INTELLIGENT SERVER HARDENING SOFTWARE

HardnBot

Project Id: 19_20-J01

Project Proposal Report

R.M.B.B Rathnayake (IT16054400)

G.G.L Anjula (IT16022416)

W.M.K.M.W Wijekoon (IT16167742)

Aruna S.H.G.R (IT16099746)

B.Sc. (Hons) Degree in Information Technology

Department of Information Technology

Sri Lankan Institute of Information Technology

Sri Lanka

August 2019

INTELLIGENT SERVER HARDENING SOFTWARE

HardnBot

Project Id: 07

Project Proposal Report

R.M.B.B Rathnayake (IT16054400)

G.G.L Anjula (IT16022416)

W.M.K.M.W Wijekoon (IT16167742)

Aruna S.H.G.R (IT16099746)

B.Sc. (Hons) Degree in Information Technology

Department of Information Technology

Sri Lankan Institute of Information Technology

Sri Lanka

August 2019

DECLARATION OF THE CANDIDATE & SUPERVISOR

We declare that this is our own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student ID	Signature
R.M.B Rathnayake	IT16054400	
G.G.L Anjula	IT16022416	
W.M.K.M.W Wijekoon	IT16167742	
Aruna S.H.G.R	IT16099746	

Supervisor

……..

Mr. Amila Senarathne

ABSTRACT

Server Hardening is the one of the most important tasks to be handled on servers. Server hardening, which is also referred to as operating system hardening, is the process of making the server stronger and more resistant to security issues. Server hardening is an inexpensive and simple task to improve the overall operating system for maximum performance and to reduce expensive failures. Hardening is a Process requires many steps, all of which are critical to the success of the hardening system. The more steps a user follows, the safer and more resilient the system will be.

Using strong passwords, establish a password lockout policy, establish data encryption mechanism are sample tasks that are just the tip of the iceberg as far as server hardening is concerned. Many more tasks must be completed, and each is more complex. For proper execution and to maximum protection, professional assistance from an expert is needed.

Currently there is no fully automated server Hardening tool in information system. Although there are many vulnerability scanners and other auditing tools available for Application and Network audits, there are no proper tools for OS hardening.

This research is on building an automated tool which is capable of conducting a complete information systems compliance scan and classify compliance issues with a trained machine learning model. Which predicts the risk score efficiently and perform system hardening automatically to classified compliance issues utilizing minimum resources.

Keywords: hardening, Ansible, Linux, automate, audit, machine learning, rollback, abnormal

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Background

Massive amounts of data are created daily across the planet. By 2021, the annual global Internet Protocol (IP) traffic is predicted to reach 3.3 zettabytes. To match this huge data environment, the data center industry is anticipating unprecedented growth. Data is the most precious asset in data centers. Data centers require abilities to ensure data service works properly; many technologies are used in data centers to achieve this goal. Data centers are supported to run 24/7/365 without interruption. Planned or unplanned downtime can cause business users serious damage.

Most data centers include Linux servers; Ubuntu Server, Red Hat Enterprise Linux and CentOS. Datacenter includes about 200 live servers it is very difficult to do the operating system hardening manually.

There is no fully automated hardening platform implemented yet. Even the network administrators plan to do the Hardening processes manually it might take more than 6 hours for the complete harden processes for only for one and may contain lots of paper work. By the way there can be mistakes and faults in the hardening process that can effect the live Servers.

Sometimes a company may have to hire an external party to perform the hardening or they may have to outsource their systems to external organizations to assess their compliance. This will cost them in advance. Cost of maintaining compliance and governance may higher than the risks associated with these systems. And also there could be risks in outsourcing critical information systems.

1.2 Literature Review

Prior to our research project, we have conducted a Literature survey on the existing platforms with similar functionalities and technologies. Some of the relevant researches are reviewed here.

i. An effective modified security auditing tool (SAT)

In this research they have explained how to identify an exploitable vulnerability of an operating system via a security audit. This tool gathers much information from a remote hosts and network services such as ftp, NFS and according to those gathered information it will check for any security flaws, misconfigurations and other poor policy implementations that will put data at risk. As solutions, this tool can either report on this output data or it can use a rule-based system to investigate any potential security problems. However, according to this research, their main function of this tool is to iterate future data collection of secondary hosts using the initial data collection and user configurations for the next audit process. Furthermore, this tool can also analyze a complicated network and make practically informed decisions about the security level of the systems involves [1].

ii. Design and Implementation of Secure Auditing System in Linux Kernel

This research is about a tool to audit the kernel in a Unix based system. Although there is a log collection mechanism in Unix based systems, they are only based on application-level. A typical example for such subsystem is the “syslogd” daemon. It mainly receives important information of restricted services and process according to the configuration files. However, in this research, there main goal is to go beyond the typical user-state auditing and provide with a more detailed security audit result which contains both name of the system calls and related object of that. Since the current log files can be accessed, it will be a security issue and, in this paper, they also discuss about the security of all audit logs of this kernel auditing component as well. Later system administrators can view a completely in-depth detailed kernel state and user state logs for taking decisions for the system [2].

iii. A Design of Trusted Operating System Based on Linux

Time to time information security has become a research focus. The security of operating system at the base of information system. Because of that trusted operating system can help solve the information security problems.

A design process of trusted operating system based on Linux was developed by the china Standard Software Company (CS2C), and it’s Still researching furthermore, Double-key authentication and architecture provide in this project.

This operating system selects the improve/enhance method to implement Designing Method. The benefit of trusted operating system is to offer users a trusted computing environment [3].

<table>
<tr><td rowspan="3">User layer</td><td>shell layer</td><td colspan="3">shell program</td></tr>
<tr><td rowspan="2">utility layer</td><td rowspan="2">general applications</td><td colspan="2">trusted processes</td></tr>
<tr><td>expansion of the original procedures</td><td>new programs</td></tr>
<tr><td rowspan="2">Secure core layer</td><td>system call layer</td><td>security-related system calls</td><td>security-unrelated system calls</td><td>new system calls</td></tr>
<tr><td>core layer</td><td>security-related entities</td><td>security-unrelated entities</td><td>new entities</td></tr>
<tr><td>Hardware layer</td><td colspan="4">Hardware interface</td></tr>
</table>

FIGURE 1 : DESIGN PROCESS OF TRUSTED OPERATING SYSTEM BASED ON LINUX

Architecture of trusted operating system based Linux shown three layers of architecture, they are hardware layer in the bottom, secure core layer in middle and application layer in top. Architecture was explained in these categories.

- Sign and identification
- Discretionary Access Control (DAC)
- Mandatory Access Control (MAC)
- Least Privilege Management
- Audit
- Trusted path
- Trusted software

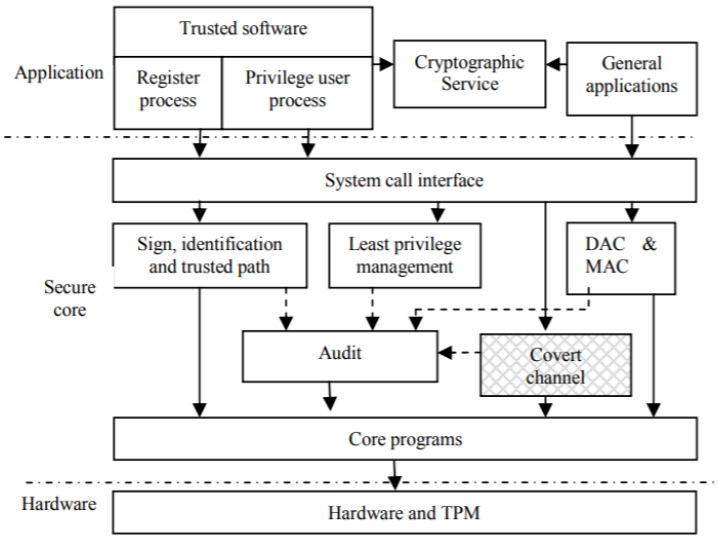

FIGURE 2 : ARCHITECTURE OF TRUSTED OPERATING SYSTEM BASED ON LINUX

- Double key authentication system is an application of the trusted operating system. It's a highly reliable and flexible authentication system for user permissions. User password and USBKEY both are used in it [3].

