Florida International University [FIU Digital Commons](https://digitalcommons.fiu.edu/?utm_source=digitalcommons.fiu.edu%2Fetd%2F2791&utm_medium=PDF&utm_campaign=PDFCoverPages)

[FIU Electronic Theses and Dissertations](https://digitalcommons.fiu.edu/etd?utm_source=digitalcommons.fiu.edu%2Fetd%2F2791&utm_medium=PDF&utm_campaign=PDFCoverPages) [University Graduate School](https://digitalcommons.fiu.edu/ugs?utm_source=digitalcommons.fiu.edu%2Fetd%2F2791&utm_medium=PDF&utm_campaign=PDFCoverPages)

11-12-2004

A novel neural network based system for assessing risks associated with information technology security breaches

Monica DeZulueta *Florida International University*

DOI: 10.25148/etd.FI14062261 Follow this and additional works at: [https://digitalcommons.fiu.edu/etd](https://digitalcommons.fiu.edu/etd?utm_source=digitalcommons.fiu.edu%2Fetd%2F2791&utm_medium=PDF&utm_campaign=PDFCoverPages) Part of the [Electrical and Computer Engineering Commons](http://network.bepress.com/hgg/discipline/266?utm_source=digitalcommons.fiu.edu%2Fetd%2F2791&utm_medium=PDF&utm_campaign=PDFCoverPages)

Recommended Citation

DeZulueta, Monica, "A novel neural network based system for assessing risks associated with information technology security breaches" (2004). *FIU Electronic Theses and Dissertations*. 2791. [https://digitalcommons.fiu.edu/etd/2791](https://digitalcommons.fiu.edu/etd/2791?utm_source=digitalcommons.fiu.edu%2Fetd%2F2791&utm_medium=PDF&utm_campaign=PDFCoverPages)

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.

FLORIDA **INTERNATIONAL UNIVERSITY**

Miami, Florida

A **NOVEL NEURAL** NETWORK **BASED** SYSTEM FOR **ASSESSING** RISKS **ASSOCIATED** WITH INFORMATION **TECHNOLOGY** SECURITY BREACHES

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ELECTRICAL **ENGINEERING**

by

Monica DeZulueta

2004

To:Dean Vish Prasad College of Engineering

This dissertation, written by Monica DeZulueta, and entitled **A** Novel Neural Network Based System for Assessing Risks Associated with Information Technology Security Breaches, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved

Melvin Ayala

Tadeusz Babij

Malcolm Heimer

Robert Oikawa

Ana Pasztor

Naphtali Rishe

Malek Adjouadi, Major Professor

Date of Defense: November 12, 2004

The dissertation of Monica DeZulueta is approved.

Dean Vish Prasad College of Engineering

Dean Douglas Wartzok University Graduate School

Florida International University, 2004

(Copyright 2004 by Monica DeZulueta

All rights reserved.

DEDICATION

To my family, Julian, Elizabeth, Julian Alexander, Concepcion, Luis Alberto, Eugenia, Isabel, and Luis Alberto, Jr..

ACKNOWLEDGMENTS

I wish to acknowledge my sincere appreciation and thanks to my advisor Dr. Malek Adjouadi who has patiently guided me for so many years. My deepest thanks go to Dr. Melvin Ayala whose contribution with Neural Studio was invaluable. Many thanks go to Dr. Robert Oikawa for his assistance and knowledge of security principles. I would also like to thank my committee members Dr. Malcom Heimer, Dr. Naphtali Rishe, Dr. Tadeusz Babij and Dr. Ana Pasztor for their extra support and attention.

I am very thankful to have had the opportunity to have worked with some of the very best graduates from the Center for Advanced Technology and Education (CATE). I am also very thankful for the expertise and assistance that several of my Microsoft colleagues such as Greg Roberts and Kimberly Matlock offered.

Most importantly, I wish to extend my sincere gratitude and love to my family for their undying support for my work and studies. Their love and patience has been immense.

This work has been supported by Microsoft that generously supplied the access to data that was used to evaluate the analysis approaches and to the National Science Foundation and Office of Naval Research for their invaluable support provided under grants **EIA-9906600** and N-00014-99-1-0952.

v

ABSTRACT OF THE **DISSERTATION**

A **NOVEL NEURAL** NETWORK BASED SYSTEM FOR **ASSESSING** RISKS **ASSOCIATED** WITH INFORMATION **TECHNOLOGY** SECURITY BREACHES

by

Monica DeZulueta

Florida International University, 2004

Miami, Florida

Professor Malek Adjouadi, Major Professor

Security remains a top priority for organizations as their information systems continue to be plagued by security breaches. This dissertation developed a unique approach to assess the security risks associated with information systems based on dynamic neural network architecture. The risks that are considered encompass the production computing environment and the client machine environment. The risks are established as metrics that define how susceptible each of the computing environments is to security breaches.

The merit of the approach developed in this dissertation is based on the design and implementation of Artificial Neural Networks to assess the risks in the computing and client machine environments. The datasets that were utilized in the implementation and validation of the model were obtained from business organizations using a web survey tool hosted by Microsoft. This site was designed as a host site for anonymous surveys that were devised specifically as part of this dissertation. Microsoft customers can login to the website and submit their responses to the questionnaire.

vi

This work asserted that security in information systems is not dependent exclusively on technology but rather on the triumvirate people, process and technology. The questionnaire and consequently the developed neural network architecture accounted for all three key factors that impact information systems security.

As part of the study, a methodology on how to develop, train and validate such a predictive model was devised and successfully deployed. This methodology prescribed how to determine the optimal topology, activation function, and associated parameters for this security based scenario. The assessment of the effects of security breaches to the information systems has traditionally been post-mortem whereas this dissertation provided a predictive solution where organizations can determine how susceptible their environments are to security breaches in a proactive way.

