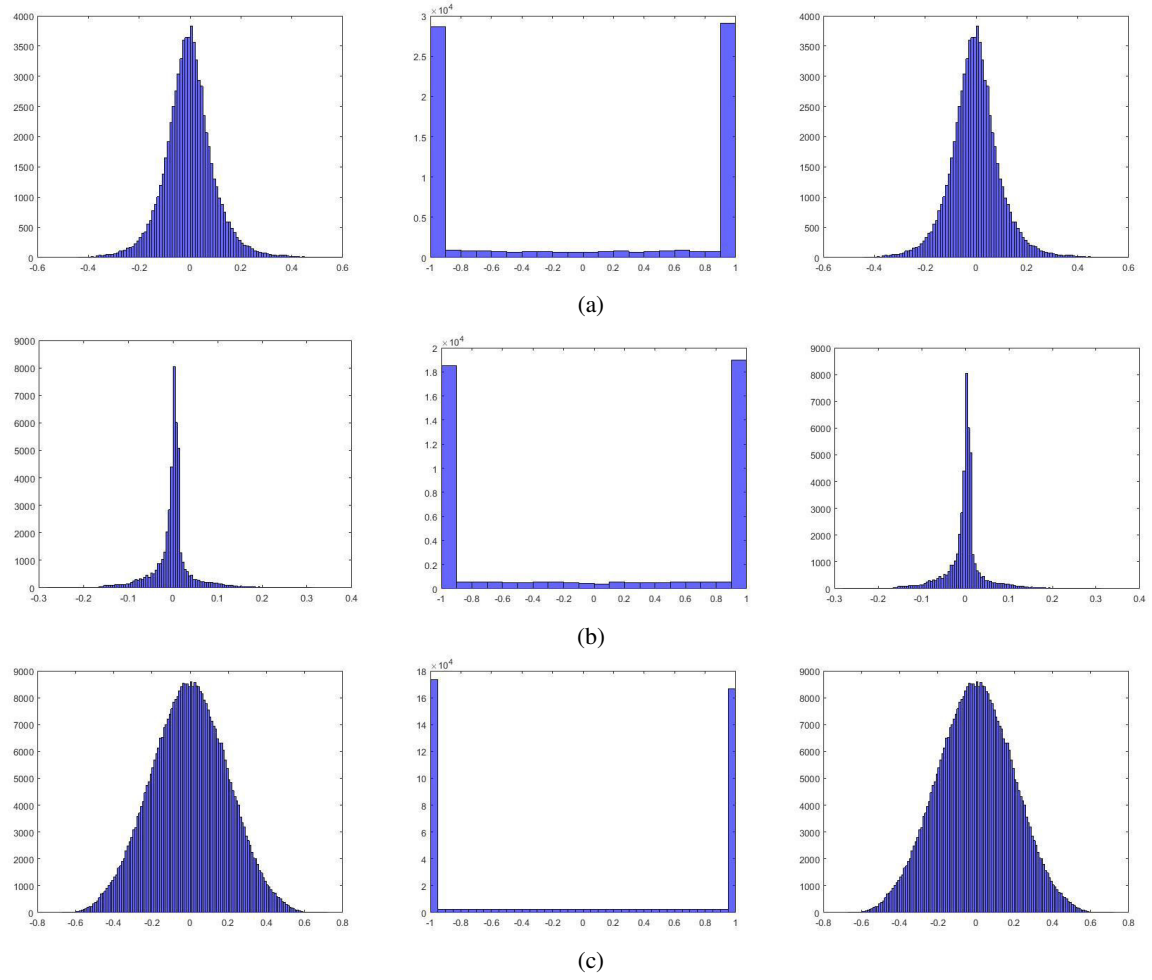


Fig. 5: Audio signals histogram (Original in Left – Encrypted in Middle – Decrypted in Right)
(a) Music (b) Speech 1 (c) Speech 2

4.3 Correlation test

The auto-correlation function identifies the chaotic system that produces a strong encryption [23]. A useful measure to assess the encryption quality of any cryptosystem is correlation coefficient between similar segments in the original signal and the encrypted signal. It is calculated as [16]:

$$r_{xk} = \frac{C(x,k)}{\sqrt{V(x)}\sqrt{V(k)}}$$

Where c (x, k) is the covariance between the original signal x and the encrypted signal k.v (x) and V (K) are the variances of the signals x and k.

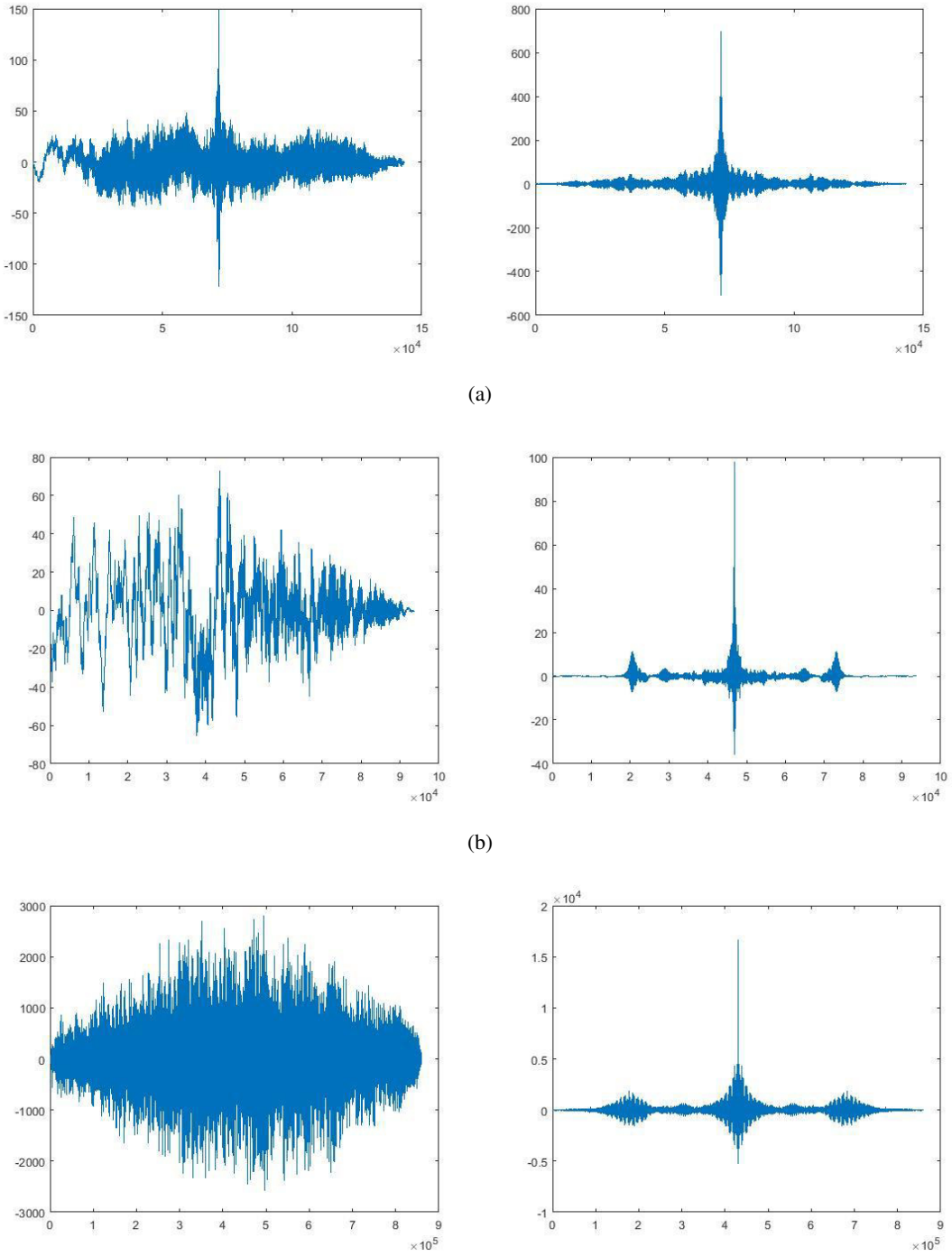


Fig. 6: Audio signals correlation: Original/Encrypted (Left) and Original/Decrypted (Right)
(a) Music (b) Speech 1 (c) Speech 2

4.4 Power spectrum

The following figures show that the power spectrum of original (a) and decrypted (c) audio signal are identical [21].

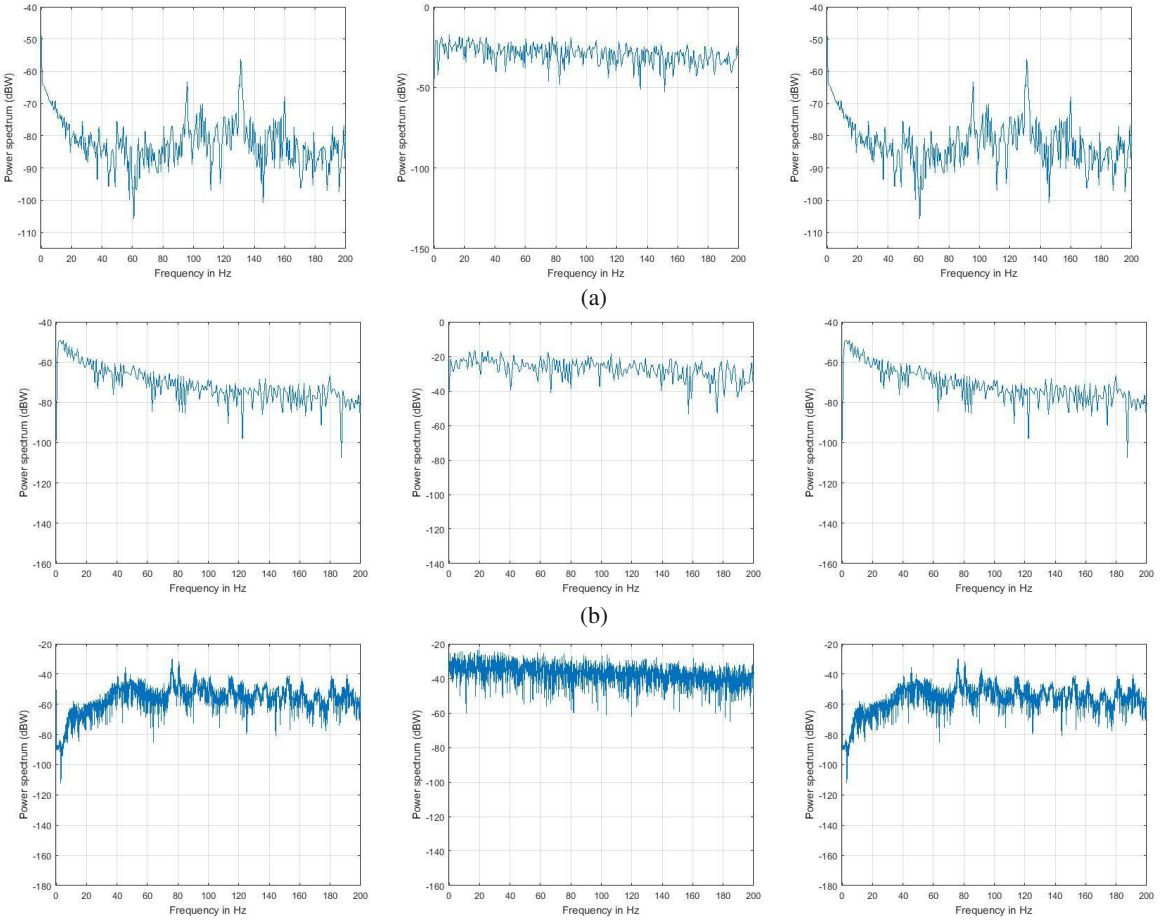


Fig. 7: Power spectrum of audio signals (Original in Left – Encrypted in Middle – Decrypted in Right) (a) Music (b) Speech 1 (c) Speech 2

4.5 PSNR test

Peak signal-to-noise ratio (PSNR) is the ratio between the maximum possible power of original speech signal and the power of encrypted signal [22]. PSNR is a calculation of encryption quality of the original signal. A higher PSNR indicates that the encryption or reconstruction is of higher quality. The PSNR is obtained from:

Table 2: PSNR Coefficient For Audio Data

	Encrypted	Decrypted
Music	47.3372	Inf
Speech 1	47.3837	Inf
Speech 2	47.2371	Inf

PSNR high means: Mean square error between the original and reconstructed signal is very low. It implies that the audio data been properly restored. In the other way, the restored signal quality is better; in our case, the value of PSNR is as follow:

$$\text{PSNR (Original/Decrypted)} = \text{Inf}$$

Contrariwise, a low PSNR means: Mean square error between the original signal and encrypted signal is very high. It implies that the audio data been correctly encrypted. In our case the value of PSNR is shown is Table 2.

The result is much closed with the correlation coefficient.

- The correlation coefficients for the original and decrypted signal are identical. The value of PSNR (Original/Decrypted) means that the decrypted audio data is identical to original data.
- The correlation coefficients for the original and encrypted signal are very different. The PSNR(Original/Encrypted) means that the encrypted audio data is totally different of the original data.