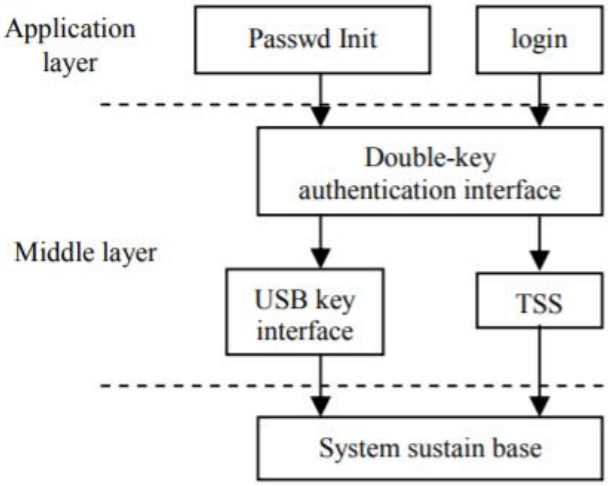

FIGURE 3 : DOUBLE-KEY AUTHENTICATION STRUCTURE

iv. System Hardening Architecture for Safer Access to Critical Business Data.

Now a day's cybercrimes are growing rapidly in significant amount. Therefore, those researchers present a research paper about system hardening and system hardening architecture. This will be a guide to system administrators for implementing multi-layers of in-depth protective mechanism over the stored data." System hardening is a strategy to increase security of the system, which applied many different security measures to different layers for detect vulnerabilities of the system layers and defeat them before it damages to the system. The proactive protective mechanism of system hardening architecture, applied to the host, application, operating system, user, and the physical layers. This system hardening security strategy can be implemented to the organization, to decrease the breaches and also safer access data.

Here they develop a system hardening architecture. Create security functions and combine them independently with relevant module. These functions are applied independently and separately, however here they try to implement their own security mechanism levels and use those mechanisms to relevant places which the relevant module located in the system. Because of that mechanisms if an attacker breaks the level one security he had to break number of security mechanism levels to access the relevant data. Because of the higher number of security levels mechanisms, attackers had spent much more time and also wasting resources they may give up and find some other efficiency way to do their malicious activities [4].

v. Prowse, D. (2010). *CompTIA Security+ Cert Guide: OS Hardening and Virtualization.*

At present security of operating systems has gain a major concern as the security misconfigurations of OS. It causes severe impact, when it exploited. In this paper they discuss about Operating system hardening. Operating system hardening is the concept of making OS secure by patches vulnerabilities manually or automatically. In this paper

they mainly consider windows operating system and how to make it secure by performing OS hardening using OS security audit method. In order to perform OS hardening, as a first step they perform vulnerability assessment over windows, then perform the security audit and fully analysis system logs. Further they explained how important it is to perform periodic security audit over windows OS in order to track vulnerabilities and take relevant countermeasures.

As benefit of doing OS hardening, they printed out how it helps to reduce the risk, improve the performance, eliminates vulnerable entry points and mitigate security risks. As this paper status OS hardening can be done using techniques such as program clean-up, service packs, patch management, group policies, see templates and configuration baselines. Further for more user friendliness windows provides facilities to prioritize vulnerabilities as high, medium and low.

To strengthen the security of OS, they discussed manual technology as well as semi-automated terms under manual techniques. Preparing checklist for security parameters, reviewing security configuration aspects, manually set security configuration and explaining OS as per configuration parameter included. In semi-automated way they are using scripts for audit such as .bat, .ps, set security configuration using script, exploiting OS scripted pay load. In discussion they showed how important it is to perform periodic audits to identify security issues, prioritize those and treat in order to mitigate risk over operating systems [5].

vi. Robotic Process Automation for Auditing. (2018). 15th ed. Robotic Process Automation for Auditing.

In this paper researchers have introduced RPA - Robotic Process Automation to the auditing in order to improve the quality of the audit. In auditing as it requires greater thinking skill, RPA would help in that as well as doing perfunctory task replacement. As they describe RPA is simply automating human task in order to higher the reliability, accuracy, efficiency so on. In auditing RPA can be used to automate tasks such as internal control testing, reconciliation, detail testing. They have discussed on current state of auditing which use automated tools like excel macros, Case Ware IDEA, scriptable languages such as python and R. Further they explained the process of implementing RPA in auditing phase by phase.

Once the implementation done and embedded RPA to the auditing, it provides a lot of benefits to the audit procedures. Mainly it improves the business value, compliance, process automation as well as service quality. Even though it provides many benefits, in other hand it has risky side too in security when it comes to cyber threats. Also applying RPA to auditing will result less human interaction where robots replace human job. Thus RPA will enhance auditing will lead hiring less human asset for organizations. Overall as they pointed out RPA will be bench mark on auditing in future [6].

vii. An automated approach for mitigating server security issues [7].

Many types of servers exist, such as mail server, web servers, application server, etc.., that store many sensitive information such as project details, media information, personal data, national security related information etc.., if such sensitive data gets in to wrong hands. Business and the reputation of the organization will be damaged. Therefore, need to automate security mechanism to detect, prevent and protect the server from the attackers. Security policies play an important role in network security and server security. An Automated Approach for Mitigating Server Security Issues proposes a framework that would ease the work of an administrator. It focuses on designing an automated tool which would perform an audit of the servers and check if it is compliant with all the prescribed security policies. As there are multiple platforms upon which the servers run, the tool is designed to adapt to heterogeneous environment.

This work was carried out at Hewlett Packard India Software Operations Pvt.Ltd., Hewlett Packard Enterprise reserves all rights to this work.

Related work

- Detecting the malware in Linux machine.
- Implemented a Security auditing tool for find the latest vulnerability.
- Introduced the network security technology.
- Discuss the various computer techniques.
- Describe the Linux, analyze well known weaknesses of Linux operating system.

Basic Server security step

- Every organization should have their security policies defined.
- There should be an apt network protection mechanism such as the firewall technology, anti-virus technology, Intrusion Detection System (IDS), Virtual Private Network (VPN) and data encryption technology.
- Use of automated tool that would keep track of server's security policies and their compliance.
- Use of secure administration and maintenance processes, which includes application of patches and upgrades, monitoring of all the logs, backing up of the server data and operating system.
- Installation and configuration of secure operating system and software in the server.
- Use of vulnerability scanner to perform security testing.
- Malware detection, mainly during insertion of infected devices through USB.
- Configuring access control.

Proposed Work

- Checking if the Windows server is running the approved up-to-date anti-malware solution.
- Checking if the approved antivirus shield is seen in the taskbar for a system running Windows.
- List the version of the Anti-Malware running on the systems for Linux and Windows operating system.

- Check if any personal removable media present for both Linux and Windows Server, if so list their names and timestamp of the device insertion.
- List all the security patches applied to the Windows operating system.
- List the installed versions of software running on the Windows System.
- Check if the event logs have been enabled or disabled.

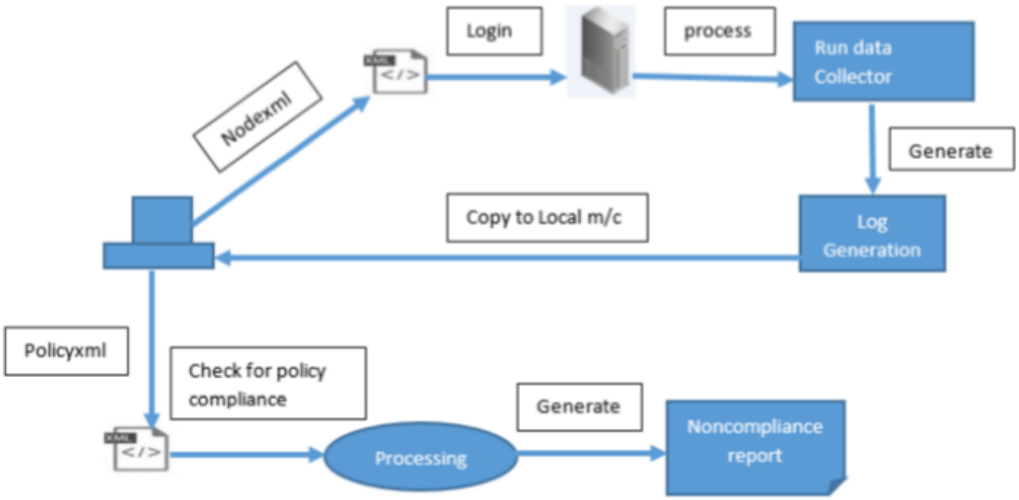

FIGURE 4 : PROPOSED SECURITY COMPLIANCE TOOL

There will be two XML files. It's Containing IP address, type of the server, login credentials, policy details. The tool will logins to the server and collect the data of server Using the XML file
And checked for policy compliance and a non-compliance report will be generated.