TABLE OF **CONTENTS**

LIST OF **TABLES**

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

CHAPTER 1

Introduction

There have been great strides and innovations in **the field of** information technology **as in the case of the increases in** computing **power as well as the proliferation of** computing **power.** Over **the past decades, the access to computers has** changed **the way people** do business and **even interact with each other as can be seen in the extensive usage of** email **and** instant messaging. **The increased usage of** computers **in people's daily life has led to** an **increased dependency on** computers. **This** means **that if there are** problems encountered with the information technology, people's lives are disrupted. Security breaches have been **a great** source **of** these **disruptions to the** computing **world.**

Security is still very reactive, and **this study proposes** more **proactive** measures **when dealing with** security. Examples **of the reactive** nature **of** security measures are **the facts that most** companies **do not** implement security countermeasures until they have suffered an attack and **that** companies sometimes **fail to apply patches let alone test the patches in a** timely manner. Sometimes **a patch to a** security **breach is available six months prior to the release of a** virus **that exploits the** vulnerability, and **the** computer personnel **fail to install the patch the** removes **the** vulnerability. One **reason that** companies **do not deploy patches in a** timely manner **is that they feel** overwhelmed **with the** number **of patches that are** being **released.**

Before **detailing the proactive or even reactive** measures, **it is best to have** an understanding **of what the** various security **breaches are.** It **can often be seen in the** media **that the attacks such** as viruses, worms, **or** Trojan **horses have affected** organizations, **but there is a failure** to **really differentiate** between **the attacks. For** example, **a** virus **is** an **attack that requires a host**

to propagate or to even be active. Common examples of viruses are the **ILOVEYOU** virus which is a Visual Basic Script (VBS) that is sent as an e-mail attachment. **[1]** This virus required that an individual run the vbs script by means of clicking on the attachment. Since the attachment is a VBS script, the operating system would make the association that the VBS scripts are invoked and executed using the script engine. This association was done when a user opened the attachment; thereby, executing the script which in turn would iterate through the user's mailing list contacts and email itself to all the contacts in the mailing list. In addition to this, it would overwrite each graphic file that it found on the user's machine with the copy of the script.

As opposed to a virus which requires the use of a host, the worm is a stand alone attack, thereby, not requiring the use of a host. This malicious code self-propagates itself across the networks. A common example of a worm is the Blaster attack which exploited a buffer overflow vulnerability in Microsoft's implementation of the Distributed Common Object Model **(DCOM)** Remote Procedure Call (RPC) interface within the operating system. [2]

Another type of attack that is common is known as the Trojan horse. A Trojan horse refers to a program that is placed on the computer without being detected. The Trojan horse gives the attacker who placed the malicious code on the computer unrestricted access to data. Trojan horses have been known to access confidential data and transmit it back to the attacker. Back Orifice is an example of a Trojan horse which installs its server piece on a unsuspecting victim's object and is controlled **by** its client piece on an attacker's machine. [3] Unfortunately, most anti-virus software applications fail to detect the Trojan horse because the anti-virus engine is primarily focused on the scanning for viruses.

Three different classes of maiware based attacks have been mentioned, but a key factor is that these attacks could have been controlled and not have had an impact on the victimized corporations had certain security measures been present. Security relies on three key important factors: people, process, and technology. Although the attacks that were discussed earlier are directed toward software, technology is not the only way to mitigate against such attacks. In fact, 70% of attacks can be prevented by having the people and process in place, meaning that the use of technology accounts for only 30% of mitigating against a security attack. For instance, the ILOVEYOU virus relied on the host running with administrator privileges. If a corporation had the policy of least privilege in place meaning that users are not allowed to use administrative privileges to run applications, the virus would not have been able to iterate through the contacts in mailing list or overwrite the graphic files. Most of the users that were not affected by the ILOVEYOU virus were spared because the user did not click on the attachment rather than because they had the right policy in place.

Information security breaches have cost organizations in the upwards of 200 millions of dollars in the US alone. This cost encompasses physical attacks such as theft as well as software security attacks. The cost associated with property theft is roughly 7 million dollars while the software attacks has costs of over 65 million dollars. It is evident that these are staggering costs, yet there are a great deal of costs that do not even get reported. [4] The costs listed apply to US firms. Organizations in other countries are also plagued by security breaches. In fact, the Department of Trade and Industry reported that security breaches have cost businesses in the United Kingdom losses in the billions of pounds. [5]

Currently, there is no easy way to gauge how susceptible a corporation is to a security attack. In this study, a way of determining an organization's susceptibility to a virus attack is proposed, implemented, and detailed. The study includes determining which are key parameters in establishing and maintaining a secure system. People, process and technology are the three important aspects, and there are parameters associated with each of these areas. In order to discern what are the key parameters that have an affect on the outcome which is a secure system, a sensitivity analysis was performed. The study also undertook the effort of determining the relationship between the parameters in the areas of people, process, and technology to security. In essence, this study has three goals which are to determine what is the relationship between people, process, and technology to security, to perform a sensitivity analysis to discern what are the key inputs needed to have a secure system, and to build a predictive model that can determine the degree of risk to a virus attack an organization has.

Although part of the study is focused on assessing the degree of vulnerability that an organization may be susceptible to a virus, the results of the study can be utilized to assess a corporation's vulnerability to security breaches such as Trojan horses and worms. These security breaches can be passive or active. Passive attacks refer to an attack that does not change the state of the information system and is designed only to gain access to the system whereas active attacks do result in unauthorized changes to the state of the information system. This study is focused on the active attacks, namely, worms, but the assessment results can be utilized to determine how vulnerable an organization is to a passive attack.

The motivation to pursue this course of study was the fact that there is no way to properly assess the degree of risk that an organization may be facing when a security breach takes

place. It is most astounding that there is no predictive or proactive methodology considering the exorbitant amount of money that a security breach costs an organization. Sometimes, the security breaches not only affect the cost, but they also affect the branding and the perceived good will of the organization. There have a been documented attacks in which the attacker obtained confidential information such as the social security numbers of customers or credit card numbers and the attacker published the confidential information. The public loses faith in the company to safeguard the confidential information. This loss of faith usually results in lost business. An example of this is Wells Fargo case where customers' security numbers were released. [6]

It is important to note that there is no efficient or automated proactive methodology to assess the risk; in fact there is on going research that centers on monitoring network traffic patterns and determine if a passive attack is being attempted. This type of research is centered on reactive measures. An expert system monitors network traffic for load assessments and attempts to discem patterns to see if an attack is being performed. Modem day intrusion detection systems work in a similar fashion. An intrusion detection system **(IDS)** is defined as a defense mechanism that attempts to discem network activity. **By** its very definition, an **IDS** is a defense mechanism meaning it is entirely reactive. Another flawed aspect of an IDS in terms of its ability to be proactive is that it only addresses the technological aspects of security. It is imperative to emphasize that technology is only an element of security. In fact, it only constitutes about 30% of security. [8]

Because of the necessity of security especially in these days of post September 11 and the fact that security is often viewed in a reactive manner, this is why the study was proposed in

order to view security in a more proactive manner. Since a more proactive methodology was the goal of the project, a model that serves as a predictor was developed. It was determined that an artificial neural network (ANN) is a viable fit in this endeavor because ANNs can provide a predictor value given the proper training. The development and training of the model requires much data. In this study, the data addressed the three key facets of security: people, process, and technology. A critical factor of the innovation of this study was that it took into account other facets of security beyond the normal scope of previous studies which just focus on the technological aspect of security.