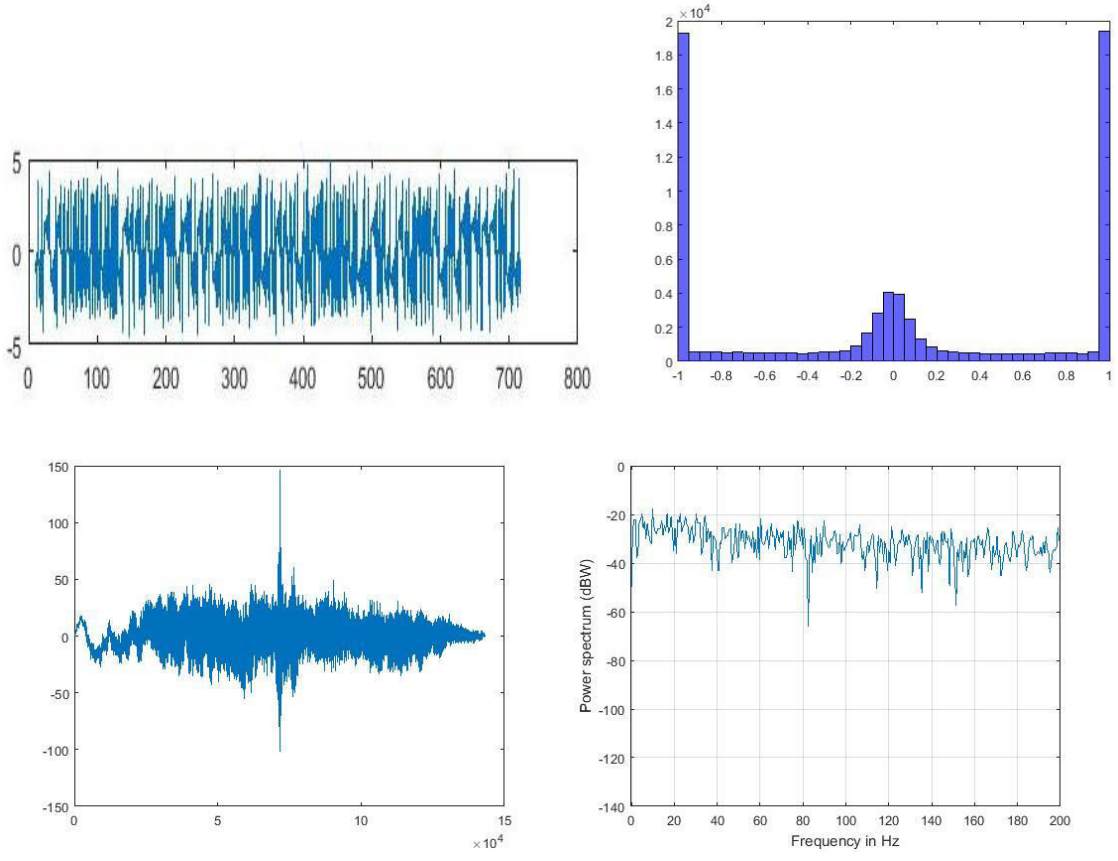
4.6 Security analysis

For testing the sensitivity of the proposed cryptosystem, the encrypted signal is decrypted with the reverse process of encryption method using the six hyperchaotic system by modifying the initial conditions of the system (1) with 10^{-9} as $_{O2}$ = [1, 0, 3, −1, 2, 4.000000001].

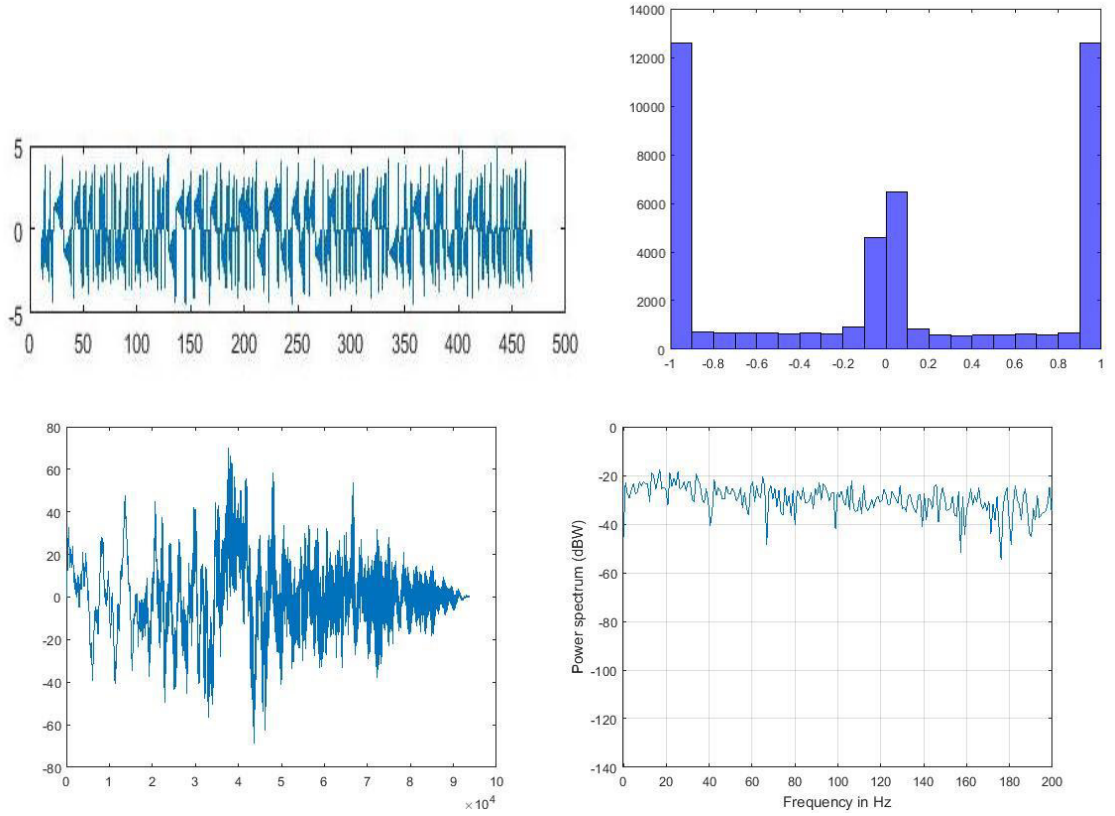
The decrypted signals are totally wrong, as shown in figure 8(a) (b) (c) (Top Left). The corresponding histogram, cross-correlation and power spectrum prove that the decrypted signals are totally different from the original ones. The Table 3 shows that the PSNR values are close to the encrypted signals. Hence the sensitivity of the encryption key is proven.

Table 3: PSNR Coefficient for Wrong Decrypted Audio Data

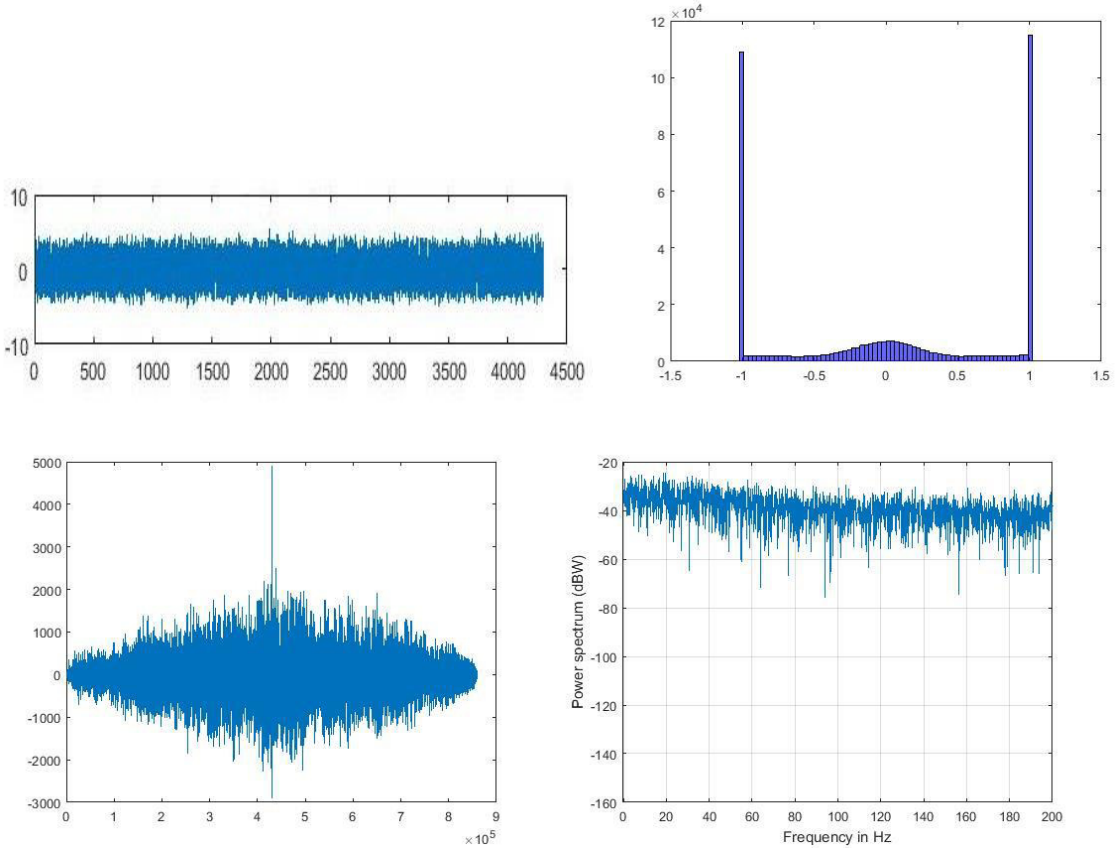
	Decrypted with wrong key
Music	48.7479
Speech 1	48.9013
Speech 2	48.6589



(a)



(b)



(c)

Fig. 8: Decrypted audio signals with $_{O_2}$ = [1, 0, 3, −1, 2, 4.000000001]

V. CONCLUSION

A new encryption system for audio files was presented. Speech encryption using hyperchaotic generator is a proven model. In this method, the three different audio data are tested. The histogram of the encrypted signal shows that more sensitivity entails more security. The simulation results showed the audio signal proposed encryption method has high level of security and can recover the original signal quickly with good audio quality. The results endorse that the speech signal is highly masked from eavesdroppers. Statistical analysis using histograms, PSNR, correlation, and power spectrum showed that the algorithm is powerful.

REFERENCES

1. Bhaskar Mondal and Tarni Mandal, "A Multilevel Security Scheme using Chaos based Encryption and Steganography for secure audio communication", Jharkhand.
2. Anoop (2007), "Public key cryptography—applications algorithm and mathematical explanations".
3. T. Thongpon & K. Sinchai. "Accelerating asymmetric-key cryptography using parallel-key cryptographic algorithm". 6th International Conference on Computer and Information Technology, 2, 812–815, (2009).
4. D. López-Mancilla & C. Cruz-Hernández, "Output synchronization of chaotic systems: model-matching approach with application to secure communication", Nonlinear Dynamics and Systems Theory, 5 (2), 141- 15 (2005).
5. J. Fridrich. "Symmetric ciphers based on two-dimensional chaotic maps". International Journal of Bifurcation and Chaos, Volume 8(6), 1259–1284, (1998).
6. KG. Gopalan, DS. Benincasa, SJ. Wennd. "Data embedding in audio signals". IEEE Aerospace Conference Proceedings (Cat. No.01TH8542). Vol. 6, pp. 2713–2720, (2001).
7. C. Cruz-Hernández & A.A. Martynyuk, "Advances in chaotic dynamics with applications", Cambridge Scientific Publishers Ltd., Vol. 4, (2009).
8. U. Feldmann, M. Hasler and W. Schwarz, "Communication by chaotic signals: the inverse system approach", Int. J. Circuits Theory and Applications, 24, 551-579 (1996).
9. L. M. Pecora and T.L. Carroll, "Synchronization in chaotic systems", Phys.
10. S. N. Lagmiri, N. Elalami, J. Elalami. "Three Dimensional Chaotic System for Color Image Scrambling Algorithm". International Journal of Computer Science and Information Security (IJCSIS), Vol. 16, No. 1, January 2018.
11. M. Delgado-Restituto, M. Linan and A. Rodriguez-Vazquez, "CMOS 2.4pm chaotic oscillator: experimental verification of chaotic encryption of audio", *Electronics Letters,* Vol. 32, Issue 9, pp.795-796, 1996.
12. Wenwu Yu and Jinde Cao, "Cryptography based on delayed chaotic neural networks", *Physics Letters A*, Vol. 356, Issues 4-5, pp. 333-338, August 2006.
13. Chang CC, Lee RTC, Xiao GX, Chen TS (2003). "A new Speech Hiding Scheme based upon sub-band coding". Proceedings of the 2003 Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing and Fourth Pacific Rim Conference on Multimedia. Vol. 2, pp. 980– 984.
14. L. Gámez-Guzmán, C. Cruz-Hernández, R.M. López-Gutiérrez, and E.E. García-Guerrero, "Synchronization of Chua's circuits with multiscroll attractors: Application to communication", Commun. Nonlinear Sci. Numer. Simulat. 14, 2765–2775 (2009).
15. Chen S, Leung H, Ding H (2007). " Telephony Speech Enhancement by Data Hiding". IEEE Transactıons On Instrumentatıon And Measurement. Vol. 56, no. 1, pp. 63–74.
16. Shujun Li, Guanrong Chen, Kwok-Wo Wong, Xuanqin Mou and Yuanlong Cai, "Baptista-type chaotic cryptosystems: problems and countermeasures", Physics Letters A, Vol. 332, Issue 5-6, pp 368-375, November 2004.

17. Dipu KHM, Alam SB (2010). “Hardware based real time, fast and highly secured speech communication using FPGA”. IEEE International Conference on Information Theory and Information Security, pp. 452–457.
18. L. M. Pecora and T.L. Carroll, “Synchronization in chaotic systems”, Phys.
19. Shujun Li, Guanrong Chen, Kwok-Wo Wong, Xuanqin Mou and Yuanlong Cai, “Baptista-type chaotic cryptosystems: problems and countermeasures”, *Physics Letters A*, Vol. 332, Issue 5-6, pp 368-375, November 2004.
20. Xiaogang Wu, Hanping Hu and Baoliang Zhang, “Analyzing and improving a chaotic encryption method”, *Chaos, Solitons & Fractals*, Vol. 22, Issue 2, pp. 367-373, October 2004.
21. S. N. Lagmiri, J. Elalami, N. Elalami. "Hyperchaotic system for encryption & decryption audio communications". The 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (IEEE TrustCom-18), 1-3 August, 2018. New York, USA.
22. S. N. Lagmiri, N. Elalami, J. Elalami. "Audio encryption algorithm using hyperchaotic systems of different dimensions”. 1st International Conference on Networking, Information Systems & Security, April 27-28, 2018, Tangier, Morocco.
23. Matej Salamon (2012), “Chaotic Electronic Circuits in Cryptography, From the book Applied Cryptography and Network Security”, InTech.
24. S. N. Lagmiri, N. Elalami, J. Elalami. "Color and gray images encryption algorithm using chaotic systems of different dimensions". International Journal of Computer Science and Network Security, Vol.18, No.1, January 2018.
25. S. N. Lagmiri, N. Elalami, J. Elalami. "Three Dimensional Chaotic System for Color Image Scrambling Algorithm". International Journal of Computer Science and Information Security, Vol.16, No.1, January 2018.