The threats must be analyzed and must be recoded to the prevent future Cyber-attacks. We must take right decision to protect the servers in Difficult situations. Security must be given highest priority. The aim of effective policy compliance tool to check if the servers were in compliance with all the policies, and help he administrator to fix the loopholes and keep them secure [7].

viii. A Remote Backup Approach for Virtual Machine Images [8].

When we are considering cloud computing, virtualization is playing a major role, because of hosting several applications and services in virtual machines (VM) which were hosted in cloud environments. Security become a prior requirement in virtualized applications. In this research, mainly focused area is high availability issue in virtual machines. LiveRB (Live remote backup) is the proposed remote backup approach. The purpose of the Live RB is to save the running state of the VM in an online manner known as "Live Migration". This backup process will happen the background of the hosted cloud applications of the VM and is transparent to them. A virtual block device will be designed and will be used to cache of I/O Operations in memory, in order to save the incremental virtual disk data.

LiveRB will be implemented on KVM virtualization platform in order to evaluate effectiveness and efficiency using a set of comprehensive experiments. These experiments are all related to Cloud Computing and the security issues that come along with this and the key points considered in order to have successful cloud computing are security, availability & fault tolerance. The commonly used solution to handle Fault Tolerance & High Availability is using snapshots or checkpoints that periodically record the states of the software for backup and rollback the cloud applications to the previously backup up state. This procedure will be carried out when encountering Failures or Errors of the original system.

Most currently existing VMs stop the VM to take snapshots. Some VMs need to be shut down too take snapshots which this affects the ability to provide the service/ result in abnormal cloud application behavior. Some VMs suspend the current process and save the current progress onto local disks to be transferred onto remote servers later which sometimes result in data loss if a hardware failure is encountered.

The above issues can be resolved using the Live RB since it works by not stopping the VM to do the backup process. Results of this process indicate that Live RB can be used on a VM to do the backup task from VM onto a Remote Server with only a slight reduction in performance

ix. An Adaptive Technique for Dynamic Rollback in Concurrent Event-Driven Fault Simulation [9].

In here it is discussing about automatic rollback based on an adaptive mechanism which is including advanced network/system status recording system. Time can be any time, that mean before changing of a system or after changing a system this status recording can be apply. Main feature of this research is user can define the rate for maximum acceptable level for rollback. This approach takes the average time to minimum level, that means very short time of rollback process.

To come up with proposed technique, researchers were used existing methods such as incremental backups, journal files, checkpoints, rollback, roll forward which were found on different applications, different operating systems as well as different databases. Mainly the status of the network/system is record on disk and run for negative time period to analyze previous status. If needed user can run for a positive time period as well. Those time periods are for compare with current status of the network/system. To make it happens above approaches need some fine tunes as well.

x. Research and Implementation of Data Storage Backup [10].

With use of applications which were depend on big data, the usage of data storage backups was became more important. Therefore, the methods used to backup should be more flexible and can be able to ensure of security and reliability of backup contents and also backup and restore should be in a convenient manner. There are several backup methods such as data backup, system backup, application backup etc. The backup contents are guaranteed to be confidential, complete and effective.

There are several specific performances in a backup,

I. Backup should be upgradable, capacity expansion

II. Management without affecting other application in the system
III. Implement a backup storage system combining SAN (storage area networks) and NAS (network attached storage) storage networks.
IV. Provide several backup methods such as data backup, system backup, application backup,
V. Backup contents should be secure and restore operations should be done in a convenient manner.

System backup

Reefers to the backup of the end-point operating system, server operating system and other systems. In here core files and system's registry are backed up as a data. In a matter of system crash or operation mistaken the backup can be restoring to the previous state.

Virtual tape library

Virtual tape library (VTL) considered as a world's leading modern technology to create a backup system. It can rapidly backup and rapidly recover a system that we want to backup. Main feature is no manual intervention of this technology. VTL storage media is a SATA disk and its data transfer rate is 150MS/s. That means approximately it takes 10 seconds for transferring 1.5GB data to the backup storage.

xi. A Cost-effective Dependable Microcontroller Architecture with Instruction-level Rollback for Soft Error Recovery [11].

This tool is developed for detect soft errors using electronic design automation (EDU) which generates optimized soft error detecting logic circuits for flip-flops. When a soft error is detected that signal goes to a developed rollback control module (RCM). That RCM will reset the CPU and restores the CPU's register file from a backup register file using a rollback program guidance. After that CPU will able to restart from the state which is before the soft error occurred. In here researchers were developed another two modules called error reset module (ERM) that can restore the RCM from soft errors and error correction module (ECM) that corrects errors in RAM after error detection with no delay overhead. In above mentioned soft error means, which are random transient errors. Those errors are the main cause of failures in microcontrollers which include reversal of a memory element's bit data due to factors such as alpha rays in a package, neutron strike and noise of the environment.

xii. Machine Learning Techniques for Predicting Web Server Anomalies [12].

The basic idea between servers on the web is to provide requests made by the client through the web using different transmission methods such as Services. Businesses relying on these services require the web servers to have reliability, availability and security in order to provide constant quality in the service provided. This document describes the quality ensured in these services.

The assumption made for this problem is mainly due to Resource Starvation. Resource Starvation is when a process that functions in Concurrent Computing is unendingly denied the necessary resource to continue & process the rest of its work. Resource Starvation is measured by the response time taken to cater requests under artificial workloads while collecting data on other resource parameters. The research provides proof that these recordings gathered from different artificial workloads can be applied to real world entities as well

Machine Learning is used to monitor & correlate the high response time and this is done by observing the system data. The goal of this analysis is to resolve issues of this variety in Web Servers, Operating Systems or in VM (virtual machine) Rejuvenation.

Based on the statistics provided by the Internet World Statistics, we could clearly notice a rapid rise QoS (quality of service) Internet Service Usage users and this gave several companies & industries to exist in the current world. The below listed out Companies/ Industries who gets affected by these figures since their prime business is offering Internet QoS,

- Cloud Computing
- Data Storage
- Hosting Providers
- Content Delivery
- Application Performance Management & other

Due to this high demand and dependence on network QoS, it is important for a particular service to be aware of its own deteriorating quality. Currently there are several self-monitoring network products that ensure that the QoS of services offered through the internet. The goal of this this research is to increase this area.

The benefits taken from this research can be applied to other areas as well and they have been listed down below,

- Proactive Software Rejuvenation
- Web Server Workload Balancing
- Web Server Performance Testing

xiii. Vulnerability Profile for Linux [13].

In this research, they talk about profiling identified vulnerabilities according to the CVE score of them. In their classification scheme, they consider four types of classification schemes namely,

1. Confidentiality violation
2. Integrity violation
3. Availability violation
4. System compromised

If a confidentiality violation occurs, it allows an attack to directly steal information from the system. Integrity violations allows an attack to directly change the information passing through the system. Availability violation results an attack that

limit the genuine access to a genuine user (human or machine), Denial of service attacks (DOS) can be taken as an example. According to their research system compromised attacks gives the attacker the privilege to access the system in four different levels such as: run an arbitrary code, elevate privilege, account break-in, and finally root break in which can be the worst-case scenario.