Using data that was collected detailing the current environments of Microsoft's customers globally, a neural network was developed, trained, and validated. The neural network was developed using a tool, Neural Studio, developed at the Center for Advanced Technology and Education at Florida International University. The data were obtained with the customers' consent in fact, a website was established so customers would voluntarily submit profiles about their environments. The data were voluntarily submitted by customers who were eager to see a study of this magnitude be successful. The customers filled a survey which would give the research an understanding of the people, process, and technology that was in place at the customer site. The survey consisted of over forty questions. Since the answers to some of the questions were qualitative in nature, answers were quantified. Prior to feeding the data into the Neural Studio, the answers were also normalized meaning they were adjusted to values ranging from 0 to 1.

In order to better communicate the contribution of this work, this dissertation has the following structure:

- * Chapter 1 **provides** an **introduction of the** current problem **of being reactive** with **respect to** security and **why there is a need to** make improvements **in** security. In addition, **it provides a brief** overview **of what** are **different** types **of security breaches.**
- **"** Chapter **2 introduces the three key factors of** security **which are people, process** and technology. **It details what contributions each factor provides in ensuring a secure** environment. **This chapter also establishes** the **security** terminology **that are used** throughout this **dissertation.**
- * **Chapter 3 focuses on the approach that was utilized to build the predictor** model. **It provides reasons** as **why** an artificial **neural** network (ANN) **was chosen to develop** the **predictor** model. **A** background **on** artificial neuron networks **is** furnished **in order to provide a** foundation **of** knowledge **for the subsequent chapters.**
- * **Chapter** 4 **details** the **how the** predictor model **was** designed, **trained, and validated.** In **this chapter, the** methodology **was validated using generated data. This chapter discusses the selection of the** neural network **topology, the activation** functions, and the **activation** function parameters.
- * **Chapter** 5 **builds upon the** knowledge **gained by** implementing **the** model **using generated data, and it applies such** insight **to building a** model **based on actual** customer furnished **data. This chapter** discusses **the collection of data** and **the** normalization procedures **that were applied to the data, in addition to detailing the selection of the topology, activation** functions, and **the associated** parameters.
- Chapter 6 provides a discussion of the results. The results associated with the model \bullet that was built using generated data are discussed as well as providing a detailed discussion and analysis of the results pertaining to the model that was implemented using actual collected data
- Chapter 7 consists of concluding remarks as well as listing future enhancements to \bullet this already significant body of work.

CHAPTER 2

Security Principles

Before there is an attempt to address the security problems that plague companies, it is imperative that the overall security principles be well understood. There is a tendency to attack security problems from a technology perspective only instead of taking a holistic stance. In this study, security was viewed from three key sides: people, process and technology. Before delving into the implementation details of this study, the key security principles will be discussed and some key security terminology will be defined.

2.1 Security Terminology

The goal of security professionals is to ensure that an information system meets the following requirements:

- **Availability**
- Integrity
- Confidentiality

These requirements are the most basic ones that need to be met, and ensuring that there are controls to certify that three requirements have been addressed is fundamental. Availability centers around the information systems being able to be used and accessed. For instance, if there is a security breach that causes the database server to be down, the information stored on that server is rendered unavailable. With availability, there are usually two associated topics of discussions, denial of services and loss due to a natural disaster. Denial of services

often refers to an intruder **whose actions** burden **a** computing system **so** much **that valid users** cannot **utilize the** system. **Controls that authenticate** and **authorize only valid users to** perform **duties can be** useful **to ensure** availability. **Loss of** computing services due **to a** natural disaster **is address by contingency** planning and **disaster** recovery methods.

Integrity **refers to protecting the state of the** information and ensuring **that the** information **has not been** tampered **with.** Systems **where** integrity **is of the** utmost importance **are** government **systems such as air traffic control** systems and military **weapons** systems. **Controls associated with the integrity of data** are:

- **e Least privilege**
- * Separation **of duties**
- **Rotation of duties**

Least privilege refers to granting **access** to only **what is needed to** perform **the job duties. Users are often** given administrative **access when their job** function **does** not deem **it** necessary. This **is often the case in the** type **of access that a user has to his/her own** computer. **It is** important **to note that users should** only modify **data** through well-formed transactions **that** are **well** documented **to ensure and** maintain **data** integrity. Separation of **duties is** targeted **at** ensuring **that** no **one individual is given control of a transaction** from **the beginning to the end. Control is** segmented, and **control of each** segment **is** assigned **to different individuals. This control is to** ensure **that** personal **gain** cannot **be achieved by having access to the entire** transaction. **The rotation of duties refers to periodically changing**

job assignments to avoid users from collaborating in order to gain access to the entire transaction.

Confidentiality refers to keeping information private and not disclosing that information. Privacy is critical to industries such as the healthcare and financial industries where patient information or credit information needs to be kept confidential and its access closely guarded. These industries are subject to governmental oversight. Controls focused on authentication and authorization can help address the confidentiality concerns so that only authenticated and authorized users can have access to data.

In order to meet the three requirements of availability, integrity and confidentially, technical controls are usually put in place to address the three key requirements. These controls often center on the following:

- **Authentication**
- **Authorization**
- Auditing
- Non-repudiation

Authentication refers to validating that the user is who he/she says he/she is whereas authorization is giving permission to the validated user to do what he/she needs to do. The first thing to do is validate the user before allowing the user to carry out any actions. Auditing is the ability to track the user's activities. This can be accessing a file or attempting to access a server or even attempting to enter a physical location such as the

11

server room. Non-repudiation is the ability to ensure that a transferred message was received or sent by parties claiming that they did perform said actions. For instance, a user may deny sending an email. If the email was signed, it can easily be verified that he indeed sent that email.