An Overview of Live Detection Techniques to Secure Fingerprint Recognition System from Spoofing Attacks

Munish Kumar & Dr. Priyanka

D.C.R. University of Sciences & Technology

ABSTRACT

Due to unique characteristics and permanence of fingerprints, Fingerprint Recognition (FPR) Systems has been extensively employed in security, identification, and verification of a person in forensic and commercial applications. External attacks or spoofing attacks are among the main challenges in using FPR systems. Spoofing means to make the system fool by producing fake finger as a real one to the fingerprint sensor. So, there is a need to develop efficient techniques to distinguish between fake and real fingers against these attacks, so that FPR systems can be secured. These types of attacks can be identified by measuring life signs and live detection techniques together with FPR techniques. This paper presents a brief review of work carried out in the field of live detection in FPR systems with various spoofing and anti-spoofing techniques. To make FPR systems more reliable, secure & efficient, it is necessary to detect and protect the system against any unpredictable spoofing attacks.

Keywords: FPR, BMR, Spoofing Attacks, live detect.

Classification: K.6.5

Language: English

London Journals Press

LJP Copyright ID: 975711
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology

2 4 U K

Volume 18 | Issue 1 | Compilation 1.0

An Overview of Live Detection Techniques to Secure Fingerprint Recognition System from Spoofing Attacks

Munish Kumar[α] & Dr. Priyanka[σ]

I. ABSTRACT

Due to unique characteristics and permanence of fingerprints, Fingerprint Recognition (FPR) Systems has been extensively employed in security, identification, and verification of a person in forensic and commercial applications. External attacks or spoofing attacks are among the main challenges in using FPR systems. Spoofing means to make the system fool by producing fake finger as a real one to the fingerprint sensor. So, there is a need to develop efficient techniques to distinguish between fake and real fingers against these attacks, so that FPR systems can be secured. These types of attacks can be identified by measuring life signs and live detection techniques together with FPR techniques. This paper presents a brief review of work carried out in the field of live detection in FPR systems with various spoofing and anti-spoofing techniques. To make FPR systems more reliable, secure & efficient, it is necessary to detect and protect the system against any unpredictable spoofing attacks.

Author α σ: ECE Department, D.C.R. University of Sciences & Technology, Murthal.

Keywords: FPR, BMR, Spoofing Attacks, live detect.

II. INTRODUCTION

In the present scenario there are various applications in which protection of sensitive data, services or facilities from illegal access is required this is usually achieved by some type of security means like password, card, key, PIN (Personal Identification Number) and biometric identification. The objective of using any one of these security means that the system is accessed by an only legal person or user [1]. The password, card, key, and PIN can protect the system against fraudulent access, but they can be cracked or guessed by hackers. Sometimes the single password or PIN is employed for access to multiple resources. If one can crack or guess that password or PIN, then all the resources can be easily accessed by an unauthorized person. Thus there is a necessity of more secure and reliable techniques to protect resources from fraudulent access.

To overcome the limitations of these recognition techniques means Biometrics Recognition (BMR) System started to find applications in security systems [2]. BMR systems use human characteristics for authentication purpose to the system and they can't be stolen easily and don't require any password to remember by the user. Biometric systems work on the human unique characteristics that are a fingerprint, voice, palm geometry, face, signature, and iris etc [3-6].

Fingerprint Recognition (FPR) systems are preferred over all the BMR systems used for authentication in personal and commercial applications due to their reliability and uniqueness in nature. Though the biometrics information can't be easily stolen, forget or lost; but due to advancement in technology one can access the system by fake identity made from a real one using various methods. BMR systems are more accurate and secure than all traditional security methods but also have various vulnerabilities that make the system less secure [7-8].

Biometrics systems are affected by various types of attacks at different levels. These attacks may be at the sensor level, replay attack, feature level, tempering stored template, matching process & matching score level, attacking the channel at a different level and overriding the final decision etc. [9]. Among all the vulnerabilities, spoofing or direct attacks have been considering for the present study. These types of attacks are also known as presentation attacks because in these attacks a fake identity of the genuine user is presented to the sensor for accessing the system. This paper focuses on the sensor level attacks, which affect the system easily without knowing the internal structure of the system.

These attacks are done by presenting fake fingerprints to the sensor for authentication purposes. An unauthorized person uses fake fingerprints are made from the impression of a real finger on artificial materials such as plastic, clay, play-doh, putty, paraffin wax, silicon, and gelatin material to access the system with this fake fingerprint. To avoid these types of attacks and to ensure right person at a right place one needs to use live detection techniques in the FPR system [10]. In this study, consideration of the different types of sensor level spoofing attacks and their detection techniques have been discussed and compared. In section II brief study about various types of attacks are discussed. Section III depicts different live detection techniques. Section IV gives a comparative study of work carried out by researchers in this area. In section V the study has been concluded.

III. SPOOFING ATTACKS IN FRS

Researchers have shown that one can easily spoof the FPR system [11], which means a false acceptance of user by using fake evidence to the FPR or biometric system. By spoofing the FPR system an unauthorized person can access the system and manipulate the data or information without knowing to an authorized user. One can manipulate the data or information in FPR system by two form of attacks as shown in fig.1 [12]. Direct attacks are performed at first stage of FPR system i.e. at sensor by making a fake finger from real one and these attacks are also known as spoofing or presentation attacks. Hence these are found outside the digital limit of FPR systems. In indirect attacks manipulation of digital information i.e. features, database, channel & score etc. is done at various stages of FPR system by the specialized hackers to illegally access the system.

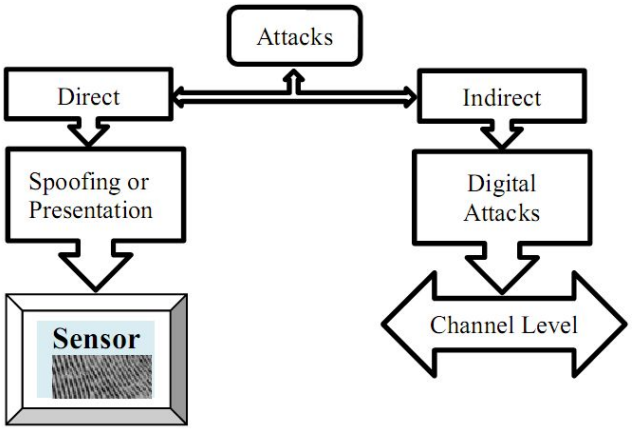


Fig. 1: Attacks in FPR System

Spoofing attacks are more popular because there is no need of any technical knowledge about the internal structure of the FPR system. These attacks are performed by making an artificial mold or fake finger of a real finger with cooperation or without the cooperation of user. There are two categories by which fake fingerprints are created [13]:

a) With the cooperation of person – directly created from the real finger.
b) Without the cooperation of person – created from dead fingerprints.

In 'with cooperation' the user itself provided his/her finger for generating fake finger or impression so these prints are made with cooperation of user from live finger and in 'without cooperation' fake finger or impression are generated from dead finger, so these prints are made from without cooperation of user i.e. latent fingerprint, fingerprint reactivation, fingerprint synthesis, and cadaver.

The main attention is carried on fake finger attacks. These attacks have been implemented by employing fake finger to the sensor made from an impression of real one. There are several categories from which fake fingerprint can be produced. Various materials incorporated for fake fingerprint are plastic print, clay print, Play-Doh print, silicon mold, gelatin material, print photo, ink print and mikrosil mold etc. According to the input data, there is various process by which fake fingerprints are created as shown in fig.2.

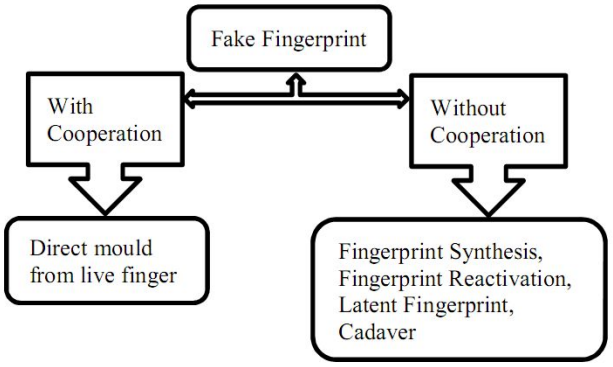


Fig. 2: Various Spoofing Methods

In fig. 3 basic block diagram of FPR system is shown with different types of attacks at different stages/level by which one can affect the system or spoofing the system with a fake identity.

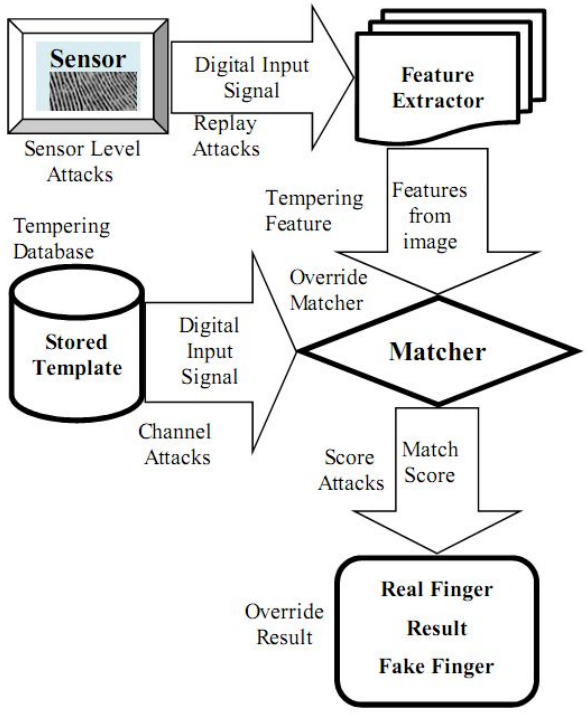


Fig. 3: Showing attacks in FPR system at different stages

Based on the literature survey any person can generate a fake fingerprint and easily spoof the FPR system. Thus, to avoid spoofing attacks, live detection techniques must be incorporated into the system [14-17].

IV. LIVE DETECTION TECHNIQUES IN FPR SYSTEM

Live detection in FPR system is imperative to ensure right person at the right place or refrain from unauthorized access to the system. There are two types of Live detection methods first is based on hardware used during acquisition stage and second is software based used during processing stages [18]. Table 1 shows live detection methods used in FPR system based these approaches [19-20].

Table 1: Live Detection Methods

S. No	Hardware-Based	Software-Based
1	Finger temperature, pulse	Skin deformation
2	Pulse oximetry	Image quality based
3	Blood pressure	Pore based
4	Electric resistance	Perspiration based
5	Odour	Combined based

In hardware-based approach interfacing of the extra device to the sensor is required to detect the real finger by acquiring life signs [21]. Biometrics sensor cannot compute life sign directly, so they require further hardware for detection. Biomedical sensors are employed for measuring life sign in hardware-based approaches [22]. Which in turn makes the system more bulky and expensive. In software- based approaches there is no need of additional hardware is required which make the system flexible for future changes/ adoption. Software-based approaches are inexpensive than a hardware based.

Table 2 shows the various spoofing detection methods used for distinguishes between real and

fake fingerprint based on the extracted feature from the input.

These features are extracted from real and fake finger based on predefined attributes. A classifier is learned based on the training set. These methods give the output in the form of numerical values and depend on these values, classifier authenticates the input as real or fake [18].

Table 2: Spoofing detection methods based on the attributes and features extracted from the fingerprint [23-36].

S. NO.	Author (s)	Feature extraction used	Attributes
1	Moon et al [23]	Coarseness analysis using noise residue.	Coarseness
2	Coli et al. [24]	Power spectrum analysis.	
3	Tan et al. [25]	Wavelet –based statistics.	
4	Abhyankar et al. [26]	Perspiration analysis using wavelet.	Perspiration
5	Marasco et al. [27]	Fusion of morphology & perspiration analysis.	
6	Marcialis et al. [28]	Statistics related to fingerprint pore analysis.	Anatomical
7	Espinoza et al. [29]	Pore analysis	
8	Tan et al. [30]	Fusion of ridge signal & valley noise analysis.	
9	Nikam et.al. [31]	Grey level cooccurrence matrix.	
10	Nikam et.al. [32]	Local binary patterns.	Textural
11	Ghiani et.al. [33]	Local phase quantization.	
12	Jia et.al. [34]	Local ternary patterns.	
13	Sansone et.al. [35]	Weber local descriptors.	
14	Ghiani et.al. [36]	Binary statistical image features.	

In table 2 four attributes coarseness, perspiration, anatomical and textural are considered for detection of spoofing attacks. Serial no. 1to 3 is based on coarseness attribute with different feature extraction methods to detect input attacks in FPR system. In serial no. 4 to 5 perspiration analysis is done using wavelet and fusion of morphology on the input image to find the real and fake finger. In serial no. 6 to 8 pore and noise analysis are considered under anatomical attributes based on that system detect input finger is real or fake. At last textural attributes with feature extraction is used to detect spoofing attacks.

In recent years software-based approaches gain more attention due to various types of algorithms that can be applied to detect spoofing attacks. Existing FPR systems can be easily protected by simply upgrade the system with a new algorithm. Thus, by applying software-based techniques existing FPR systems are protected. These solutions are easy to use, fast, less costly, easily upgradeable, and more efficient.

IV. COMPARATIVE STUDY OF WORK CARRIED OUT BY AUTHORS

In the last year, the researcher gave significant efforts in live detection techniques used in FPR systems. A comparison of different techniques used for live detection for FPR system is shown in Table 3.

Table 3 is divided into two tables i.e. (a)(b). Table 3 (a) provides a comparative study of work carried out by different researchers based on the rate of classification as a parameter to increase the performance of the system with different feature extraction techniques used for fingerprint images. And in table 3 (b) based on accuracy, error and novel material detection to find spoofing attacks with different feature extraction techniques used for fingerprint images is provided.