Furthermore, these classifications are again grouped according to the severity level.

Damage Type	Severity Level		
	High	Medium	Low
Confidentiality	- Disclosure of information and system configuration in root/super user level	- Disclosure of system information and configuration in user level	- Disclosure of some no relevant information.
integrity	- Information and system configuration changed in root/super user level	-Information changed in user level	- Non-relevant information changed in other user level
availability	-Whole system crash or unavailable	- Some services unavailable -System temporary unavailable	- Some services temporary slow down with flooding
System compromised	-Root break-in -Account break-in -Run arbitrary code by root/super user privilege	- Privilege gain in some domain - Run arbitrary code by user privilege	- Run arbitrary code by other user privilege

xiv. Automatic Classification for Vulnerability Based on Machine Learning [14].

This research paper is based on vulnerability classification using machine learning methods based on LDA model and SVM. Word location information is introduced in to LDA model called WL-LDA (Weighted location LDA), which is somewhat better than typical language processing algorithms because it generates outcomes from vector space on themes other than on word, and a multi-class classifier called HT-SVM (Huffman Tree SVM) is developed that it could make a faster and more stable classification on the vulnerabilities.

The main idea of this research is that they classify vulnerabilities with new models based on existing LDA and SVM models and obtain more accurate and more effective outcome.

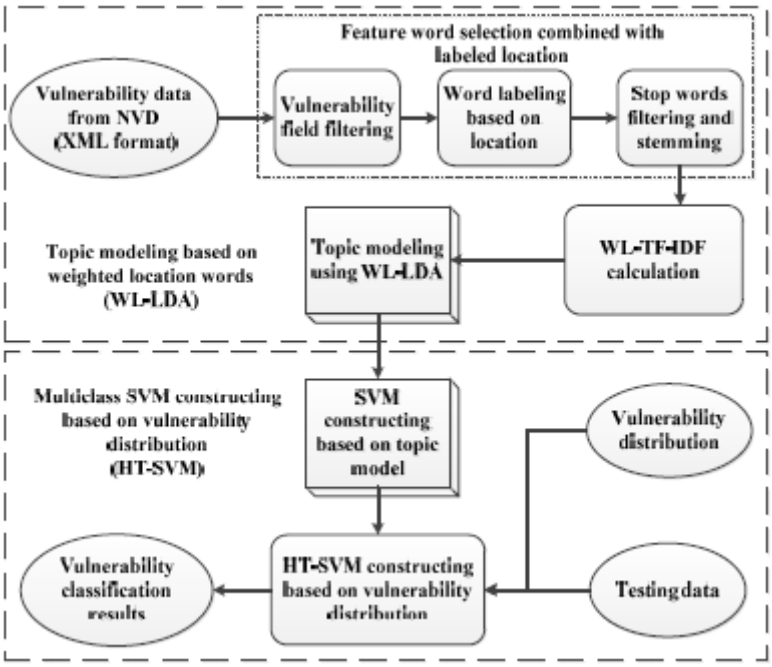

FIGURE 5 : WORK FLOW

xv. Password Strength Prediction using Supervised Machine Learning Techniques [15].

This research is targeted on the password strength of a system and predict password strength of a system whether it's a strong password or a weak password using supervised machine learning techniques such as classification (discrete) and regression (continuous).

Here the password strength prediction is modeled as classification task and supervised machine learning techniques were used as mentioned above. In this research they mainly used some commonly used classification models such as

1. Decision tree classifier
2. Multilayer Perception
3. Naïve Bayes Classifier

4. Support Vector Machine (SVM)

For testing and select the best classification model for the performance of the task. After some performed tests, they identified Support Vector Machine (SVM) as the most suitable model for this task.

xvi. Vulnerability Severity Prediction with Deep Neural Network [16].

Multiple deep learning methods for vulnerability text classification evaluation are proposed in this research paper. Three kinds of deep neural networks,

1. CNN,
2. LSTM,
3. TextRCNN

and one kind of traditional machine learning method

1. XGBoost

are used. Here all parameters tuned via experiments to improve the accuracy of the task. However, in this research they said that the deep neural network methods evaluate vulnerability risk levels better, compared with traditional machine learning methods but it costs more time to train. This research says that they scored 93.95% accuracy level when training the model.

The flow of this task is below.

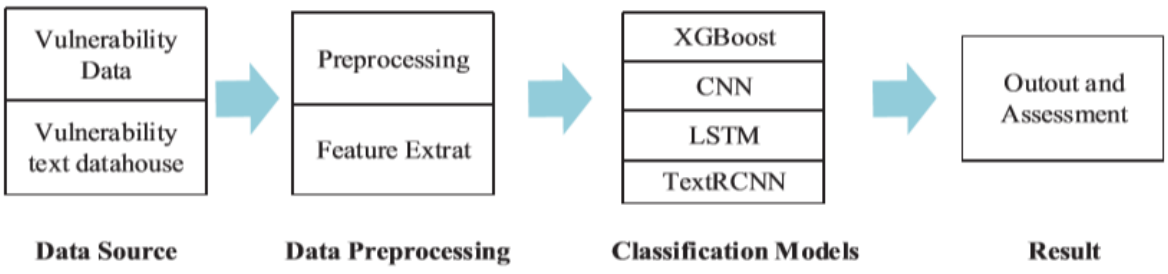

FIGURE 6 :WORK FLOW OF DEEP NEURAL NETWORK

xvii. Research of Information Security Risk Prediction based on Grey Theory and ANP [17].

Risk prediction is an important part of the information security system. In accordance with the information security risk assessment process and combination of assets, threat, vulnerability and safety control measures, to strengthen the correlation among these factors and make the prediction results more objective for the target, the authors put forward a model based on the combination of the grey theory and analytic network process(ANP) with information security risk prediction. Establish the weight of each

risk assessment element through the analytic network process (ANP) by analyzing interdependency and feedback. Finally, set up systematic risk fuzzy comprehensive calculation to process data and build accurate mathematical model by combining with the risk assessment level.

firstly, the authors grasp the development law of information system through the processing of raw data and the establishment of the grey model, and confirm the preliminary scientific quantitative prediction for the system's future state; Secondly, use the network analysis method of ANP to compare each independent elements, so that the authors are able to calculate the weight value of each risk factor which affects the system security, reorder the weights, and propose more targeted and objective improvement measures; Finally ,combining with the weight value, to analyze risk objects, the authors obtain fuzzy membership matrix of judgment matrix and build the fuzzy mathematical model, calculate the value of the risk factors comprehensively, and treat it as the guidance, so that reliable guarantee for information system security can be provided. The model realized the grey theory prediction model and was applied in the field of information security, calculate accurate comprehensive weights of various risk factors in information system. In the system, internet elements are interdependent and give feedback to each other, thus combining the theory of fuzzy mathematics, satisfying the requirements of the objectivity and complex of information system, forecasting result is scientific and accurate, instructive significance as well.

xviii. Information Security Risk Assessment and Management Method in Computer Networks [18].

This suggested a method for quantitative information security risk assessment and management in computer networks. This process evaluates an impact and possibility value for specific threats using fuzzy logic and analytic hierarchy process to evaluate. Using fuzzy rules and fuzzy interference system, evaluation vulnerabilities under the uncertainty.

Consider such types of assets - information, host, servers, and telecommunication equipment, IT-services (confidentiality, integrity and availability)

Consider three groups of external socio-political impact, internal impact, and direct financial losses. Most of these are qualitative, thereby they use the analytic hierarchy process for their quantitative evaluation. Evaluate priority weights of information asset regarding to confidentiality, integrity or availability.

Possibility Evaluation for Specific Threat

They suggest a method for quantitative evaluation of threat’s exercising possibility, which is based on the questionnaires. These questionnaires include questions about possibility factors for specific threat and some possible answers to these questions. Answer for every questions. After the answering all questions assign number of points and they will assign possibility of the threat.

Vulnerability Evaluation

They suggest a few methods for vulnerability risk assessment. It is based on common Vulnerability Scoring system (CVSS).and they suggest new vulnerability assessment method based on expert judgments, fuzzy production rules and fuzzy logic.