Other key terms that are often associated with security breaches also need to be reviewed. These key security terms are:

- **Threat**
- Safeguard
- Vulnerability
- Attack
- Patch

A threat represents a potential danger to business assets. All threats should be calculated in terms of business risk. For instance, a threat can be the unauthorized access to a customer's credit information. A threat can manifest itself in the loss of privacy, availability, and/or data integrity. A safeguard is a measure that counters the threat. This measure can be technology, people, or process. An example of a safeguard may be the process of requiring employees to use a badge to gain entrance to a facility so the threat of non-badge personnel accessing the facility is lessened. A vulnerability is a weakness in a safeguard or an absence of a safeguard that can be exploited. Leaving ports open in a firewall is classified as a vulnerability. An attack refers to a threat that was carried out **by** exploiting the

vulnerability. The various viruses that have penetrated information systems and corrupted files such as the ILOVEYOU virus are considered examples of attacks. Patches are designed to eliminate the vulnerabilities that may exist. Usually upon the discovery of a vulnerability, platform vendors will release a software update which is referred to as a patch to address the vulnerability before an attack can be made to exploit it.

2.2 People and Process

Although people and process are two distinct factors in security, they are closely tied together in that the processes are created by people and implemented for people to follow. This section will detail how people and process can impact security.

People and their behavior play a pivotal and critical role in maintaining a strong security. Social engineering demonstrates how attacks on computer information systems may be successfully carried out by exploiting weaknesses in humans. An example of this is where unassuming individuals have pretended to be government officials and have been given access to secured areas because the security guards thought that the visiting personnel were federal officers. [9] One way to guard against people being the weakest link is having processes and policies encompassing training and awareness in place to ensure that people become the best defense instead of the weakest link.

A key person that can assist in ensuring that processes are in place is the security officer. Companies should invest in a qualified chief security officer that can provide the needed guidance and set the overall security policy for the company. The security officer would create the overall security policy and ensure that the guidelines, standards, baselines and procedures adhere to the policy.

The security policy refers to the overall corporate policy. In essence, it establishes the company's attitude toward security. Standards are needed to detail the mandatory actions or rules that support the corporate policy that every employee must adhere to. For instance, a standard may dictate that all computers in a corporation must run a specific anti-virus software. In addition to standards, companies also need baselines which provide mandatory descriptions how to implement security packages in order to ensure a consistent deployment. The guidelines refers to recommendations on how to deal with security issues whereas the procedures are the step by step instructions on how to handle security issues. If an organization wishes to have an environment that exhibits security awareness, it needs to establish a corporate security policy, standards, baselines, guidelines, and procedures. It is not enough to just have these documents created. **[10]** Employees need to live by them and use them. Ironically, some organizations have these documents, but they do not enforce them which renders them useless. For instance, there may be procedures when an employee is terminated. The procedure may dictate that all access, whether it be physical or remote, to network resources be terminated as well as having the employee escorted **by** security guards. If this procedure is not followed, the terminated employee may implant a virus or erase data on an application server.

2.3 Physical Security

An often overlooked factor of security is physical security. Theft of computer property has cost US companies an excess of 7 million dollars [4]. Physical security should be considered a fundamental part of security. When securing physical equipment, servers should be the first priority. Most if not all, software technical controls can be defeated easily if an attacker has physical access to a machine. For instance, a malicious user could boot a machine using another operating system running off a CD, grab the password file and crack it. This demonstrates how physical access trumps software technical controls. There should be guidelines and procedures in place that detail the secured physical access to the server rooms. Access to information needs to be restricted whether it is technological restrictions and/or physical restriction.

As part of physical security, the use of escorting terminated employees should be considered as well as escorting guests. These circumstances needed to be addressed in the guidelines and standards.

Theft is not the only consideration of physical security; environmental concerns should also be considered as part of physical security. For instance, redundant network connections, fire suppression, and power connections should also be part of the physical security plan because of their impact on availability. [10]

2.4 Training

Before the policies, guidelines, and procedures can be enforced, they need to be shared and communicated to the employees. Training is how an organization can effectively communicate the security policies, the guidelines, and the procedures. Security awareness training is needed for all employees and should be done periodically so that security becomes foremost in people's minds. Those individuals in the information technology (IT) and security departments should have specialized training focused on technologies that safeguard information. Since technology changes at a staggering pace, those in IT and security need access to current training, periodically.

Training is only part of the story. In order to achieve the goal of a secure organization, security needs to be foremost in the employees' minds so that it becomes just second nature. This means that even the development of software should also incorporate security best practices as well as the physical access should incorporate best practices.

In addition to offering security awareness training and specialized security training, it is important to ensure that personnel are well versed in security practices. Key security personnel should be experienced professionals with experience in establishing security policies, guidelines, standards, baselines and procedures in addition to having experience in enforcing them as well.

2.5 Privileges

Other processes that need to be in place revolve around least privilege. This refers to running resources using non-administrative privileges. There is a tendency to run applications with administrative privileges even if the application does not require such privileges. For instance, many employees run as local administrator on their Windows machines when all they wish to do is use a word processor. Malware may run at a privilege

16

level equal to that of the logged-on user, and **thus gain the** same **ability to control** and administer **the** computer **as the local** administrator. By **enforcing the use of least privilege, the** company may **reduce the** damage **that** a **security breach can do.**

Another **facet of privileges is ensuring that no one** employee **has control over an entire** transaction **of work. If one** employee **is given that** much **power, there** may be an opportunity to **violate or** bypass any **controls that exist in the** transaction. This type **of** opportunity **can** be **avoided** by **applying separation of duties where no one individual has control of** an **entire process. For instance,** an enforcement **of separation of duties would require** two **separate parties to approve** expenditures **and release the** monies **instead of having just one person approve the money** and **having physical access** to **the** money. **It would** be **too easy for the person to** make fake **approvals** and **keep the cash.**

Rotation **of duties refers to the** recommendation **that duties should be routinely rotated so collaboration can be avoided.** In an environment **where there is separation of duties, collaboration is needed to gain control over** an **entire** transaction. **Too** much **time on the** same assignment may **lead to** complacency **and lend itself to** employees **collaborating, thereby** making **the separation of duties ineffectual.**

2.6 Patch Management Process

Companies **need to have a patch** management **process. This process should detail the** procedures **for testing patches** targeted **for both client and** server computers and **the** procedures **for deploying patches to both clients** and servers. Along **with** the procedures, **there** should **be deadlines as to when the patches will** be fully **deployed across the enterprise.**