Table 3(a): Comparative study based on classification rate

S. NO.	Author (s)/Year	Feature Extraction Techniques used for Fingerprint images.	Remarks
1	Shankar et al., IEEE 2008, [37]	Texture & Wavelet Based	Very efficient & Classification rate is 97 %.
2	Shankar et al., IEEE 2008, [38]	Wavelet Energy Signature(WES) & Gray Level Co-Occurrence Matrix (GLCM)	Classification rates for WES is *94.35*% to *96.71*% & GLCM is *94.82*% to *97.65*%
3	Shankar et al., IEEE 2008, [39]	Ridgelet Transform	Classification rate from 94.35 % to 97.41% & slightly better.
4	Pereira et al., IEEE 2012 [40]	Multilayer Perceptron & Support Vector Machine.	System performance is increased.
5	George et al., IEEE 2012 [41]	Modular architecture based on image quality.	Reduced misclassification rate by 64.0% and performance of system increased by 49.61%.
6	Pereira et al., 2013 [42]	Spatial surface coarseness analysis.	Fingerprint classification accuracy is increased by 70.09%.
7	Jing-Wein Wang et al., 2015 [43]	DWT and Gaussian template.	Ridge structures clarity and continuity improves.
8	Kulkarni et al., IEEE 2016 [44]	Local Binary Pattern and Discrete Shearlet Transform	The method used is efficient than the others.

Table 3(b): Comparative study based on accuracy and error

S. NO.	Author (s)/Year	Feature Extraction Techniques used for Fingerprint images.	Remarks
1	Emanuela et al., IEEE 2010 [45]	Technique using multiple textural features	Reduce Average error rate to 12.47%.
2	Ankita et al., IEEE 2012 [46]	Histogram features.	Error rate reduces to 7.58%.
3	Akhtar et al. 2015 [47]	LUCID, CENTRIST, POEM features based detection	HTER - Iris: 0:01%, Face: 0:1%, Fingerprint: 0:25%.
4	Silva et al., IEEE 2015 [48]	Pores, statistical features, and quality	Achieved 97.3% accuracy
5	Rattani et al., IEEE 2015 [49]	Weibull-calibrated SVM (W-SVM)	Improvement in novel material detection up to 44%.
6	Dongju et al., IEEE 2015 [50]	Adopting ridge tracing features.	Show the improvement on spoof finger detection.
7	Marasco et al., IEEE 2015 [15]	Convolutional Neural Networks	Emprovement in accuracy- Caffe Net (96.5%), Google Net (96.6%), Siamese (93.1%).
8	Ding et al., IEEE 2016 [52]	One Class Support Vector Machine (OC-SVM) approach	Emprovement in detection accuracy & easily detect unseen material,

V. CONCLUSION

Recently the demand for biometrics systems has been increased for security and authentication purposes due to their unique characteristics. Among all biometrics, human fingerprints are most reliable and employed in large-scale applications. Biometrics systems are sometimes influenced by various attacks such as spoofing attacks. In last few years these types of attacks have become a challenge to the use of biometric systems. Hence researchers gave more attention to detect these types of attacks, that is to develop techniques to differentiate between fake and real finger and ensure the genuine user for the system. In this paper, an overview of various types of spoofing attacks and live detection techniques proposed by different researchers have been discussed. There is need to develop generalized

live detection algorithm for unpredictable and unseen spoofing attacks. One way to predict or detect these types of attacks in FPR systems is to add more features to the system for recognition, so the system becomes more secure and efficient.

REFERENCES

1. www.cedarbuffalo.edu/govind/presentations/ Fingerprints Overview.
2. Priyanka "Fingerprint Recognition Techniques and its Applications" IEEE International Conference on Advances in Engineering & Technology Research (ICAETR - 2014), August 01- 02, 2014.
3. Daugman J (1994) Biometric personal identification system based on iris analysis. United States Patent, 5291560
4. Singh D, Singh A (2010) A secure private key encryption technique for data security in a modern cryptosystem. In: BIJIT—BVICAM's International Journal of Information Technology Bharati Vidyapeeth's Institute of Computer Applications and Management BVICAM), New Delhi.
5. Kumar, M. & Priyanka.: 'Various image enhancement and matching techniques used for fingerprint recognition system', Int. j. inf. tecnol. (Springer Singapore Print ISSN 2511-2104), 2017, https://doi.org/10.1007/ s41870-017- 0061-4.
6. Panganiban A, Linsangan N, Caluyo F (2011) Wavelet- based feature extraction algorithm for an iris recognition system. J Inf Process Syst 5(3): pp.-425–434.
7. A. Jain, A. Ross, and S. Pankati, "Biometrics: A tool for information security," IEEE Trans. Inform. Forensics Security, vol. 1, no. 2, pp. 125–143, June 2006.
8. Galbally, J. Fierrez, and 1. Ortega-Garcia, "Vulnerabilities in Biometric Systems: Attacks and Recent Advances in Liveness Detection," Proceedings of Spanish Workshop on Biometrics, SWB, Girona, Spain, June 2007.
9. S. Marcel, M. Nixon, and S. Z. Li, Handbook of Biometric Anti-Spoofing. New York: Springer, 2014.
10. D. Maltoni, D. Maio, A. Jain, and S. Prabhakar, Handbook of Fingerprint Recognition. New York, NY, USA: Springer-Verlag, 2009.
11. B. Tan and S. Schuckers, "Liveness detection using an intensity-based approach in fingerprint scanner," Proceedings of Biometric Consortium Research Symposium, Crystal City, September 2005.
12. Abdenour Hadid, Nicholas Evans, Sébastien Marcel, and Julian Fierrez, "Biometrics Systems Under Spoofing Attack," IEEE SIGNAL PROCESSING MAGAZINE, pp. 20-30, 2015.
13. T. Putte and 1. Keuning "Fingerprint recognition don't get your fingers burned," Fourth Working Conference on Smart Card Research and Advanced Applications, Kluwer Academic Publishers, pp. 289-303, 2000.
14. H. Kang, B. Lee, H. Kim, D. Shin, and J. Kim, "A Study on Performance Evaluation of the Liveness Detection for Various Fingerprint Sensor Modules," Knowledge-Based Intelligent Information and Engineering Systems Lecture Notes in Computer Science, Springer Berlin Heidelberg, Vol. 2774, pp. 1245-1253, 2003.
15. H. Kang, B. Lee, H. Kim, D. Shin, and J. Kim, "A Study on Performance Evaluation of the Liveness Detection for Various Fingerprint Sensor Modules," Knowledge-Based Intelligent Information and Engineering Systems Lecture Notes in Computer Science, Springer Berlin Heidelberg, Vol. 2774, pp. 1245-1253, 2003.
16. J. Galbally, R. Cappelli, A. Lumini, G. Gonzalez-de- Rivera, D. Maltoni, J. Fierrez, 1. Ortega-Garcia, D. Maio "An evaluation of direct attacks using fake fingers generated from ISO templates, " Pattern Recognition Letters, Vol. 31, pp. 725-732, 2009.
17. T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, "Impact of artificial 'gummy' fingers on fingerprint systems, "Proceedings of SPIE Vol. 4677, Optical Security and Counterfeit Deterrence Techniques IV, January 2002.
18. Ajita Rattani, Walter J. Scheirer, and Arun Ross, "Open Set Fingerprint Spoof Detection Across Novel Fabrication Materials", IEEE

Transactions on Information Forensics and Security, Vol. 10, No. 11, November 2015.
19. B. DeCann, B. Tan and S. Schuckers "A Novel Region Based Liveness Detection Approach for Fingerprint Scanners, "Advances in Biometrics Lecture Notes in Computer Science, Springer Berlin Heidelberg, Vol. 5558, pp. 627-636,2009.
20. A. Abhyanka and S. Schuckers, "Wavelet-based Approach to Detecting Liveness in Fingerprint Scanners," Proceedings of the SPIE Vol. 5404, Defense and Security Symposium, Biometric Technology for Human Identification, pp. 278-286, April 2004.
21. H. Choi, R. Kang, K. Choi, and J. Kim, "Aliveness Detection of Fingerprints using Multiple Static Features ", World Academy of Science, Engineering and Technology 28, 2007.
22. H. Choi, R. Kang, K. Choi, and J. Kim, "Aliveness Detection of Fingerprints using Multiple Static Features ", World Academy of Science, Engineering and Technology 28, 2007.
23. Y. S. Moon, J. S. Chen, K. C. Chan, K. So, and K. S. Woo, "Wavelet-based fingerprint liveness detection," Electron. Lett., vol. 41, no. 20, pp. 1112–1113, Sep. 2005.
24. P. Coli, G. L. Marcialis, and F. Roli, "Power spectrum- based fingerprint vitality detection," in Proc. IEEE Int. Workshop Autom. Identification. Adv. Technol. (AutoID), Alghero, Italy, Jun. 2007, pp. 169–173.
25. B. Tan and S. Schuckers, "Liveness detection for fingerprint scanners based on the statistics of wavelet signal processing," in Proc. Conf. Comput. Vis. Pattern Recognit. workshop, Jun. 2006, p. 26.
26. A. Abhyankar and S. Schuckers, "Integrating a wavelet- based perspiration liveness check with fingerprint recognition," Pattern Recognit., vol. 42, no. 3, pp. 452–464, 2009.
27. E. Marasco and C. Sansone, "Combining perspiration- and morphology-based static features for fingerprint liveness detection," Pattern Recognit. Lett., vol. 33, no. 9, pp. 1148– 1156, 2012.
28. G. L. Marcialis, F. Roli, and A. Tidu, "Analysis of fingerprint pores for vitality detection," in Proc. 20th Int. Conf. Pattern Recognit. (ICPR), Aug. 2010, pp. 1289–1292.
29. M. Espinoza and C. Champod, "Using the number of pores on fingerprint images to detect spoofing attacks," in Proc. Int. Conf. Hand-Based Biometrics, Hong Kong, 2001, pp. 1–5.
30. B. Tan and S. Schuckers, "Spoofing protection for fingerprint scanner by fusing ridge signal and valley noise," Pattern Recognit., vol. 43, no. 8, pp. 2845–2857, 2010.
31. S. B. Nikam and S. Agarwal, "Wavelet energy signature and GLCM features-based fingerprint anti-spoofing," in Proc. IEEE Int. Conf. Wavelet Anal. Pattern Recognit., Hong Kong, Aug. 2008, pp. 717–723.
32. S. B. Nikam and S. Agarwal, "Local binary pattern and wavelet-based spoof fingerprint detection," Int. J. Biometrics, vol. 1, no.2, pp. 141–159, 2008.
33. L. Ghiani, G. L. Marcialis, and F. Roli, "Fingerprint liveness detection by local phase quantization," in Proc. 1st Int. Conf. Pattern Recognit., 2012, pp. 537–540.
34. X. Jia et al., "Multi-scale block local ternary patterns for fingerprints vitality detection," Proc. Int. Conf. Biometrics, Jun. 2013, pp. 1–6.
35. D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva, "Fingerprint liveness detection based on Weber local image descriptor," in Proc. IEEE Workshop Biometric Meas. Syst. Secure. Med. Appl., Sep. 2013, pp. 46–50.
36. L. Ghiani, A. Hadid, G. L. Marcialis, and F. Roli, "Fingerprint liveness detection using binarized statistical image features," in Proc. IEEE 6th Int. Conf. Biometrics, Theory, Appl. Syst., Sep./Oct. 2013, pp. 1–6.
37. Shankar Bhausaheb Nikam and Suneeta Agarwal, "Texture and Wavelet-Based Spoof Fingerprint Detection for Fingerprint Biometric Systems," First International Conference on Emerging Trends in Engineering and Technology, 978-0-7695-32 67-7/08 $25.00 © 2008 IEEE, DOI 10.1109/ ICETET.2008.134, pp. -675-680.