Risk Assessment

They assess the information security risk for specific threat and specific vulnerability by following way.

Risk (Threat) = impact (threat). Possibility (treat).RL

xix. Using a Prediction Model to Manage Cyber Security Threats [19].

Cyber-attack is an attempt to exploit computer systems and networks. Cyber-attacks use malicious codes to alter algorithms, logic, or data. Securing information systems is thus critical. Multiple countermeasures need to be built The CVSS is an industry framework that helps quantify the vulnerability impact. This paper demonstrated a mathematical model to predict the impact of an attack based on significant factors that influence cyber security. Vulnerability and network traffic were selected as the influencing factors to predict CVSS score. Based on the score, the technical analyst can analyze the impact and take necessary preventive actions. This model also considers the environmental information required. It is thus generalized and can be customized to the needs of the individual organization.

TABLE I: Project data points.

Y CVSS Score	*X*1 Vulnerability	*X*2 Network Traffic
2.1	20	324
5.3	53	623
1.0	15	235
8.0	85	932
2.9	28	438
3.0	25	498
3.8	38	391
1.0	18	132
1.2	16	177
5.9	63	823
4.3	39	579
2.8	30	455
1.1	14	231
4.2	35	725
5.4	51	740
1.9	21	345
2.0	25	432
4.1	37	467
6.2	58	845
1.1	15	111
2.3	22	191
1.2	16	182

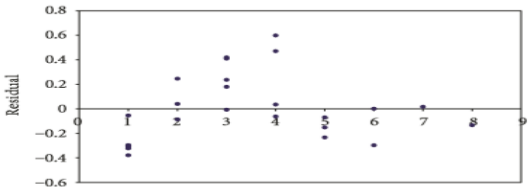

FIGURE I: Residual plot.

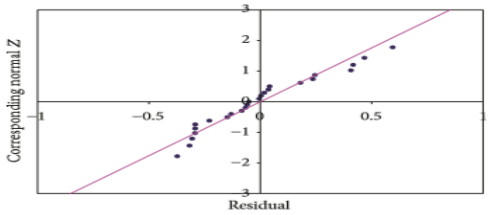

FIGURE 2: Normal probability plot.

FIGURE 7 :CVSS SCORE PREDICTION

Y is the overall CVSS score. CVSS is the predicted based on the environment and system characteristics of the target application. X1 is the number of vulnerabilities, namely, the total number of vulnerabilities detect by the static and dynamic vulnerability detection tool for target application.X2 is the average input network traffic.

In this regression model, CVSS score predict by the using two variables network traffic and vulnerability. Vulnerability and network traffic have no influence over CVSS score. No mirror pattern can be found (residual plot). Probability plot shown in figure 2is approximately linear. CVSS score is impacted positively both vulnerability and by network traffic.

Predicted CVSS score = intercept + Vulnerability * number of vulnerabilities +network traffic *average input networks

Intercept, vulnerability, network traffic can be calculate using regression equation.

xx. Quantitative Assessment of Cyber Security Risk using Bayesian Network-based model [20].

This paper proposes a quantitative model for assessing cyber security risk in information security. The model can be used to evaluate the security readiness of firms in the marketplace through qualitative and quantitative tools. We propose a Bayesian network methodology that can be used to generate a cyber-security risk score that takes as input a firm's security profile and data breach statistics. The quantitative model enables cyber risk to be captured in a precise and comparable fashion. The objective of the scoring model is to create a common reference in the marketplace that could enhance incentives for firms to invest and improve their security systems. This paper concludes with a demonstration of scoring an intrusion detection network.

The Scoring mechanism determine from questionnaires are generated, the network is complete in both its qualitative and quantitative assessments. The scoring mechanism proceeds with a series of calculations to determine the score of a higher child node and similarly to the resource-driven security score.

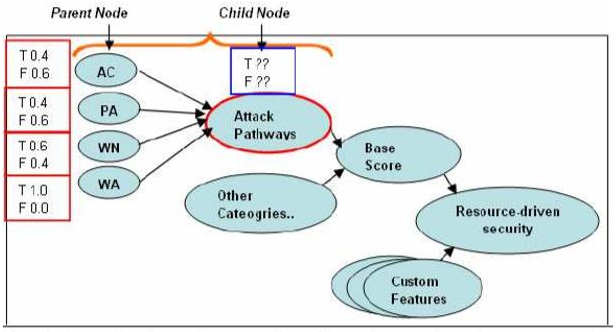

Fig. 8. Intrusion detection network to determine attack pathway prevention sub-score.

FIGURE 8 : INTRUSION DETECTION NETWORK PREVENTION PATHWAY

Scoring mechanism perform using Bayesian methodology and probability theorem.

1.2.1 Products Available in the Market

I. Lynis Software

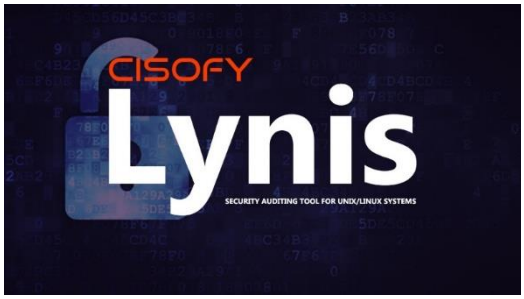

FIGURE 9 : LYNIS SOFTWARE LOGO

Lynis is a battle-tested security tool for systems running UNIX system, macOS, or Unix-based software. It performs an in depth health scan of your systems to web hardening and compliance testing. The project is open supply package with the GPL license and offered since 2007

- Security auditing
- Compliance testing (e.g. PCI, HIPAA, SOx)
- Penetration testing
- Vulnerability detection
- System hardening

.

II. Yolinux

FIGURE 10 : YOLINUX LOGO

Security configuration and set-up for UNIX operating system servers exposed to the web: Any laptop connected to the internet would force steps and precautions to be taken to scale back the exposure to hacker threats. Web, mail and DNS servers are particularly vulnerable. Massive operations can hide behind a CISCO firewall for many of their protection. The UNIX operating system, server should be designed for network security and have its applications and services configured for security. This tutorial covers the steps and tools which might want to monitor and counteract hacker threats. Simply put;

- web server security.

- Secure Shell (encrypted telnet session)
- block repeated failed logins
- Restricted shell for use with OpenSSH sftp (optional chrooted account)
- monitoring network probes and hack attacks and stopping them.
- security monitoring your system for changes and questionable files.

Performing A Security audit:

- Hunt for Trojan commands, worms and known exploits
- Performing a network vulnerability scan/security audit of your system.
- Performing a vulnerability scan/security audit of a WordPress site (core and plugins).

III. Open-AudIT

FIGURE 11 : OPEN-AUDIT LOGO

Open-AudIT is meant to be run on a server (Windows or Linux) and to scan your networks for devices. Once a tool is found, Open-AudIT runs a series of commands upon it and stores the following data during a} very info. This data is then gettable for varied coverage functions. Open-AudIT comes with a list of over fifty reports with any kind, of any reports able to be created by the user.

- Network Discovery
- Computer Auditing including software, hardware, configuration.
- Configuration Change Detection and Reporting
- Geographical Maps
- File Auditing

IV. Pentana Software

FIGURE 12 : PENTANA SOFTWARE LOGO

Ideagen created the Pentana automated audit software. This is suitable for businesses of all sizes and all industries. This isn't support only for Windows operating systems. It also supports for web-based systems. (Windows 7, Windows Vista, Windows XP, Web browser (OS agnostic), Windows 2000, Windows 8, Windows 10)

Key features of Pentana,

1. Compliance management
2. Exceptions management
3. Issue management
4. Risk assessment
5. Internal controls management functionalities

V. Netwrix auditor

FIGURE 13 : NETWRIX AUDITOR LOGO

Netwrix free Netwrix Auditor 8.0. The new version of the IT auditing platform simplifies detection of security threats and enables organizations to realize management over critical data across all levels of IT environment, including hybrid cloud and storage appliances.