There have been instances **where** software **vendors have released patches six** months **prior to the creation of a** virus **that can exploit the** vulnerability, and organizations **have not deployed the patch in a** timely manner. Deploying **is only one aspect of the patch** management **process; testing of the patch needs to** occur **before** any **roll out is done. For** instance, **a patch** may somehow **affect a third-party or in-house developed application.** Thorough **testing of a patch needs to occur** before any deployment **is** performed **to ensure that the patch does not affect the** operating **applications.** Figure **2-1 depicts** a **flow of the patch** management **process that can be utilized to roll out patches to** production servers. **A** similar **process can be done with client** machines. **[11] There** are software **packages such as** Microsoft's System Management Software **that can** perform **the** deployment **to clients** machines **without requiring physical access to a client machine. Such tools also provide auditing** features **that detail if the** deployment **was successful** and **when it was** performed.

2.7 Password Policy

Passwords are **useful** and common **tools for validating users.** Because **they are so** common, **it is** necessary **to** ensure **that they are not** shared. **There needs to be a policy that enforces the uses of** password. Using **passwords is not** enough; **there needs to be a** requirement **that the password** be a **strong password. Strong passwords** impose minimum limits **on size** and **the characters that can** be **used and how often a portion of the password can be reused.** There **is a tendency for users to use** an easy-to-remember **password such as their** name and **then change a portion of it periodically in a** predictable fashion **such as adding "1"** and **"2" to** remain within the company policy for changing passwords. to devise a password and just simply change a **portion of it periodically.** This means **that the** passwords **could be obtained**

with greater **ease,** thereby making the company more at risk to a security breach. Since passwords are susceptible to dictionary attacks, there needs to an enforceable policy that requires strong passwords. Dictionary attacks refers to finding password using a words that can be found in a pre-defined list of words such as a dictionary.

2.8 Software Development Process

When an organization opts for developing application software in-house, security needs to be included in the design phase and not considered at the end of the development process. [12]

One of the reasons that security is often overlooked in the design phase is that security is often considered inversely proportional to usability. Many users and even developers tend to find that a very secure system is not very user friendly. A user friendly system having no passwords may not be available because of security attacks. People need to recognize the value of security and include it in their design. The goal should be secured by design. In fact, about 50% of the security bugs are often associated to design errors. [13] Processes that reduce security flaws need to be established.

2.9 Threat Modeling

A useful methodology in assessing threats and vulnerabilities is known as threat modeling. It is a detailed process that should be incorporated early in the product life and should be

Figure 2.1: Patch Management Process

maintained throughout the product's life cycle. The process of threat modeling can be defined **by** the following set of steps:

- Create a team to perform the threat modeling
- Decompose the healthcare application into its key components
- Ascertain the potential threats to the system and its components
- Rank the threats
- Decide on effective ways to address the threats
- Evaluate which techniques may be applied to address the risks
- Establish which technologies are suitable to implement the chosen techniques

Threat modeling by its indefinite and encompassing nature is not a simple task. The team needs to include a security analyst as well as having representation from the disciplines that are involved in the solution life as well as those disciplines associated with the solution functions. Paradoxically, hackers are well suited for the role of security analyst because there is a tendency in hackers to search for the vulnerabilities. The project life cycle related disciplines may include networking, developers, operations, training, and testing; whereas the functional disciplines may include users such as clinicians and insurance representatives in the healthcare landscape.

The next step is to identify the subsystems (components) that constitute the solution. For software-based solutions, common decomposition tools can be **UML** (unified modeling

language) diagrams or the DFD (data flow diagrams). For the hardware-based systems, block diagrams can be used. The key part of this step is to identify the boundaries on the subsystems and not the functionalities. This means that software lines of code or schematics are not the object of the analysis at this junction, but it is rather the study of boundaries of the subsystems that are at the core of the analysis. As part of the decomposition it is recommended to study the interfaces and interactions among the subsystems. The interfaces and interactions may include events, triggers, requests, and responses the subsystems will generate or respond to. During this phase, the decomposition of the system should not typically go beyond four levels [13].

2.10 Classifying Threats

The threats can be classified using the STRIDE categorization [13] developed by Microsoft. Although it is usually applied in a software based scenario, it can be applied to classify threats in general. STRIDE is an acronym which stands for six classifications:

- Spoofing Identity
- Tampering with Data
- Repudiation
- Information Disclosure
- Denial of Service
- **Elevation of Privilege**

The spoofing of identity refers to an attacker impersonating a known user or server. The tampering of data means when data are maliciously altered. Repudiation refers to a user denying that he/she performed an act, yet there is evidence that the act was committed by that user. A denial of service refers to the valid user being unable to access a system because a maliciously attack has consumed the system resources. The elevation of privilege refers to an attacker obtaining privileged access giving him/her access to the system. The different classes of threats can be interrelated. For instance, spoofing of identity may result in the elevation of privilege. Another important fact to retain is that an asset may be the target of more than one class of threat.

2.11 Identifying Threats **and** Vulnerabilities

Once the system has been decomposed into components, it is necessary to determine what the system's vulnerabilities are and which threats can exploit those vulnerabilities. Threat trees can be used in identifying those vulnerabilities. Threat trees refer to applying the concept of fault trees to security. The use of fault trees is a common way of identifying failure modes in hardware. Threat trees show how an attack can manifest itself and how a system can be compromised by showing the decision-making process that an attacker may apply to attack the key components. As the components are attacked, the entire system is getting compromised. The listing of threats should contain several details in addition to just the threat title and the type threat. The list of threats should include:

- Threat $-$ a description of the threat
- \bullet Threat target key component that is susceptible to attack
- Threat type $-$ STRIDE classification
- Risk calculated consistently according to desired method
- Attack tree tree showing how an attack will manifest itself

2.12 Ranking Threats

Risk assessments are useful for assessing the risks and vulnerabilities of protected information. After the threats have been identified, the ranking of the threats is done. Such ranking is based on the risk associated with each threat. There are various methods of calculating risk. One method is called DREAD. **[13]** Calculating risk and its pre-requisites are an analysis process essential in the identification of threats and vulnerabilities that the healthcare system faces. In the complexity spectrum, DREAD is considered to be somewhat moderate in complexity in that it takes into account more factors than an overly simplified methodology which will account for a minimal of two factors such as criticality and likelihood. A very complex method will instead involve complex calculations such as regression analysis. DREAD is an acronym that stands for:

- Damage potential
- Reproducibility
- **Exploitability**
- Affected users
- Discoverability

Damage potential refers to the criticality of the vulnerability. Reproducibility signifies if the threat can exploit its target in a predictable manner with respect to time. Exploitability metric measures how much effort or how easily the asset can be exploited. The value assigned to affected users would reflect how many users would be affected by the threat. The discoverability factor refers to how easily the vulnerability will be discovered. After each metric has been assigned a value between 1 and 10 with 10 having the highest impact, the risk associated with each threat is determined by summing each metric's assigned value and dividing by the number of metrics which in this case is 5.