38. Shankar Bhausaheb Nikam and Suneeta Agarwal, "Wavelet Energy Signature And GLCM Features-Based Fingerprint Anti-Spoofing," Proceedings of the 2008 International Conference on Wavelet Analysis and Pattern Recognition, Hong Kong, 30-31 Aug. 2008, IEEE, pp.-717- 723.
39. Shankar Bhausaheb Nikam and Suneeta Agarwal, "Fingerprint Anti-Spoofing Using Ridgelet Transform," 978-1-4244-2730-7/ 08/$25.00 0D2008 IEEE.
40. Luis Filipe A. Pereira, Hector N. B. Pinheiro, Jose Ivson S. Silva, Anderson G. Silva, Thais M. L. Pina, George D. C. Cavalcanti, Tsang Ing Ren and Joao Paulo N. de Oliveira, "A fingerprint spoof detection based on MLP and SVM," WCCI 2012 IEEE World Congress on Computational Intelligence June 10-15, 2012 - Brisbane, Australia.
41. George D. C. Cavalcanti, Luis Filipe A. Pereira, Hector N. B. Pinheiro, Jose Ivson S. Silva, Anderson G. Silva, Thais M. L. Pina, Daniel B. O. Carvalho and Tsang Ing Ren, "A modular architecture based on image quality for fingerprint spoof detection," 2012 IEEE International Conference on Systems, Man, and Cybernetics October 14-17, 2012, COEX, Seoul, Korea.
42. L.F.A. Pereira, H.N.B. Pinheiro, G.D.C. Cavalcanti And Tsang Ing Ren, "Spatial Surface Coarseness Analysis: Technique for Fingerprint Spoof Detection," electronics letters 14th February 2013 Vol. 49 No. 4.
43. Jing-Wein Wang, Ngoc Tuyen Le, Chou-Chen Wang, And Jiann-Shu Lee, "Enhanced Ridge Structure for Improving Fingerprint Image Quality Based On A Wavelet Domain," IEEE Signal Processing Letters, VOL. 22, NO. 4, Pp.- 390-394 APRIL 2015.
44. Samruddhi S. Kulkarni and Dr. Hemprasad Y. Patil, "A Fingerprint Spoofing Detection System Using LBP," International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) – 2016, 978-1-4673-9939-5/16/$31.00 ©2016 IEEE, pp. -3413-3417.
45. Emanuela Marasco and Carlo Sansone, "An anti- spoofing technique using multiple textural features in fingerprint scanners," 978-1-4244-6304-6/10/$26.00 ©2010 IEEE.
46. Ankita Chaudhari and P.J. Deore, "Prevention of spoof attacks in fingerprinting using histogram features," Nirma University International Conference on Engineering, Nuicone-2012, 06-08 december, 2012, 978-1-4673-1719- 1/12/$31.00©2013 IEEE.
47. David Menotti, Giovani Chiachia, Allan Pinto, William Robson Schwartz, Helio Pedrini, Alexandre Xavier Falcão, and Anderson Rocha, "Deep Representations for Iris, Face, and Fingerprint Spoofing Detection," IEEE Transactions on Information Forensics and Security, VOL. 10, NO. 4, APRIL 2015, pp.-864-879.
48. Murilo Varges da Silva, Aparecido Nilceu Marana and Alessandra Aparecida Paulino, "On the Importance of Using High-Resolution Images, Third Level Features and Sequence of Images for Fingerprint Spoof Detection," ICASSP 2015, 978-1-4673-6997-8/15/$3l.00 ©2015 IEEE, pp.- 1807-1811.
49. Ajita Rattani, Walter J. Scheirer, and Arun Ross, "Open Set Fingerprint Spoof Detection Across Novel Fabrication Materials," IEEE Transactions on Information Forensics and Security, VOL. 10, NO. 11, NOVEMBER 2015, pp.-2447-2460.
50. Dongju Li, Hiroaki Kunieda, Supawan Kumpituck and Tsuyoshi Isshiki, "Online Detection of Spoof Fingers for Smartphone-based Applications," 2015 IEEE 17th International Conference on High Performance Computing and Communications (HPCC), 2015 IEEE 7th International Symposium on Cyberspace Safety and Security (CSS), and 2015 IEEE 12th International Conf on Embedded Software and Systems (ICESS), 978-1-4799- 8937-9/15$31.00©2015 IEEE, DOI 10.1109/ HPCC-CSS-ICESS.2015.322.
51. Emanuela Marasco, Peter Wild and Bojan Cukic, "Robust and Interoperable Fingerprint Spoof Detection via Convolutional Neural Networks," 978-1-5090-0770- 7/16/$31.00 ©2016 IEEE.
52. Yaohui Ding and Arun Ross, "An Ensemble of One- Class SVMs for Fingerprint Spoof Detection Across Different Fabrication

Materials," IEEE International Workshop on Information Forensics and Security (WIFS), 978-1-5090-1138-4/16/ 2016 IEEE.

This page is intentionally left blank

A Comparative Study of Two Dynamic Load Balancing Algorithms as a Means to Increase Performance in Shared Memory Parallel Computing

Frankline Makokha & William Okello-Odongo

University of Nairobi

ABSTRACT

The proliferation of multicore computing devices ranging from notebooks, tablets and smartphones has led to a need for load balancing to ensure optimum and full utilization of all the cores. Various algorithms exist for implementing load balancing in these multicore platforms albeit with different performance characteristics. To ensure optimum usage of all the cores, an experimental comparison of the performance of the various algorithms on real world problem domains is necessary to inform on which algorithm to use for each domain of computational problems.

This research focused on two dynamic load balancing algorithms namely, centralized dynamic load balancing algorithm and cyclic load balancing algorithm using matrix multiplication, sorting and searching as the problem domains, with the measured parameters being processing time and processor idle time.

Keywords: load balancing; parallel computing; distributed computing; distributed memory parallel systems and shared memory parallel systems.

Classification: F.2.1, C.1.4

Language: English

London Journals Press

LJP Copyright ID: 975711
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology

2 4 U K

Volume 18 | Issue 1 | Compilation 1.0

A Comparative Study of Two Dynamic Load Balancing Algorithms as a Means to Increase Performance in Shared Memory Parallel Computing

Frankline Makokha[α] & William Okello-Odongo[σ]

I. ABSTRACT

The proliferation of multicore computing devices ranging from notebooks, tablets and smartphones has led to a need for load balancing to ensure optimum and full utilization of all the cores. Various algorithms exist for implementing load balancing in these multicore platforms albeit with different performance characteristics. To ensure optimum usage of all the cores, an experimental comparison of the performance of the various algorithms on real world problem domains is necessary to inform on which algorithm to use for each domain of computational problems.

This research focused on two dynamic load balancing algorithms namely, centralized dynamic load balancing algorithm and cyclic load balancing algorithm using matrix multiplication, sorting and searching as the problem domains, with the measured parameters being processing time and processor idle time.

Based on the experiments carried out, centralized dynamic load balancing performed better in matrix multiplication in terms of processing time while cyclic load balancing algorithm performed better for sorting and searching.

From the results obtained, it is recommended to use centralized dynamic load balancing algorithm for mathematical computations while for non-mathematical computations it is recommended to use cyclic load balancing algorithm.

Keywords: load balancing; parallel computing; distributed computing; distributed memory parallel systems and shared memory parallel systems.

Author α σ: School of Computing and Informatics University of Nairobi, Nairobi, Kenya.

II. INTRODUCTION

With the increasing application of computers in almost all facets of human life, there has been a significant demand for high performance from the various computing platforms.

Due to this demand, the evolution of computers has been characterized by increasing processor speed, decreasing component size, increasing memory size, and increasing I/O capacity and speed [1]. One factor contributing to increases in processor speed is the shrinking size of microprocessor components which reduces the distance between components and hence increases speed [1]. However, the true gains in speed in recent years have come from the organization of the processor, including heavy use of pipelining, parallel execution techniques and the use of speculative execution techniques (tentative execution of future instructions that might be needed)[1].

These advances have led to development of super computers. A super computer is defined as the fastest computer currently available that provides peak performance [2]. The value of supercomputers derives from the value of problems they solve and not from the technology

they showcase [2]. It is also these advances that led to the development of vector computer systems. A vector computer/processor is a machine designed to efficiently handle arithmetic operations on elements of arrays called vectors [2].

To design computers that are effectively fast, performance of various components has to be balanced, so that a performance of some components is not dragged by low performance of others. Case in point, processor speeds have increased drastically than memory access. To ameliorate this mismatch, various techniques have been adopted, namely: use of caches, wide data path between memory and processor, and more intelligent memory chips [1].

Further, apart from increasing performance by considering the design of the internal computer components, other techniques have been developed, e.g. use of parallel computing, the use of a memory cache hierarchy, and speedup in memory access time and I/O transfer rate due to technology improvements.

However, this has also been limited by the nature of problems being solved as highlighted by Amdahl's law which states that the speedup of a parallel algorithm is effectively limited by the number of operations which must be performed sequentially [1].

Due to the above limitation, Computer Scientists have always strived to increase the performance of their computer architectures. High performance may come from fast dense circuitry, packaging technology, and parallelism [3]. This research focuses on parallelism enhanced by shrewd load balancing techniques.

Parallel computing systems are computer systems consisting of multiple processing units connected via some interconnection network plus the software needed to make the processing units work together [3].

Parallel computing systems can be classified as shared memory systems and distributed memory systems [3]. A shared memory system typically accomplishes inter processor coordination through a global memory shared by all processors while a distributed memory system combines the local memory and processor at each node of the interconnection network. Since there in no global memory in a distributed memory system, it is necessary to move data from one local memory to another by means of message passing.

One of the important mechanisms for utilizing and sharing the CPUs optimally is the policy of balancing the load amongst the processors. This type of load balancing can be achieved by transferring some of the tasks from a heavily loaded processor to a lightly loaded processor [4].

A diagrammatic representation of the two parallel computing systems is as shown in figure 1 and figure 2 [5].

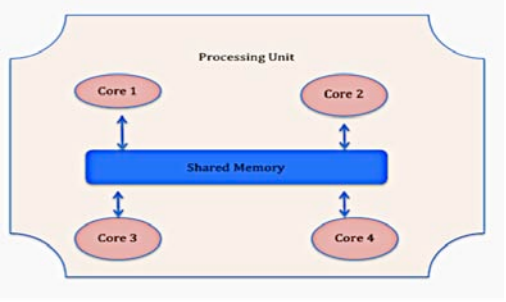


Figure 1: Shared Memory Parallel System

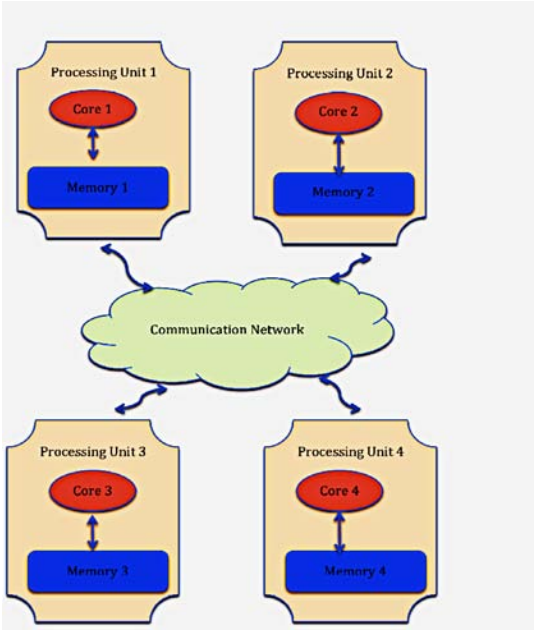


Figure 2: Distributed Memory Parallel System

III. LOAD BALANCING

Load balancing is the process of roughly equalizing the work load among all nodes of the distributed system [6].

Comprehensively, Load balancing is the distribution and/or redistribution of processing tasks among the processing nodes of a parallel or distributed system for the purpose of improving the efficiency and effectiveness of the entire processing system.

An imbalance on today's fastest supercomputers can force hundreds of thousands of cores to idle, and on future exascale machines this cost will increase by over a factor of a thousand [7].

Load balancing is one of the central problems which have to be solved in parallel computation [8]. Since load imbalance leads directly to processor idle times, high efficiency can only be achieved if the computational load is evenly balanced among the processors.

The load balancing problem can be stated as: given the initial job arrival rates at each computer in the system find an allocation of jobs among the computers so that the response time of the entire system over all jobs is minimized [9].

Load balancing algorithms can be broadly categorized as either static load balancing algorithms or dynamic load balancing algorithms. Static load balancing algorithms distribute the tasks to processing elements at compile time, while dynamic algorithms bind tasks to processing elements at run time [10].

Other authors have classified load balancing algorithms into three main classes: static algorithms, dynamic algorithms, and adaptive algorithms [11]. Static algorithms decide how to distribute the workload according a prior knowledge of the problem and the system characteristics. Dynamic algorithms use state information to make decisions during program execution. Finally, Adaptive algorithms are a special case of dynamic algorithms, which dynamically change their parameters in order to adapt its behavior to the load balancing requirements.

The various dynamic load balancing algorithms include: Centralized Dynamic Load balancing algorithm, Random (RAND), Adaptive Contracting with Neighbor, Prioritized Random (PRAND) and Cyclic Algorithm [12].

IV. BACKGROUND

Whereas work has been done in analysis of the various dynamic load balancing algorithms, most emphasis has been on distributed systems and using qualitative parameters e.g. overload rejection, reliability, predictability, adaptability, scalability, stability, waiting time, throughput etc., and thus there has been little practical emphasis on shared memory parallel systems and use of quantitative parameters like execution time and processor idle times [13].

This is attributed to the fact that during the times of those researches, shared memory parallel devices were not as prevalent as they are today.

The various works that have performed a comparative study of algorithms using qualitative parameters include: [14]; [15]; [16].

This research aimed to address this gap by performing a comparative study of two dynamic load balancing algorithms, namely Centralized Dynamic Load balancing algorithm and Cyclic Algorithm. The outcome form this research informs on the choice of load balancing algorithm to use on the various computational domains.

V. METHODOLOGY

In the comparison of the performance of the two dynamic load balancing algorithms, the following methodology was adopted.

5.1 Experimentation Methodology

The aim of this project was to practically and using real world problems compare the performance of centralized dynamic load

balancing algorithm and cyclic load balancing algorithm. The problems were implemented using each of the identified load balancing techniques and then a comparison made based on how each of the algorithm performs by measuring the processing time and processor idle time.

5.2 Experimental Problems

In this project, the identified common problems were implemented on a shared memory system for parallel execution, using each of the identified dynamic load balancing techniques.

The identified common problems are Matrix Multiplication, Sorting and Searching.

Matrix multiplication was chosen because it is an important linear algebra operation and hence a number of scientific and engineering applications include this operation as building blocks [17]. Further, matrix multiplications are important linear algebra algorithms which may simulate many real applications like image processing, video compression [11]. Due to their fundamental importance, much effort has been devoted to studying and implementing matrix multiplications. For parallel matrix multiplications, the entire task should be decomposed, this introduces various overheads. The most important are the communication and load balancing overheads.

Sorting was chosen because sorting algorithms are widely used in a broad variety of applications e.g. in commercial computing where government organizations, financial institutions, and commercial enterprises organize much of this information by sorting it [18] Further, keeping data in sorted order makes it possible to efficiently search through it.

The selection sort was chosen because it is an in-place sorting (requires no extra memory, thus ideal for small parallel systems e.g. phones and tablets where auxiliary memory is limited), and also for its simplicity in implementations.

Searching was chosen because keeping data in sorted order makes it possible to efficiently search through it. Linear search was chosen because it is extremely common in most real world applications e.g. in ruby's find_index and jQuery, further it is also the most basic search that can be found.

5.3 Experiment Design

The experiment to compare the performance of the algorithms was designed as below:

For matrix multiplication, the following structure was used:

If Matrix A: $\begin{matrix} a & c \\ b & d \end{matrix}$ and Matrix B: $\begin{matrix} e & g \\ f & h \end{matrix}$

Then the multiplication shall be:

$$\begin{matrix} (a*e)+(c*f) & (a*g)+(c*h) \\ (b*e)+(d*f) & (b*g)+(d*h) \end{matrix}$$

This results in four tasks, which are divided into chunks that are assigned to the processing entities. The experimentation was organized into chunk size 4, 8, 16 and 32.