- File/folder access audit
- File/folder permission change audit
- Local user logon auditing
- File integrity monitoring
- Video recording of user screen activity

1.3 Research Gap

The following chart shows the comparison between existing solutions or application and our proposed solution "**Automated audit & hardening solution**". Even through most of problems we identified doesn't addresses in the existing solutions.

TABLE 1 : COMPARISON BETWEEN EXISTING SOLUTIONS

Features	Lynis	YoLinux	Open Audit	Pentana Software	NetWrix	SolarWinds	Audit Prodigy
No need to Expert knowledge for doing the audit	*	*	*	*	*	*	*
Generate detailed reports	*		*		*		
Risk Assessment				*			*
Monitoring Network Devices		*	*			*	
Backup Functions for all the settings							
Compare with previous reports							
Mobile Access							*
Vulnerability detection	*	*					

1.4 Research Problem

Massive amounts of data are created daily across the planet. By 2021, the annual global Internet Protocol (IP) traffic is predicted to reach 3.3 zettabytes. To match this huge data environment, the data center industry is anticipating unprecedented growth. Data is the most precious asset in data centers. Data centers require abilities to ensure data service works properly; many technologies are used in data centers to achieve this goal. Data centers are supported to run 24/7/365 without interruption. Planned or unplanned downtime can cause business users serious damage.

Most data centers include Linux servers; Ubuntu Server, Red Hat Enterprise Linux and CentOS. Datacenter includes about 200 live servers it is very difficult to do the operating system hardening manually.

There is no fully automated hardening platform implemented yet. Even the network administrators plan to do the Hardening processes manually it might take more than 6 hours for the complete harden processes for only for one and may contain lots of paper work. By the way there can be mistakes and faults in the hardening process that can effect the live Servers.

Sometimes a company may have to hire an external party to perform the hardening or they may have to outsource their systems to external organizations to assess their compliance. This will cost them in advance. Cost of maintaining compliance and governance may higher than the risks associated with these systems. And also there could be risks in outsourcing critical information systems.

2 OBJECTIVES

2.1 Main Objectives

The aim of this research is to find an effective way to secure the datacenter RHEL 7,6 and CentOS 7,6 servers and enhance productivity for the customers and employees with a low cost and few human interactions.

Hence there is no any fully automated hardening platform implemented yet, in this system implementation a novel automated open source software is proposed. This research focuses to find an automated way to perform all Level 1 configuration profiles of CIS benchmark with Ansible automation tool.

Here fully automated free open source hardening platform is developed with the capability to detect poor or non-compliant configurations in a system (OS/DB/Application) and applying industry recommended fixes/configurations and secure systems by reducing its surface of vulnerabilities.

This research composes a great business value as it can cover 90% of an Information Systems Audit and hardening process. For internal audit, this software presents both opportunity and responsibility. By helping the organization understand and control risks and identifying opportunities/industry best practices to embrace. This software also will allow internal audit to position themselves as trusted advisors.

i. **Perform scans to detect compliance issues.**

   HardnBot perform scans to identify misconfigurations in operating systems, databases. The software will run commands and queries to retrieve configurations for relevant locations and record them. These configurations will be compared against pre-defined configuration libraries and resources which contains defective values will be identified. Then these configurations will be presented to the user in a well formatted way.

ii. **Classify compliance issues according to a risk score.**

   Compliance issues are classified in to categories named critical, high, medium and low with a machine learning trained model. Display the count of each severity.

iii. **Predict the overall risk score of the server.**

   Using the classified compliance issues an equation/algorithm is developed to predict the overall risk score of the server. This will allow users to predict the asset value to be threaten.

iv. **Backup**

   System configurations should be backed up before introducing any change to them. Otherwise if any failure occurs after applying a fix or if there is a risk associated with the particular fix, it will leave the system vulnerable. Through this software, we provide backup capability which will save the existing system configurations. These configurations can be loaded into the system if any failure occurs.

v. **Implement rollback function when abnormal behavior is detected.**

System will automatically detect any abnormal behavior in the system after the hardening is carried out. If any anomaly detected rollback function will take place and proceed to established earlier status (backup) of the server.

vi. **Automatically harden the system to classified compliance issues.**

System generates hardening scripts to correct the classified compliance issues and automatically execute in particular server. System owner doesn't have to access Unix environment anymore. The hardening configurations will be displayed to users with appropriate parameters.

vii. **Bring industry** **Bring industry best practices into system.**

For a particular parameter in the system configuration, there is an industry recognized value. For example, a password should expire at most in 90 days. A company may not adhere to these standard values; their security policies may not describe them. In such occasions, system administrators may have assigned them with default values. Through our software, we plan to introduce industry accepted values and configurations into the information systems.

2.2 Specific objectives

i. **Reduce manual work**

Through this software, we target to reduce the manual effort needed for a complete system hardening. Almost all the parts of the audit and hardening could be performed through this software. Compliance issues are automatically classified using a trained machine learning model and risk score is predicted. Reports can be generated through this software at the end of the hardening process.

ii. **Ease the internal audit**

This software is designed in a way that it could be easily used by non-technical people. Simple interfaces, pop-up guidelines and descriptive reports will make the user aware of the functionalities of the software. Even before applying a fix to an issue, user can read a full detailed report of that issue including its impact and the remediation.

This software package can be handled by a single user. Therefore, the entire system hardening process can be conducted by a single user. This requires minimal amount of organization's resources. As for the internal audit, this software will cover up to 90% of their work.

iii. **Rollback the backup when abnormal behavior is detected.**

Through this software System will automatically detect any abnormal behavior in the system after the hardening is carried out. If any anomaly detected rollback function will take place and proceed to established earlier status (backup) of the server.

3 RESEARCH METHODOLOGY

1. **Design formatted configuration libraries/Design Solution structures/libraries.**

   Formatted configuration libraries contain configurations that should be in a particular operating system or a database. Basically it is a collection of parameters and their industry accepted values. When a user runs the scan, the software will load both system configurations and these formatted configuration libraries into the memory. System configurations will be compared against these configuration libraries. If the system configurations differ from these libraries, they will be identified as misconfigurations.

   Solution libraries contain the implementation needed to change the detected misconfigurations. It is basically the coding to manipulate system values. When user clicks on the fix button for a particular issue, these libraries will be executed and the scripts on them will fix the related issue.

2. **Design OS commands/Queries.**

   Commands and queries are issued for a particular OS or a database in order to retrieve details of their configuration settings. These commands and queries vary with each OS and database. Therefore, separate sets of commands and queries should be designed and tested prior to the implementation of this audit platform.

3. **Implement scans to detect non-compliant parameters in OS/DB components**.

   Then scans will be conducted on OS/DB components to retrieve their existing configurations. Once user run the scan, the software will detect the OS/DB version installed on the host and it will select the relevant set of commands and queries needed to retrieve the configurations. Then these configurations will be compared against formatted configuration libraries in order to detect misconfigurations. Detected issue will be listed with the options to “see details” and “fix”.

4. **Implement backup function.**

   Existing system configurations should be taken into backups as a precaution if the new configurations failed. For this purpose, scripts will be designed for each OS/DB components. These scripts will generate a single backup file which can be used to restore in case of applied fixes failed.

5. **Implement Intelligent Rollback function.**

   Rollback function will take place after the hardening process is done. This function basically depends on the abnormal behaviors of the server which are appearing after the hardening process. In here there are several pre-define behavior models to detect those kind of abnormal behaviors. For that purpose, we create our own models as well as we can use existing models in various tools such as NAGIOS for this task. If those models detect any anomaly regarding to server services, there is a rollback script to run for establishing previous status (backup) of the server and it will automatically run.

6. **Risk score prediction**

   Using classified compliance issues, Overall risk score of the server is predicted. For this purpose, it is necessary to implement an algorithm to generate risk score. Using this algorithm HardnBot will predict the likelihood of Impact and probability of occurrence, to implement this algorithm, common probability functions and Bayesian methodology will be used. Potential Risk and Total asset value used to implement the Probability and impact of the risk. Using this method, the asset value to be threaten can be displayed.