2.13 Mitigating Threats

Table 2-1 showcases a sampling of techniques that can be used to mitigate the risks associated with the different threat types.

To determine which mitigation technology should be applied, it becomes necessary when designing the system to see what are the implementations associated with the technique. Examples of software authentication technologies are forms-based authentication, Windows authentication, and Kerberos authentication. Authorization technologies that are found in software implementations are access control lists, privileges, IP restrictions, and serverspecific permissions. These technologies have been used to address denial of service threats, information disclosure threats, spoofing identity threats, and tampering with data

threats. **The use of logging in both operating** systems and **application** server systems **are** useful technologies associated **with the** mitigation technique **known as audit trails** which addresses **the** threat **of non-repudiation.**

Threat Type			
	Mitigation Technique		
Spoofing identity	Appropriate authentication		
	Protect secret data		
	Don't store secrets		
Tampering with data	Appropriate authorization		
	Hashes		
	Message authentication codes		
	Digital signatures		
	Tamper-resistant protocols		
Repudiation	Digital signatures		
	Timestamps		
	Audit trails		
Information disclosure	Authorization		
	Privacy-enhanced protocols		
	Encryption		
	Protect secrets		
	Don't store secrets		
Denial of service	Appropriate authentication		
	Appropriate authorization		
	Filtering		
	Throttling		
	Quality of service		
Elevation of Privilege	Run with least with privilege		

Table 2.1: Mitigation Techniques Associated with Threat **Types**

2.14 Technology

The technologies **that** are **detailed in this section can** be **applied to the** software development **process as well as technologies that can** be **applied in general for** securing information technology systems.

To ensure that a system exhibits the characteristic of a secured system, certain technologies can be implemented where it is deemed appropriate. This section will provide an overview of the various technologies and what security features they implement.

One technology that companies often treat as the proverbial *silver bullet* that will protect them from any attacks is the anti-virus software. Although anti-virus software is important to implement across the enterprise, it is only one step and **by** no means the only step. Installing anti-virus software is only a start; ensuring that the virus signatures are always up to date is the critical factor with regards to using anti-virus software.

Firewalls are another technology that many companies tend to use. A firewall is a hardware or software defense mechanism that monitors traffic going through and determines if it adheres to a security policy. For instance, a firewall can block traffic on designated ports. Firewalls are usually placed to separate the Internet from the corporate network. These firewalls usually should block **all** ports with the exception of the ports that handle standard internet protocols such as HTTP and HTTPs. Too further protect the corporate information assets, companies tend to separate their web servers from their *backend* systems. The backend systems are usually the databases or mainframes that house business data. The firewall that separates the backend systems should block all unnecessary ports. The zone that lies between two firewalls is called the perimeter network. The perimeter network is the area where web servers are placed. There is a tendency to think that by having firewalls the company is free of problems. Again, firewalls are only one very small way of protecting the corporate resources. There are also different types of firewalls some of which are better than others at inspecting the traffic. Another important fact is that upgrades to firewall

software should always be current. Along with irewalls there is a technology that monitors the traffic and makes decisions based on the traffic patterns. These systems are referred to Intrusion Detection Systems (IDS). IDS systems usually are based on a rules-type engine such as an expert system. They monitor the traffic to see if an Internet Protocol (IP)-based attack such as SYN Flood is being done. The SYN Flood attack is when a machine continually initiates TCP/IP connections forcing the receiving machine to continually allocate memory and allocate resources to handle the connection requests. This ties up the receiving machine so that it cannot handle valid connection requests, thereby resulting in a denial of service attack.

2.15 Authentication **and Authorization**

Some key requirements for a secure system are authentication and authorization. Mechanisms such as directory services can be utilized for authentication and authorization. Directory services can determine if the user is a valid user. Based on the access control lists (ACL), directory services can determine if the validated user should be given access to system resources. Access Control Lists detail what capabilities a principal has on a resource. These capabilities may be read, write, and execute. When an authenticated principal wishes to access a resource, the system checks in the ACL to see what can the principal do on the resource. The system will grant access according to the ACL. Examples of system resources can be a file share repository and a printer. The use of both directory services and ACLs can address authentication and authorization concerns. Directory services are not always suitable for authentication as in the case of an internet web application, but it may be suitable for an intranet application and to some extent an extranet

web **application.** An **authentication** mechanism which would verify **users against** account information **stored in a database** server **is a** more **suitable** mechanism **for** an **interet web application.** [14]

In **essence, authentication requires that proof also called credentials be supplied to certify that the user is a valid user. Credentials can be:**

- **"** Something **that is** known **as in the** case **of a password or Personal Identification** Number (PIN)
- * Something **that is** owned **such as a** smart-card
- * Something **that is** unique **about the user such as a bio-metric possibly a fingerprint**

With respect **to authentication, the use of** two-factor **authentication is considered far superior** than simply **using one factor authentication such as a password. Two-factor authentication utilizes** two types **of credentials. For** example, two-factor **authentication can be** implemented **as something that is known such as** a password and something **that is** owned **such** as smart-card.

2.16 Non-repudiation

Analyzing **the** characteristics **of a** secure system, **the ability to have non-repudiation is key in that it** provides **a** guarantee **that the sender did** in **fact send the** message **as well as the recipient received the** message. Non-repudiation **is usually** accomplished **by** means **of** digital signatures. Digital signatures rely on a private and public key infrastructure. A

message having passed through a cryptographic function is said to be hashed. The hash is then encrypted with the sender's private key and sent. The recipient will decrypt the message with the sender's public key. This message is said to have been digitally signed; therefore, the recipient can verify that the sender was indeed a trusted party.