For random number generation, part of the module code was from the code distributed under the GNU LGPL License [19].

For sorting and searching, the sort and search spaces are divided into chunk sizes 4, 8, 16 and 32 at different times during experimentation e.g. if we have chunk search space or sort space 120, then using chunk size 4, this shall be divided into sort/search space of 30 and each assigned to the processing entities

5.4 Runs

During experimentation, the following procedures were adapted

For matrix multiplication, the sizes used were as shown in table 1. Each matrix size is run eight (8) times with results recorded for each run and

averaged at the end. This is in line with [20] on secrets of successful simulation studies.

Table 1: Matrix Multiplication

No.	Matrix Size
1.	(500 by 500) *(500 by 500)
2.	(800 by 800)* (800 by 800)
3.	(2000 by 2000) *(2000 by 2000)
4.	(4000 by 4000) *(4000 by 4000)

These sizes were chosen due to their ability to produce better results based on preliminary runs.

For sorting, the sort spaces used are as shown in table 2. Each sort space is run eight (8) times and results recorded for each run and averaged at the end.

Table 2: Sort Space Size

No.	Sort Space
1.	6,000
2.	8,000
3.	10,000
4.	20,000
5.	40,000

For searching, the search space used is as shown in table 3 below: each is run eight (8) times and results recorded for each run an averaged.

Table 3: Search Space Sizes

No.	Search Space
1.	40,000
2.	60,000
3.	80,000

5.5 Chunk Sizes

A chunk is an ordered, fixed-sized array of fixed-sized slots [21]. In this experiment chunk size refers to the portion of the whole problem being solved that is assigned to a given core for execution. The chunk sizes used are as shown in table 4 .

Table 4: Chunk Sizes

No.	Chunk Size
1.	4
2.	8
3.	16
4.	32

These chunk sizes were chosen because they are multiples of the physical processors in the hardware used in the experiment and also are multiples of the total threads in the system.

VI. PLATFORM DESIGN

The system was developed as below:

6.1 Softwares used

[1] Minimalist GNU for Windows (MinGW). This is an Open Source development environment for native Microsoft Windows applications.

[2] Eclipse for Parallel Application Developers. This is an IDE for Parallel Application Developers.

[3] The development language used in the implemententation was OpenMP programming language, which enables parallel execution and timing of processing time.

Parallelism was enabled by the directive: # *pragma omp parallel,* while the timing of the processing time shall be done using: *omp_get_wtime ( )* which is an OpenMP function that returns a double precision value equal to the number of seconds since the initial value of the operating system real-time clock.

For load balancing algorithm implementation, # pragma omp for schedule (type, chunk), was used.

The type shall represent the load balancing algorithm to use while chunk represented size of work allocated to a processing entity at a given time.

6.2 Algorithms

The algorithm used in the experiment for matrix multiplication is as shown below:

1.	Generate matrix size, and associated values for the matrix	
2.	time_begin = omp_get_wtime ( );	*// get time at computation start*
3.	# pragma omp parallel	*//Break the task into parallel tasks*
4.	# pragma omp for schedule (type, chunk size)	*// state the type of algorithm to use and chunk size*
5.	Using inner for loop compute first assigned chunk and wait for next chunk	
6.	idle_start=omp_get_wtime ( );	*// get time at start of waiting for new chunk just after end of outer loop*
7.	idle_end=omp_get_wtime ( );	*// get time at end of waiting for new chunk when thread exits the outer final loop*
8.	Loop again till end of whole computation	
9.	time_stop = omp_get_wtime ( )	*//get the time after all computations*
10.	Processing time= time_stop - time_begin Processor idle time = idle_end - idle_start	

The algorithm used in the experiment for searching is as shown below:

1.	Generate an array of random numbers,	
2.	time_begin = omp_get_wtime ( );	*// get time at computation start*
3.	# pragma omp parallel	*//Break the task into parallel tasks*
4.	# pragma omp for schedule (type, chunk size)	*// state the type of algorithm to use and chunk size*
5.	Using for loop Search for a given value	
6.	idle_start=omp_get_wtime ( );	*// get time at start of waiting for new chunk just after end of outer loop*
7.	idle_end=omp_get_wtime ( );	*// get time at end of waiting for new chunk when thread exits the outer final loop*
8.	Loop again till value is found or no value found	
9.	time_stop = omp_get_wtime ( ) ;	*//get the time after all computations*
10.	Processing time= time_stop - time_begin Processor idle time = idle_end - idle_start	

The algorithm used in the experiment for sorting is as shown below.

1.	Generate an array of random numbers,	
2.	time_begin = omp_get_wtime ( );	*// get time at computation start*
3.	# pragma omp parallel	*//Break the task into parallel tasks*
4.	# pragma omp for schedule (type, chunk size)	*// state the type of algorithm to use and chunk size*
5.	Using for loop Sort the given chunk of array	
6.	idle_start=omp_get_wtime ( );	*// get time at start of waiting for new chunk just after end of outer loop*
7.	idle_end=omp_get_wtime ( );	*// get time at end of waiting for new chunk when thread exits the outer final loop*
8.	Loop again till all values are sorted	
9.	time_stop = omp_get_wtime ( ) ;	*//get the time after all computations*
10.	Processing time= time_stop - time_begin Processor idle time = idle_end - idle_start	

6.3 Experimentation platform

The experiment was perfomed on a computer with specifications as indicated in table 5:

Table 5: Experiment Terminal Specifications

No.	Item	Specifications
1.	Make	Toshiba
2.	Model	Satellite S855
3.	Processor	Intel Core i7 3630QM
4.	Speed	2.4 GHz
5.	RAM	8 GB
6.	Hard Disk	1 TB
7.	Operating System	Windows 8 64 bit

VII. RESULTS AND DISCUSSION

7.1 General Results

From the experiments done, the results are as below:

Table 6: Matrix Multiplication Results

Matrix Size	Centralized dynamic load balancing		Cyclic dynamic Load balancing	
	P-time	Idle Time	P-time	Idle Time
(4000*4000)* (4000*4000)	1111.083	0	1498.398438	0
(2000*2000) * (2000*2000)	138.9112	0	185.2085375	0
(1000*1000)* (1000*1000)	17.80218	0	22.48989	0
(800*800)* (800*800)	9.201964	0.000571	11.56868	0.000571
(500*500)* (500*500)	2.236393	0.000536	2.827786	0

The processing time can be depicted graphically as shown in figure 3.

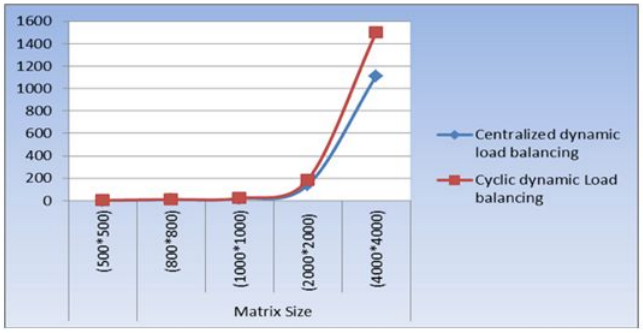


Figure 3: Matrix Multiplication results

The results for the search experiment are shown in table 7.

Table 7: Search Experiment Results

Search space	Centralized dynamic load balancing		Cyclic dynamic Load balancing	
	P-time	Idle Time	P-time	Idle Time
40,000	0.00171425	0	0.001679	0
60,000	0.0047045	0.000643	0.00275	0
80,000	0.00493915	0.000572	0.003357	0

The processing time can be depicted graphical as shown in figure 4.

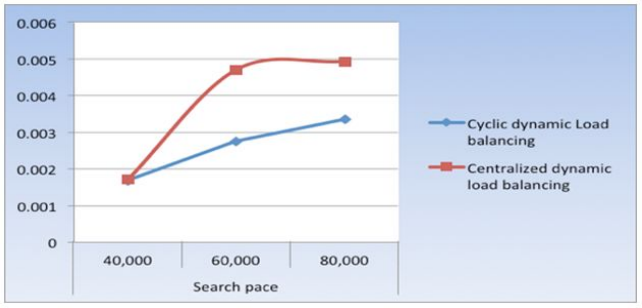


Figure 4: Search Experiment Results

From the above summaries, the cyclic dynamic load balancing performs better than the centralized dynamic on linear search.

The results for the sort experiment are as shown in table 8.

Table 8: Sort Experiment Results

Sort Size	Centralized dynamic load balancing		Cyclic dynamic Load balancing	
	P-time	Idle Time	P-time	Idle Time
40000	15.20170313	0.001031244	15.17908969	0.001031244
20000	2.936703125	0.000499994	2.9317875	0
10000	0.7455	0.002821	0.737179	0.002821
8000	0.486107	0.001143	0.477036	0.002214
6000	0.276321	0.000536	0.270036	0.002179

Graphically, the processing time can be presented as in figure 5.

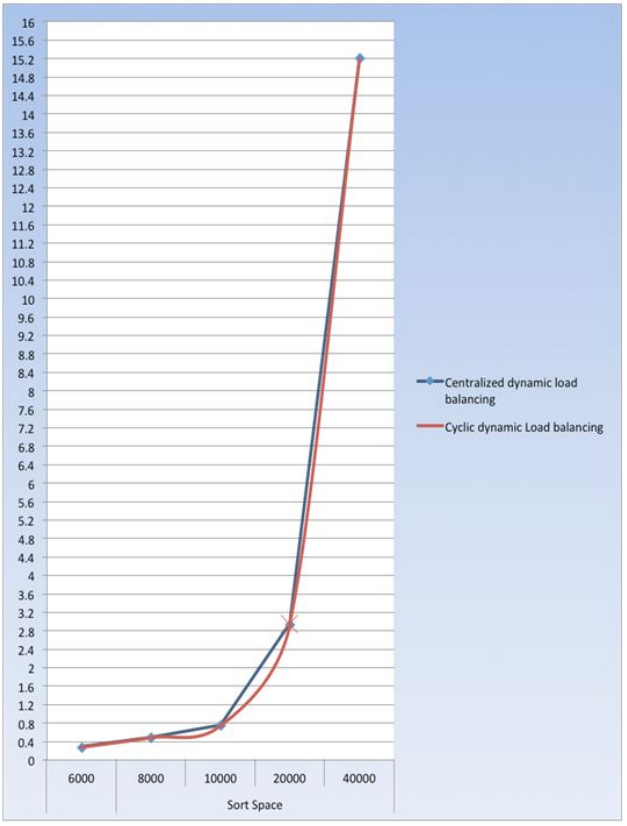


Figure 5: Sort Experiment Results

7.2 Summary of Results

1. Processing speed

For matrix multiplication, centralized dynamic load balancing algorithm performs better than cyclic load balancing as depicted in table 6. This is because for computations, which vary in computation complexities, it is better for the core to first finish its portion of assigned work before it can assign more work upon request as opposed to continuous assignment in a cyclic way. This is to reduce unbalanced workload as much as possible.

For sorting applications, the cyclic load balancing performs better than the centralized dynamic load balancing as depicted in table 8. This is because there are no varied complexities in sorting and hence waiting for the core to finish the assigned portion of work and waiting for it to request wastes time.

For searching applications the cyclic load balancing performs better than the centralized dynamic load balancing as depicted in table 7. This is because there are no varied complexities in searching and hence waiting for the core to finish the assigned portion of work and waiting for it to request wastes time

2. Processor idle time

From the conducted experiment, there was no substantial consistency in the observed processor idle time to warrant a conclusive analysis.

VIII. CONCLUSION

From the results obtained, it is recommended to use centralized dynamic load balancing algorithm for mathematical computations while for non-mathematical computations it is recommended to use cyclic load balancing algorithm.

ACKNOWLEDGMENTS

I acknowledge the assistance and guidance provided by the late Prof.Okello-Odongo, of the University of Nairobi, throughout this research before his untimely demise.

REFERENCES

1. Stallings, W. (2010). Computer Organization and Architecture: Designing for Performance. 8th Ed. New Jersey: Prentice.
2. Dongara, J. (2004) Trends in High Performance Computing. The Boole Lecture. Vol. 47. No. 4. 10th March.
3. Hesham, E. and Mostafa, A. (2005) Advanced Computer Architecture and Parallel processing. New Jersey: John Wiley & Sons, Inc.
4. Barmon, C., Faruqui, M. N. and Battacharjee, J. P. (1990/91). Dynamic Load Balancing Algorithm in a Distributed System. Microprocessing and Microprogramming. Volume 29, Issue 5.
5. Chandra, R., Dagum, L., Kohr, D., Maydan, D., McDonald, J., and Menon, R., (2001). Parallel Programming in OpenMP. San Diego: Morgan Kaufman Publishers.
6. Alakeel, A. M. (2010). A Guide to Dynamic Load Balancing in Distributed Computer Systems. IJCSNS International Journal of Computer Science and Network Security, Vol. 10 No.6.
7. Pearce, O., Gamblin, T., Supinskiy, B., Schulzy, M., and Amato, M. (2012) Quantifying the Effectiveness of Load Balance Algorithms.US: Department of Energy.
8. Horton, G. (1993) A multi-level diffusion method for dynamic load balancing. Erlangen: Elsevier Science Publishers B.V.
9. Grosu, D., Chronopoulos, T. A., and Ming-Ying L. (2002). Various Schemes of Load Balancing in Distributed Systems- A Review. In Proc. of the 16th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2002): Fort Lauderdale, Florida, USA, IEEE Computer Society Press.
10. Mandal, A. and Chandra, S. (2010) An Empirical Study and Analysis of the Dynamic Load Balancing Algorithms Used in Parallel Computing Systems: Proceedings of ICCS-2010, 19-20 Nov. West Bengal: University of North Bengal.
11. Aldasht, M., Ortega, J. and Puntonet, C. (2007) Dynamic Load Balancing in Heterogeneous Clusters: Exploitation of the Processing Power. 2nd Palestinian International Conference on Computer and Information Technology (PICCIT), Hebron, Palestine.
12. Firoj, A. & Khan, Z. (2012) The Study on Load Balancing Strategies in Distributed computing System: International Journal of Computer Science & Engineering Survey (IJCSES) Vol.3, No.2.
13. Makokha, F. and Okello_Odongo (2018) A review of Dynamic Load Balancing

Algorithms. International Journal of Computer and Information Technology. Volume 7– Issue 1, 2018.
14. Kushwaha, M. and Gupta, S (2015). Various Schemes of Load Balancing in Distributed Systems- A Review. International Journal of Scientific Research Engineering & Technology (IJSRET), Volume 4, Issue 7.
15. Manekar, S. A., Poundekar, D. M., Gupta, H. and Nagle, M.(2012). A Pragmatic Study and Analysis of Load Balancing Techniques In Parallel Computing. International Journal of Engineering Research and Applications. Vol. 2, Issue 4.
16. Willebeek-Lemair, H. M. and Revees, A. P. (1993). Strategies for Dynamic Load Balancing on Highly Parallel Computers. IEEE Transactions on Parallel and Distributed Systems. Volume 4. No 9.
17. Jin, D. and Ziavras, S. G. (2004). A Super-Programming Technique for Large. Sparse Matrix Multiplication on PC Clusters: IEICE Transactions on. Information and Systems, Vol. E87- D, issue 7.
18. Sedgewick, R. and Wayne, K. (2011) Algorithms, 4th Ed. Boston: Addison-Wesley.
19. Burkardt, J. (2008) Matrix Multiplication using C++ http://people.sc.fsu.edu/~jburkardt/cpp_src/mxm/mxm.cpp [Accessed on 17th February 2014].
20. Law, M. A. and McComas, G. M. (1991). Secrets of successful simulation studies, Winter Simulation Conference Proceedings, Phoenix, AZ, 1991.
21. Paluska, M., J. (2013). Computing with Chunks. PhD, Massachusetts Institute of Technology.