7. **Implement methods to apply fixes.**

   When user clicks on the fix option for a particular issue, software will load the correct configuration values from solution libraries and these values will be replaced on the detected misconfiguration values. For this purpose, admin privileges will be required from a particular system. Therefore, privilege management will be handled and fixes will be applied one by one, logging every change that this software does to the system.

8. **Implement follow up sheets to track audit progress.**

   The follow-up sheets will be used to track the audit progress. For a particular system, a digital follow-up sheet will be generated which contains a checklist to audit that system. User can record the status of the audit in these follow-up sheets and generate a score at the end. These follow-up sheets can be used to identify what parts are successfully configured and what parts left to be configured. These follow-up sheets also provide guidance to audit physical components of a system. This will enable a complete system auditing process.

9. **Track performance of applied fixes and logging.**

   Effectiveness of the applied fixes should be monitored. Because if the new fix causes system to malfunction or it leaves the system vulnerable, the entire audit process will be a failure. Methods will be implemented to track the changes done through this software. Every single change occurred through this software will be logged and this log could be referred to identify any issue arises along with the date, time and user.

3.1 System Diagram

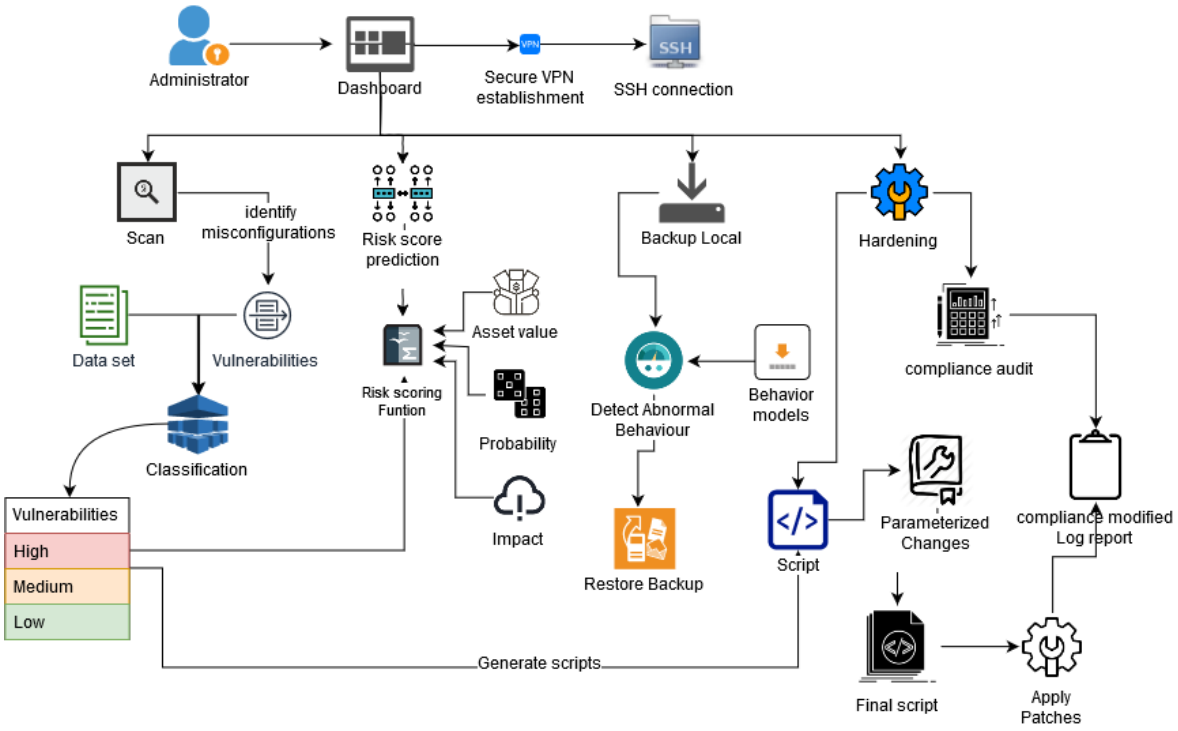

FIGURE 14: SYSTEM DIAGRAM

3.2 Tools and Techniques

Tools

- Microsoft Visual Studio IDE
- VMware Workstation
- PuTTY

Techniques

- C#
- Ansible
- Python
- Perl
- Shell Script

3.3 Testing

- Unit Testing – Each unit of the system will test by the group member who is developing that particular unit and will produce a defects free unit of coding.
- Component Testing – Several bug free units are combined together and tested. Each member combines their tested units together and test them.
- Integration Testing – In this testing level users are responsible to test whether the relationships and communication between tested components are working as expect.
- System Testing – All the components from each group member will combine together and test the whole system to verify the functionality and the performance of it.