2.17 Auditing

Auditing refers to tracking system activities and having a trail of the recorded activities. These activities can be successful and unsuccessful attempts to access a resource. Collecting the system activities is critical in the review and analysis of how secure a system is. The auditing activity is entirely reactive.

There are different types of trail or logs that are associated with auditing. For instance, platform vendors offer an event log that lists the tracked system related events. Some of the events that can be monitored are:

- Accessing system resources such as printers or file shares
- Analyzing network traffic and connections
- Attempts to logon
- Modifying records

Other types of logs are the web server logs and the database server logs. All the different types of logs need to be reviewed regularly. Another consideration is to have an alert mechanism that would notify operational personnel if certain critical events occur. These events can be high CPU utilization or extensive memory utilization for an extended period of time.

2.18 Internet **Protocol Security (IPSEC)**

Since the Transport Control Protocol/Intemet Protocol (TCP/IP) was not designed to be a secure protocol but rather a reliable protocol, packets associated with the TCP/IP stack can be easily faked. TCP/IP offers no authentication, authorization, privacy, and data integrity. IPSEC attempts to address the security problems associated with the TCP/IP stack by relying on a key exchange to offer the much needed authentication, authorization, privacy, and integrity. All traffic that is exchanged between IPSEC-secured servers is encrypted and integrity-checked.

2.19 Secure Sockets Layer **(SSL) or** Transport Layer Security (TLS)

SSL was invented by vendors, Concencus and Netscape, whereas TLS is a product of the Internet Engineering Task Force (IETF). SSL **uses** Message Authentication Code (MAC) algorithms to ensure data integrity. Data packets are encrypted as they travel between a client and server and vice versa. The sender and receiver obtain message integrity assurances by using a secret key and the message to obtain a MAC or sometimes called a hash. The sender sends the message and the hash to the receiver. The receiver will calculate a hash and compare with the hash that was received. If the two hashes are the same, then both the receiver and sender are assured that the message was not tampered with during the transit.

The technologies, SSL/TLS, are often used in web sites to provide encryption for confidential information such as credit information or login passwords on web sites. It also used for encrypted logins.

Both IPSEC and SSL/TLS can be used to address many privacy or data integrity concerns. Both technologies use cryptography to encrypt network traffic. Encrypted traffic increases the difficulty to sniff the traffic, hence keeping the data confidential and private. Data integrity is ensured by the use of MAC algorithms which certify that data are not tampered with.

This section has detailed the different security principles that encompass people, process, and technology. This has been an overview of the various technologies that can be utilized.

CHAPTER 3

Approach Proposed for Risk Assessment

After reviewing that the affects that security breaches have on organization, a course of study that would be yield a predictor for these security breaches was undertaken. The goal was to build a model that given a profile of an organization's information system environment would generate the degree of risk of being impacted by a virus.

This section will discuss the proposed approach that was developed for assessing risk in an organization. Two risks will be analyzed: one for the production server environment and the other for the environment composed of client machines. This chapter will detail the methodology that was designed for building the predictive model. This chapter begins with a discussion of the universal theorems for developing approximators and progresses to the selection of the type of model based on the universal theorems. The methodology developed for building the model is defined. The dataset collection and its preparation are also discussed in this chapter.

3.1 Universal **Approximators**

There are various methodologies that could have been utilized to build the model. The artificial neural network was selected as the methodology because neural networks are well suited for building universal approximators. Other methodologies that were considered were support vectors. Support vectors are better suited for classification scenarios. In this study, the goal was to build a predictor or approximator given certain environmental conditions.

Multilayer networks **can** be **utilized in** approximating multi-dimensional functions. This is supported **by several key** theorems **that are** universally **accepted.** [15-21] Some **of the** theorems **that support that selection of** an artificial neural network **for the** model:

- **"** Kohnogorov's Theorem
- **" Sprecher** Theorem
- * **Hecht-Nielsen** Theorem

Kolmogorov's Theorem **states that** any **continuous function defined on a closed** multidimensional **cube can be rewritten** as **a** summation **of applications of continuous** function **on** one variable. This means that for all functions $f(x_1, \ldots, x_n)$ with $x_i \in [0,1]$ there are continuous functions ψ_i and ϕ_{ij} on one variable such that:

$$
f(x_1, ..., x_n) = \sum_{j=1}^{2n+1} \psi_i \left(\sum_{i=1}^n \varphi_{ij} (x_i) \right)
$$

This **theorem only states that the** function **exists** meaning **it is only** an **existence** theorem. **It** provides no assistance in finding the functions. The functions ψ_i and ϕ_{ij} are dependent on **the mapping** function, f This theorem **refers to the** mapping **for multi-dimensional** functions **with n inputs** and 1 **output.**

The Hecht-Nielsen Theorem **is based on Kolmogorov's** Theorem, **and it shows that** any **continuous** function **can be** approximated **by a feed** forward network **with n inputs, 2n+1 hidden neurons,** and m **output nodes. Unlike Kolmogorov's** Theorem and **Sprecher's** Theorem, **both of** which **refer to having only one** output, **the** Hecht-Nielsen Theorem **states** that any continuous multi-dimensional function with n inputs and m outputs can be rewritten.

All these theorems state that an **appropriate** network **exists for such a** multi-dimensional problem, **but** they **offer** no guidance **on how to find the appropriate** network. **In** particular, **they** give **no indication of a** method **to find the appropriate weights** and **biases. A key contribution of this study is the** development **of** an **appropriate** network **to** determine **the risks in the security space.**

3.2 Conception of the Methodology

The methodology **used in building** model for **predicting the risk in a** company's **production server** environment and **client** machine environment **consists** of:

- **" Selecting the appropriate** type **of** model **for a predictor**
- * Determining **the** artificial neural network **topology**
- **" Selecting the** training **methods best suited for the predictive** model
- **"** Collecting **datasets**
- **" Preprocessing the datasets**
- Training the model
- Validating the model

Analyzing **suitable strategies for creating a predictive** model **resulted in the selection of** an artificial neural network **(ANN) since the** ANNs **are very suitable for creating** predictors **given** historical **data.** This **selection is** further **validated by the** universal approximators theorems **presented earlier.**

Before discussing **the proposed approach, a review of** artificial neural networks **will** be **performed.** This **review will define what** an artificial neural network **is , how it works,** and **where it is being used. In addition, a brief history** and **the foundation of** artificial **neural** networks will be provided.