This page is intentionally left blank

Revisiting Square Roots with a Fast Estimator

Dr. Anthony Overmars, Dr. Sitalakshmi Venkatraman & Dr. Sazia Parvin

ABSTRACT

The square root of prime numbers is an important mathematical primitive with wide applications and its computational complexity has drawn much attention among researchers. In particular, principal square root of 2 is known to be irrational. The Silver Ratio $\delta_S = 1 + \sqrt{2}$ is also irrational. These can be approximated by a ratio of rational numbers.

This paper revisits some important applications of square root and proposes a new method to compute square root of 2. The proposed estimator determines the infinite series ratio of rational numbers with infinite precision. The optimised method isolates the simple fractional part and expresses this as a ratio of two rational numbers, such that $\delta_S - 2 = \sqrt{2} - 1 = r_1/r_2$. We compare the new method with the well known Babylonian method. The experimental results show that our proposed method is efficient and outperforms the Babylonian method in both accuracy and speed.

Keywords: square root, applications, silver ratio, Babylonian method, infinite series, optimisation, efficiency.

Classification: D.1.0

Language: English

London Journals Press

LJP Copyright ID: 975711
Print ISSN: 2514-863X
Online ISSN: 2514-8648

London Journal of Research in Computer Science and Technology

2 4 U K

Volume 18 | Issue 1 | Compilation 1.0

Revisiting Square Roots with a Fast Estimator

Dr. Anthony Overmars[α], Dr. Sitalakshmi Venkatraman[σ] & Dr. Sazia Parvin[ρ]

I. ABSTRACT

The square root of prime numbers is an important mathematical primitive with wide applications and its computational complexity has drawn much attention among researchers. In particular, principal square root of 2 is known to be irrational. The Silver Ratio $\delta_S = 1 + \sqrt{2}$ *is also irrational. These can be approximated by a ratio of rational numbers.*

This paper revisits some important applications of square root and proposes a new method to compute square root of 2. The proposed estimator determines the infinite series ratio of rational numbers with infinite precision. The optimised method isolates the simple fractional part and expresses this as a ratio of two rational numbers, such that $\delta_S - 2 = \sqrt{2} - 1 = r_1/r_2$. *We compare the new method with the well known Babylonian method. The experimental results show that our proposed method is efficient and outperforms the Babylonian method in both accuracy and speed.*

Keywords: square root, applications, silver ratio, Babylonian method, infinite series, optimisation, efficiency.

Author α σ ρ: School of Engineering, Construction & Design, Melbourne Polytechnic, Victoria, Australia.

II. INTRODUCTION

Square root is one of the most useful and frequently used operations in many different real world applications such as computer graphics, multimedia, data processing, cryptosystems and many scientific calculation applications [1] [2]. Several mathematical studies have determined that the theoretical complexity of such scientific calculations can be reduced by adding square roots to the basic operations [3]. In particular, square root of 2 and its associated Silver ratio have fascinated mathematicians since ancient times due to their computational complexity and applications in many real-world scenarios including architecture, modern printing, nature, music and even today's information system applications such as security and big data [4].

It is a classical hard problem to get an accurate result for computing square root and methods are available in literature such as Rough estimation, Babylonian method, Taylor-series expansion algorithm and Newton-Raphson method [5]. However, in this information age with mountains of data, applications require improved methods that are sufficiently fast with high precision [6]. Various applications make use of square root of different numbers and understanding their requirements would help in making use of s suitable square root estimator. This paper proposes a new method that we derive for estimating square root with high accuracy. In particular, we show how square root of 2 can be calculated very fast and compare with the well-known Babylonian method. The associated Silver Ratio is also estimated and discussed.

The remaining paper is organized as follows. Section III provides an overview of the various real-life applications square root. The proposed method is derived in Sections IV and optimised in Section V. The Babylonian method is described in Section VI. Experimental results for benchmarking our proposed method are summarized in Section VII. Finally, we present our conclusions in Section VIII.

III. SQUARE ROOT APPLICATIONS

Several real world applications require computing square root and square root of two [7] [1]. The

most common application of square root in everyday life is found in standard paper sizes. The base A0 size of paper is defined as having an area of 1 m^2 and a side ratio of 1 by √2,. Successive paper sizes in the series A1, A2, A3, and the frequently used A4 paper sizes have a nice property that they are derived by halving the preceding one along the larger dimension. Similarly, folded brochures of any size can be made by using sheets of the next larger size. For example, A5 sixe brochures can be made by folding A4 sheets along the length of the paper. The standard system of basing a paper size upon an aspect ratio of √2 has several advantages of not only scaling from one size to but also in fast processing of office photocopiers.

Another real life example of square root is to calculate the interest of saving account. The factor by which the account grows is exponential. This means that if it grows by some factor in a given time, it will grow by the square of that factor in double the time. So, to compute the growth in half the time, square root is used.

Square roots have been traditionally used to calculate time and distance, such as in the applications of object falling and land area triangulation measurements. For falling objects, square root can be used to calculate the time t it takes for something to fall x distance when it is dropped. From Newton's laws with g as acceleration due to gravity, we have,

$$t = \sqrt{\frac{2x}{g}}$$

Similarly. let us consider calculating the period T of a swing that is safe for children amusement park without a fall. The period of a swing with length L is calculated by the following equation:

$$T = 2\pi\sqrt{\frac{L}{g}}$$

In land areas, in order to measure the diagonal of any rectangular space when the dimensions of the sides are known, Pythagorean theorem is usually applied to measure it as the square root of sum of square of the two sides. In this context, the aspect ratio of square root application is even more prominent in today's multimedia world of graphic designs and web designs. The reciprocal of square root is also an important operation with applications in three-dimensional graphics or scientific computations.

Square root is commonly used to transform data which has applications in computer processors, statistics, big data and information security. It is possible to reduce the theoretical complexity of certain problems by adding square roots to the basic operations in the computer processor. There exists a number of situations in which square roots are used as common transformations, for e.g. square roots are computed to be faster than divides. Similarly, while both the logarithm and square root transformations are commonly used for power transformations of positive data, the square root is preferred because there is no requirement to use special treatment for zeros. Power transformations have wide applications in big data and information security. The square roots of small integers are used in both the SHA-1 and SHA-2 hash function designs in cryptosystems to achieve information security [8] [9]. Square roots are used in decryption in order to compute the encrypted text called the cipher text. If the cipher text c is modulo of the primes p and q, then square roots can be more easily calculated by choosing $p \equiv q \equiv 3 \ mod \ 4$ in order to decrypt the message. Hence, the basic step of using square roots in public-key cryptosystems is in decryption where if the cipher text is given, one can extract the keys using square root. The metallic numbers, namely Golden ratio and Silver ratio have been used extensively in cryptography [10] [11].

In summary, square root is an important mathematical primitive having many applications. Due its computational complexity there are some difficulties to ensure its security and efficiency in distributed computing [12]. Hence, several algorithms have evolved, such as Goldschmidt's algorithm using linear approximation, Manuel Liedel algorithm through the fixed-point

arithmetic framework of Catrina/Saxena. Goldschmidt's algorithm for square root is preferred over the traditional Newton-Raphson iterations due to faster computation as iteration contains fewer dependent multiplications with the same computational complexity. In this paper we address the problem of computational complexity by proposing a faster optimised method with high precision. for estimating square roots, in particular square root of two as it has wide applications in the computing field.

IV. PROPOSED METHOD

We consider square root of as an illustration for our proposed method of estimating square roots. In Fig. 1, we pictorially represent square root as the hypotenuse of a right angled triangle with base and height to be equal of value 1 unit each.

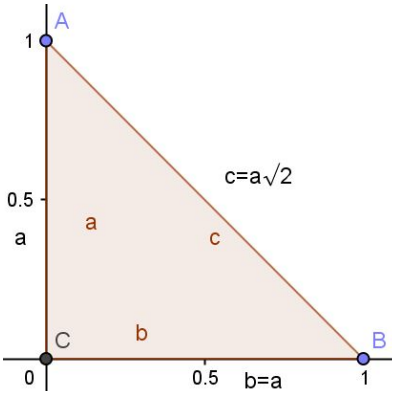


Figure 1: Square Root of 2 as a right angle triangle

We now consider the three propositions of representing square root of two as shown in Fig 2. The first proposition increases more rapidly than the others because of the $+2a$ term. We can re-express this in terms of a :

$$a = a, \quad b = a + 1, \quad c = \sqrt{2a^2 + 2a + 1}$$

Let us now express sides as $lim_{a\to\infty}$:

$$a = a, \; b = lim_{a\to\infty}\, a + 1 = a \,,$$

$$c = lim_{a\to\infty}\, \sqrt{2a^2 + 2a + 1} \; = a\sqrt{2}$$

$$\sqrt{2} = \tfrac{c}{a} = \tfrac{c}{b} \qquad (1)$$

This is consistent with Fig. 1. From previous work [13] [14] it can be shown that sides (a, b, c) can be expressed as three Diophantine equations in terms of (m, n) : $P(a, b, c) = P(m, n)$ where

$$a = 2n^2 + 2n(2m - 1) \qquad (2)$$

$$b = 2n(2m - 1) + (2m - 1)^2 \qquad (3)$$

$$c = 2n^2 + 2n(2m - 1) + (2m - 1)^2 \qquad (4)$$

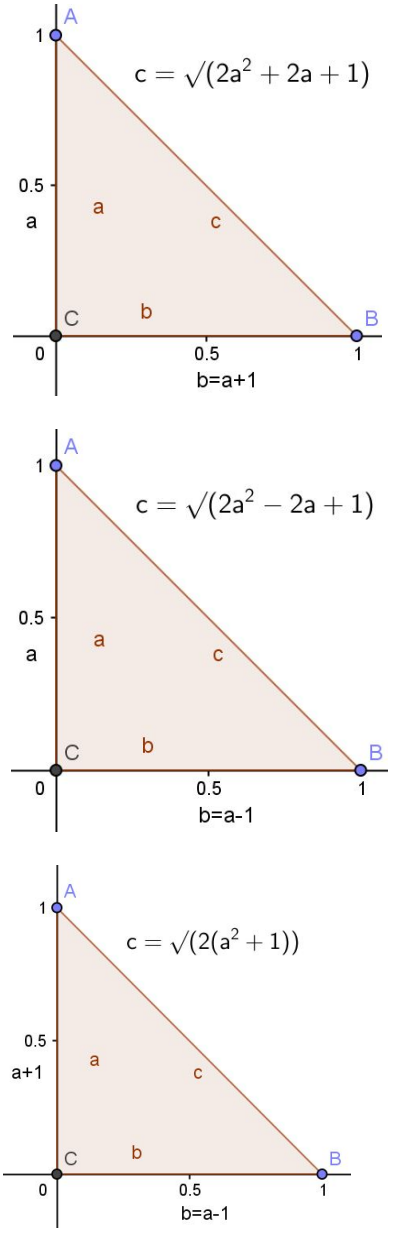


Figure 2: Approximations of Square Root of 2

Note the special condition that we specified in proposition:

$$(1)\; b = a + 1 \qquad (2)\; b = a - 1$$

Case (1): Substituting in for a *and* $b = a + 1$:

$$2n(2m-1) + (2m-1)^2$$

$$= [2n^2 + 2n(2m-1)] + 1$$

This gives us the special condition:

$$2n^2 - (2m-1)^2 + 1 = 0 \quad (5)$$

$$n = \sqrt{\frac{(2m-1)^2-1}{2}} \quad (6)$$

Solving this gives: $m = 2,\ n = 2$

Case (2): Substituting in for a *and* $b = a - 1$:

$$2n(2m-1) + (2m-1)^2$$

$$= [2n^2 + 2n(2m-1)] - 1$$

This gives us the special condition:

$$2n^2 - (2m-1)^2 - 1 = 0 \quad (7)$$

$$n = \sqrt{\frac{(2m-1)^2+1}{2}} \quad (8)$$

Solving this gives: $m = 4,\ n = 5$. $m = 1,\ n = 1$ is also a solution.

$P(m,n)$: $P(1,1) = (3,4,5)$

We observe the following:

$P(1,1) = (3,4,5),\ \ P(4,5) = (119,120,169)$.

In equation (8) $m = 120 \Longrightarrow n = 169$, is true.