3.4 Gantt Chart

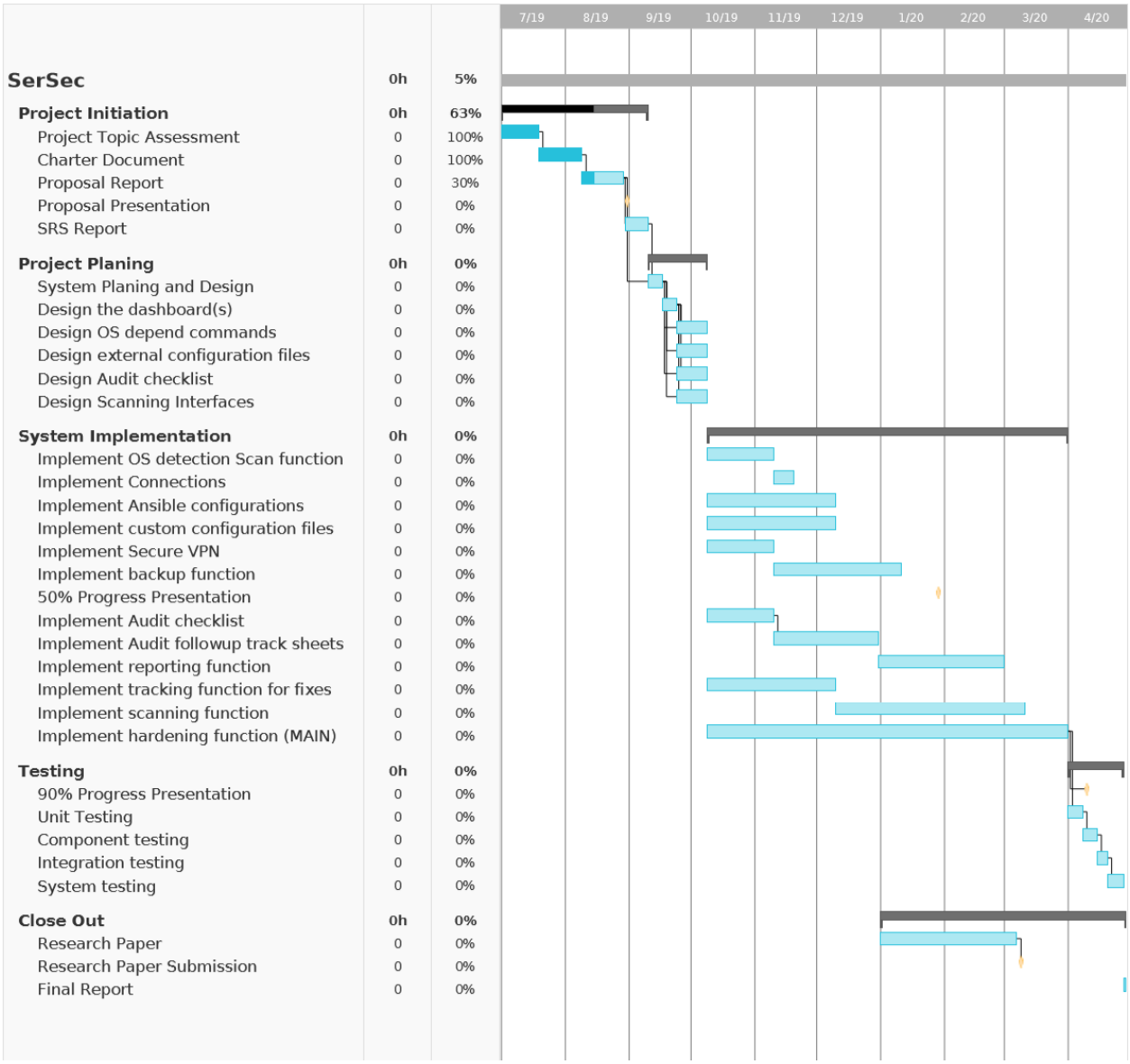

FIGURE 15 : GANTT CHART

3.5 Work Breakdown Structure

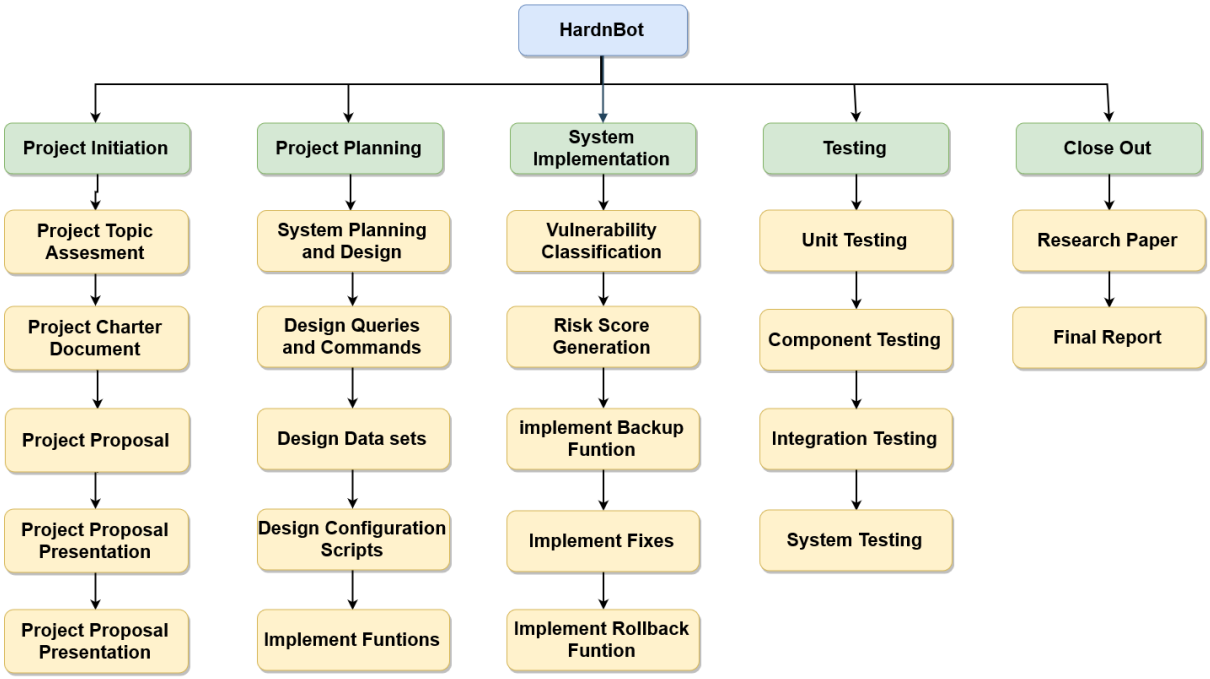

FIGURE 16 : WORK BREAKDOWN STRUCTURE

4 DESCRIPTION OF PERSONAL AND FACILITIES

TABLE 2 : DESCRIPTION OF PERSONAL AND FACILITIES

Member	Task
R.M.B Rathnayake	• Design interface for hardening and Implement hardening function. • Compliance audit. • Implement functions and interfaces to pass parameters and generate hardening scripts. • System integration.
G.G.L Anjula	• Setup SSH connection to servers through secure VPN tunnel. • Implement backup function. • Implement rollback function when abnormal behavior is detected.
K Wijekoon	• Design interface for Scanning. • Scan the server and identify failed compliance issues • Classify those compliance issues according to a risk score (Critical, High, Medium, Low) • Display overall risk score with percentages.
Aruna S.H.G.R	• Using classified compliance issues, Predict the overall risk score of the server • Display the asset value to be threaten. • Design interface for Reporting.

REFERENCES

[1] A. M. B. Mohamed, "An effective modified security auditing tool (SAT)," in *IEEE*, Pula, 2001.

[2] K. Zhao, Q. Li, J. Kang, D. Jiang , L. Hu, "Design and Implementation of Secure Auditing System in Linux Kernel," in *IEEE*, Fujian, 2007.

[3] H. Li , Y. Lan, "A Design of Trusted Operating System Based on Linux," in *IEEE*, Wuhan, 2010.

[4] Ibor, A. , Obidinn,J., SYSTEM HARDENING ARCHITECTURE FOR SAFER ACCESS TO CRITICAL BUSINESS DATA, Nigerian Journal of Technology, 2015.

[5] Prowse, D., CompTIA Security+ Cert Guide: OS Hardening and Virtualization, 2010.

[6] Robotic Process Automation for Auditing., 2018.

[7] S. Patra, N. C. Naveen , O. Prabhakar, "An automated approach for mitigating server security issues," in *IEEE*, Bangalore, 2016.

[8] Zhe Wang, Jin Zeng, Tao Lv Bin Shi, Bo Li, "A Remote Backup Approach for Virtual Machine Images," in *IEEE*, 2016.

[9] L. Farinetti, P. L. Montessoro, "An Adaptive Technique for Dynamic Rollback in Concurrent Event-Driven Fault Simulation," in *IEEE*, 1993.

[10] Ning Lu, Yongmin Zhao, "Research and Implementation of Data Storage Backup," in *IEEE*, 2018.

[11] Teruaki Sakata, Teppei Hirotsu, Hiromichi Yamada, Takeshi Kataoka, "A Cost-effective Dependable Microcontroller Architecture with Instruction-level Rollback for Soft Error Recovery," in *IEEE*, 2007.

[12] D. R. Avresky, M. I. Marinov, "Machine Learning Techniques for Predicting Web Server Anomalies," in *IEEE*, 2011.

[13] Ratsameetip Wita, Yunyong Teng-Amnuay, "Vulnerability Profile for Linux," in *IEEE*, 2005.

[14] Bo Shuai, Haifeng Li, Mengjun Li, Quan Zhang, Chaojing Tang, "Automatic Classification for Vulnerability Based on," in *IEEE*, 2013.

[15] Vijaya MS, Jamuna KS, Karpagavalli S, "Password Strength Prediction using Supervised Machine Learning Techniques," in *IEEE*, 2009.

[16] Kai Liu, Yun Zhou, Qingyong Wang, Xianqiang Zhu,, "Vulnerability Severity Prediction With Deep Neural," in *IEEE*, 2019.

[17] Qian Yu, Yongjun Shen, "Research of Information Security Risk Prediction," in *IEEE*, 2016.

[18] I. V. Anikin, "Information Security Risk Assessment and," in *IEEE*, 2015.

[19] Venkatesh Jaganathan, Priyesh Cherurveettil, Premapriya Muthu Sivashanmugam, "Using a Prediction Model to Manage Cyber Security Threats," in *ResearchGate*, 2015.

[20] Sheung Yin Kevin Mo, Peter A. Beling, Kenneth G. Crowther, "Quantitative Assessment of Cyber Security Risk using Bayesian," in *IEEE*, 2009.

[21] R. Wita , Y. Teng-Amnuay, "Vulnerability profile for Linux," in *IEEE*, Taipei, 2005.

[23] L. Zeng, Y. Xiao , H. Chen, "Linux auditing: Overhead and adaptation," in *IEEE*, London, 2015.

[24] Y.Tian, J.Lawall and D.Lo , "Identifying Linux Bug Fixing Patches," in *IEEE*, Zurich, Switzerland, 2012.

[25] CompTIA Security+ Cert Guide: OS Hardening and Virtualization..