Therefore, the following conjecture holds:

$$\begin{aligned}
&Consider\ P(m_i,n_i) :\ P(m_i,n_i) = (a_i,b_i,c_i)\ and\ P(m_{i+1},n_{i+1}) :\ P(m_{i+1},n_{i+1}) = (a_{i+1},b_{i+1},c_{i+1})\\
&P(m_i,n_i) = (a_i,b_i,c_i)\\
&\Longrightarrow P(b_i,c_i) = (a_{i+1},b_{i+1},c_{i+1}) = P(m_{i+1},n_{i+1})\\
&P(m_i,\ n_i) = (a,\ m_{i+1},n_{i+1})\\
&\Longrightarrow m_{i+1} = b(m_i,\ n_i)\ and\ n_{i+1} = c(m_i,\ n_i)\\
&P(m_i,n_i) :\ P(m_i,n_i) = (a_i,m_{i+1},n_{i+1}) \qquad (9)
\end{aligned}$$

From equation (1), we have

$$\sqrt{2} = \frac{c}{a} = \frac{c}{b}$$

And equations (2), (3), (4), we get

$a = 2n^2 + 2n(2m-1)$,

$b = 2n(2m-1) + (2m-1)^2$,

$c = 2n^2 + 2n(2m-1) + (2m-1)^2$

Re-expressing a : $a = 2n^2 + 2n(2m-1) \quad \Longrightarrow$

$$a = 2n\,(2m + n - 1)$$

Substituting the above expressions, we get

$$\sqrt{2} = \frac{c}{b} = \frac{2n^2+2n(2m-1)+(2m-1)^2}{2n(2m-1)+(2m-1)^2} = \frac{2n^2+b}{b} = \frac{2n^2}{b} + 1$$

From equation (1) $\frac{c}{a} = \frac{c}{b}$ $\therefore a = b$

$$\Longrightarrow \sqrt{2} = \frac{2n^2}{b} + 1 = \frac{2n^2}{a} + 1$$

Recalling $a = 2n\,(2m + n - 1)$

$$\frac{2n^2}{a} + 1 = \frac{2n^2}{2n\,(2m+n-1)} + 1 = \frac{n}{2m+n-1} + 1$$

$$\sqrt{2} - 1 = \frac{n}{2m+n-1}$$

Expressing this in terms of the following, we get

$$\sqrt{2} - 1 = \delta_S - 2 :\ \frac{r_1}{r_2} = \frac{n}{2m+n-1} \quad (10)$$

Now using equation (10)

$$i_1 = \frac{n_1}{2m_1+n_1-1} = \frac{5}{2(4)+5-1} = \frac{5}{12} = 0.41|\ 6..$$

$$i_2 = \frac{169}{408} = 0.41421|\ 5\ ..$$

$$i_3 = \frac{195025}{470832} = 0.41421356237|\ 4\ ..$$

$$i_4 = \frac{259717522849}{627013566048} = 0.41421356237309504880168 9|\ 6\ ..$$

$$i_5 = \frac{4605992036830504954151105}{1111984844349868137938112} = 0.414213562373095048801688724209698078569671875377|..$$

$$i_6 = \frac{144866192049726070675423463550278814198100428556 9}{3497379255757941172020851852070562919437964212608}$$

$$i_6 = 0.414213562373095048801688724209698078569671875376948073176679737990732478462107038850387534327764160|$$

$$Actual\ \sqrt{2}$$

$$= 1.414213562373095048801688724209698078569671875376948073176679737990732478462107038850387534327641570..$$

From equations (9) and (10) the results of which are shown in Table 1.

Table 1: Precision per iteration

i	1	2	3	4
m	4	120	137904	183648021600
n	5	169	195025	259717522849
.	2	5	11	24

i	5	6
m	32569282033 3408821261 504	10243586676303402326 3330860828388738872 8479963520
n	4605992036 8305049541 5105	144866192049726070675 42346355027881419810 04285569
.	48	97

V. OPTIMISING THE METHOD

It can be seen from Table 1 that the decimal accuracy doubles up each iteration and that after the 6th iteration we have an accuracy of 97 decimal places. Since the sides of the triangles are co-prime [13]: the sides have a gcd(c,b)=1. From equation (10) it can also be shown that $\sqrt{2} - 1 = \frac{n}{2m+n-1}$ has no common factors and is irreducible.

We now optimise the number of arithmetic operations per iteration for the new method.

Recall equation (9) as follows;

$$P(m_i, n_i) : \ P(m_i, n_i) = (a_i, m_{i+1}, n_{i+1})$$

We are only concerned with b_i *and* c_i to determine m_{i+1} and n_{i+1} so only equations (3), (4) are needed. Recall the following equations:

$$m_{i+1} = b_i = 2n_i(2m_i - 1) + (2m_i - 1)^2$$

$$n_{i+1} = c_i = 2n_i^2 + 2n_i(2m_i - 1) + (2m_i - 1)^2$$

This can now be optimised to minimise the number of arithmetic operations.

$$c_i = 2n_i^2 + 2n_i(2m_i - 1) + (2m_i - 1)^2,$$
$$b_i = 2n_i(2m_i - 1) + (2m_i - 1)^2 = (2m_i - 1)(2n_i + 2m_i - 1)$$
$$c_i = 2n_i^2 + b_i$$
$$m_{i+1} = b_i \ , \ n_{i+1} = c_i$$
$$m_{i+1} = (2m_i - 1)(2n_i + 2m_i - 1) \quad (11)$$
$$n_{i+1} = 2n_i^2 + m_{i+1} \quad (12)$$

Equation (11) has 5 arithmetic operations and equation (12) has 2 so this is 7 operations per iteration. The final division (costly) is only required in the last calculation, equation (10) which adds further 4 operations. Table 2 summarises the number of operations.

Table 2: Operations per iteration

i	1	2	3	4	5	6	7
.	2	5	11	24	48	97	≈192 ?
Ops	11	18	25	32	39	46	53

From table 2 we can approximate the operations per iteration as 7 with a corresponding accuracy of number of decimal digits of precision. Each m, n can be determined without any divisions being required until the last calculation in equation (10).

$$Operations: Op = 7i + 4,\ i = \frac{Op-4}{7}$$
$$Accuracy: A \approx 3(2^{i-1}) \Rightarrow A \approx 3(2^{\frac{Op-11}{7}})$$

The accuracy can be estimated per operation as: $A \approx 3(2^{\frac{Op-11}{7}})$

We can test this with 53 operations and estimate that the accuracy will be ≈192 decimal places.

$$Op = 53 \Rightarrow A \approx 3(2^{\frac{Op-11}{7}}) = 3(2^{\frac{53-11}{7}}) = 192$$

VI. THE BABYLONIAN METHOD

There are a number of algorithms for approximating √2, which in expressions as a ratio of integers or as a decimal can only be approximated. The most common algorithm that has been used in many computers and calculators is the Babylonian method [15] of computing square roots. In this section, we present the highlights of the Babylonian method for estimating square root of 2. By choosing, $a_0 > 0$, we iterate through the recursive computation for this method:

$$a_{n+1} = \frac{a_n + \frac{2}{a_n}}{2} = \frac{a_n}{2} + \frac{1}{a_n} = \frac{a_n^2+2}{2a_n} = \frac{f}{g}$$

Each iteration approximately doubles the number of correct digits. Starting with $a_0 = 1$ the next approximations are:

$$\sqrt{2}(1) = \frac{3}{2} = 1.5$$

$$\sqrt{2}(2) = \frac{17}{12} = 1.41 \mid 6..$$

$$\sqrt{2}(3) = \frac{577}{408} = 1.41421 \mid 5..$$

$$\sqrt{2}(4) = \frac{665857}{470832} = 1.41421356237 \mid 4..$$

$$\sqrt{2}(5) = \frac{886731088897}{627013566048}$$

$$= 1.41421356237309504880168 \mid 6..$$

$$\sqrt{2}(6) = \frac{1572584048032918633353217}{1111984844349868137938112}$$

$$= 1.41421356237309504880168872420969807856967187537 \mid 2..$$

The results of which are shown in Table 3.

Table 3: Babylonian method - precision per iteration

i	1	2	3	4
f	3	17	577	665857
g	2	12	408	470832
.	1	3	5	11

i	5	6
f	886731088897	1572584048032918633353217
g	627013566048	1111984844349868137938112
.	24	48

Though the divide operation is usually costly, in this case, a division by 2 can be simply implemented with a binary shift left operation of the whole binary number. This usually takes one clock cycle. For Big Integers, this will take longer, but can be efficiently performed using the above method. The divide is very expensive computationally and a simple addition is all that is required to obtain the iteration. So, we have three arithmetic operations per iteration with accuracy as given below:

$$Operations: Op = 3i,\ i = \frac{Op}{3}$$
$$Accuracy: A \approx 3(2^{i-1}) \Rightarrow A \approx 3(2^{\frac{Op}{3}})$$

VII. COMPARISON OF RESULTS

The accuracy of our proposed new method per operation can be estimated as: $A_O \approx 3(2^{\frac{Op-11}{7}})$ as compared to the accuracy of the Babylonian method given by $A_B \approx 3(2^{\frac{Op}{3}})$.

The result of operations for our proposed method is shown in Table 4.

Table 4: Results of proposed method

i	1	2	3	4	5	6	7
$Ops(A_O)$	11	18	25	32	39	46	53
A_O	2	5	11	24	48	97	192
$Ops(A_B)$	3	6	9	12	15	18	21
A_B	1	3	5	11	24	48	96

The Babylonian method achieves an accuracy of 48 decimal digits with 18 operations. This same level of accuracy requires the proposed method to perform 39 arithmetic operations. At the outset, the Babylonian method may appear more efficient by a factor of 2 per iteration. However, the Babylonian is specifically optimised to only solving the $\sqrt{2}$. For larger accuracies the cost of the two divides per iteration plays a major factor on its efficiency [16] [17]. Comparatively, the proposed method only needs one divide at the very end, while the Babylonian method requires it for every iteration. It is also very efficient in determining a series of rational numbers, which can be expressed as a quotient, that when divided, very accurately represents $\sqrt{2}$ *and* δ_S. The proposed method is also capable of solving more general cases of the $\sqrt{x}$.

VIII. CONCLUSIONS

The computation of square roots have wide real-life applications. This paper proposed a new method for estimating square roots efficiently. We compared our method with the well-known Babylonian method. The experimental results shows that the Babylonian method achieves an equivalent accuracy to the proposed method with one additional iteration. Even though it is about 2 times more efficient for small quotients, for larger factors the cost of the two divisions per iteration is impeding its performance. The Babylonian method exhibits a specific optimisation to only solving the $\sqrt{2}$. The proposed method considers only the fractional component and is capable of expressing this as a series of quotients whose accuracy need only be determined by one division at the end of the desired resolution. The advantage of our proposed method is that with a simple addition, both $\sqrt{2}$ *and* δ_S can both be determined. This is achieved by performing an addition of 1 for $\sqrt{2}$, and an addition of 2 for the Silver ratio, δ_S. The proposed method is also capable of solving more general cases of the $\sqrt{x}$ problem, which will be presented in future papers.

REFERENCES

1. Sutikno T., An Efficient Implementation of the Non Restoring Square Root Algorithm in Gate Level, International Journal of Computer Theory and Engineering, Volume 3, Issue 1, pp. 46-51, 2011.
2. Wang X., Variable Precision Floating-Point Divide and Square Root for Efficient FPGA Implementation of Image and Signal Processing Algorithms, PhD Thesis, Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts, 2007.
3. Dong-Guk H., Choi D., Kim H., Improved Computation of Square Roots in Specific Finite Fields, IEEE Transactions on Computers, Volume 58, pp. 188-196, 2009.
4. Llamocca-Obregon D. R., A Core Design to Obtain Square Root Based on a Non-Restoring Algorithm, IBERCHIPS Workhsop, Salvador Bahia, Brazil, 2005.
5. Chu W. and Y. Li;, Cost/Performance Tradeoff of n-Select Square Root Implementations, in 5th Australasian Computer Architecture, 2000.
6. Lachowicz S. and Pfleiderer H. J., Fast Evaluation of the Square Root and Other Nonlinear Functions in FPGA,. 4th IEEE International Symposium on Electronic Design, Test and Applications, DELTA 2008, pp. 474-477, 2008.
7. Kornerup, P. Digit selection for SRT division and square root. IEEE Transactions on Computers, Volume 54, Issue 3, pp. 294-303, 2005,

8. Buchmann J.A., Introduction to Cryptography (2nd ed.). Springer. 2005.
9. Kelly S. and Frankel S., Using HMAC-SHA-256, HMAC-SHA-384, and HMAC-SHA-512 with IPsec, Internet Engineering Task Force (IETF), May 2007.
10. Overmars A. and Venkatraman S., A New Method of Golden Ratio Computation for Faster Cryptosystems, IEEE Cybersecurity and Cyberforensics Conference (CCC), London, UK, 21-23 Nov. 2017.
11. Sudha K. R., Sekhar A. C. ,and Reddy P.V.G.D, Cryptography protection of digital signals using some recurrence relations, Int. J.of Comp. Sci. and Network Security, Volume 7, pp. 203-207, 2007.
12. Tahghighi M., Turaev S., Jaafar A., Mahmod R. and Md.Said M., On the Security of Golden Cryptosystems", Int. J. Contemp. Math Sciences, Volume 7, pp. 327 – 335, 2012.
13. Overmars A. and Ntogramatzidis L., A new parameterisation of Pythagorean triples in terms of odd and even series, Cornell University, arXiv:1504.03163 [math.HO], pp. 1-9, 2015.
14. Overmars A. and Venkatraman S., Pythagorean-platonic lattice method for finding all co-prime right angle triangles, International Journal of Computer and Information Engineering World Academy of Science, Engineering and Technology, Volume 4, Issue 11, pp. 1-4, 2017.
15. David F. and Eleanor R., Square Root Approximations in Old Babylonian Mathematics: YBC 7289 in Context, Historia Mathematica, Volume 25, Issue 4, pp. 366–378, 1998.
16. Tommiska M T, Area-efficient implementation of a fast square root algorithm, Proceedings of the Third IEEE International Conference on Devices, Circuits and Systems, S18/1-S18/4, 2000.
17. Takagi N, A hardware algorithm for computing reciprocal square root, Proceedings of the 15th IEEE Symposium on Computer Arithmetic, pp. 94 –100, 2001.