3.3 Artificial Neural Network Fundamentals

An artificial neural network **is a** system **that** models **the** structure **of the brain. It uses general** mathematical functions **to** model the functions of neurons in the brain. **A** multitude **of layers consisting of simple processing** elements **called neurons** are **used to** simulate **the brain. Each processing** element **or** neuron **is** connected **to certain** neighboring elements. **With each** connection **there** are **coefficients that represent the strength of the** connection. The coefficient of strength is called the weight, and it usually is multiplied **by** the input signal. **A learning phase is needed in order to adjust these strengths to cause the overall** network **to** output **appropriate results.** This learning **phase is often called training the** network.

Artificial neural networks **have been used** in many **industries.** In **the** financial services, ANNs are **used for** predict **stock** market **fluctuations. The** military **has** made **extensive use**

of ANNs because **of its ability to facilitate** pattern **recognition.** The **power utility** companies often **leverage** ANNs **for fault detection.**

3.4 Artificial Neural Network Overview

Similarly **to biological** neuron network, **the** artificial neural network **has** as **its basic** element, **the neuron.** This **artificial neuron has four basic** elements, which **accepts inputs, applies** weights, and **applies** an **operation,** and outputs **the results. The** figure **below shows** an **artificial** neuron and **its** four **basic** functions. **[22] The inputs to the** network are **referred to** as $x(n)$ and are multiplied by weights, w_n . There may also be a bias which lets the output be **biased without** accounting **for the inputs. In this** simplified **case, the products of inputs** and weights are **just** summed. **The** summation **is processed** through **a** transfer function **also called** an **activation** function and **outputted.**

Figure **3.1.** Simplified Artificial Neuron

Figure 3.2 McCulloch-Pitts Model

The Hebb net is considered to be the first supervised training neural network. In **this** type **of** network, **the** weights change **according to the** training and **are not fixed like its predecessor, the** McCulloch-Pitts model. **[22]** In **the Hebb** model, **the** weights **are initialized to zero** and **are updated** as **the** network **is trained. The Hebb** model **uses the following rule:**

$$
w_i(new) = w_i(old) + x_i t
$$

where wi is the weight **associated with the input, x;,** and **t is the target** output. **The Hebb** model **accepts either** binary **or bipolar inputs. The** problem **with using** binary **is the** model will **not** learn **if the** target **is** 0.

The Perceptron builds upon the Hebb model. The **Perceptron uses** an **iterative** learning procedure **to converge to the correct** weights. The **correct weights are defined as the ones that** will yield **the correct** output **for all input training** patterns. **There is the** assumption **that such** weights **do exist. The output of the Perceptron is defined as:**

$$
y = f(y_{in})
$$

where

$$
y_{in} = b + \sum x_i w_i
$$

The activation function is

$$
f(y_{in}) = 1 \t\t\t\tif y_{in} > \theta
$$

$$
= 0 \t\t\t\tif -\theta \le y_{in} \le \theta
$$

$$
= -1 \t\t\t\t\tif y_{in} < .\theta
$$

It should be noted that there two thresholds, θ , and, θ . These two thresholds define two decision boundaries and the three output spaces. In the Perceptron, the weights and the bias are adjusted when an error occurs. An error refers to the output not matching the target function. The weights are adjusted **by** using the following:

$$
w_i(new) = w_i(old) + \alpha x_i t
$$

where α denotes the learning rate and lies in the range between 0 and 1.

With the Perceptron as more training patterns yield the correct response, the learning is lessen since the weights are adjusted when errors occur.

The Adaptive Linear Neuron (Adaline) has much similarity to the Perceptron. In fact, the net input, yin, is calculated in the same way. The activation function for Adaline is different and is defined as:

$$
f(y_{in}) = 1 \quad \text{if} \quad y_{in} \ge 0
$$

$$
=-1 \text{ if } y_{in} < 0
$$

The weights are calculated using:

$$
w_i(new) = w_i(old) + \alpha(t - y_{in})x_i
$$

A major difference between the Perceptron and the Adaline is in the way the weight is calculated. In Adaline, the weight is proportional to the delta or difference between the target and the actual output whereas in the Perceptron the weight is the target output.

The Madaline model is simply an extension of the Adaline. The Madaline applies the Adaline algorithm to multiple layers. **[23]**

3.5 Design

The design of a neural network can be complex and is often iterative until a satisfactory design is developed. The following steps are considered typical of any artificial neural network:

- Determining how many layers are necessary
- Arranging the neurons in the various layers
- Selecting the type of connections between the layers as well as within a layer
- * Deciding the strength of the connection **by** permitting the network to learn what are the suitable weights for the connection
- * Determining **how the** neurons **receive their inputs** and yield **their output**
- Selecting what is the most suitable training method

3.6 Layers

Biological neural networks **are** constructed **in a** three dimensional **way capable of a multitude of** connections. **Such** complexity **is not** true **of the** artificial neural network. **ANN** are simple **clustering of** artificial neurons. The **clustering is accomplished by creating** inter-connected and **intra-connected layers. The** connections between **layers** and within **layers vary.** Artificial neural networks **have at least three layers:** an **input layer, one or** more **hidden layers, and** an **output layer. The input layer accepts real world inputs, and the output layer delivers the** network's output **to the real world. The** number **of hidden layers vary.** Unfortunately, determining **the** number **of hidden layers is usually** accomplished **via trial** and **error.** If **the** number **of hidden** neurons **is too** large, **the neural net will have** difficultly **generalizing which is also** known **as** an **over fit. The** training **set will then be memorized** and **the** network **will not work with new data sets. The** figure **below shows the** typical topology of an artificial **neural network. [22]**

Figure 3.3. Layering Structure of an Artificial Neuron

3.7 Communications **and** Connections

In order for the communications to exist between the layers, there needs to be connections between the neurons. The way neurons communicate is the output of one neuron feeds the input of another. This usually reflects uni-directional communications. If bi-directional communications is desired, the output of one neuron is connected to the input of another and a separate connection going the opposite direction also exists between the neurons.

The connections between neurons can be either inhibitory or excitatory. If the connection is inhibitory, the output of the neuron causes the activity of the receiving neuron to reduce whereas in an excitatory connection the output of the neuron causes the activity to increase. In other words, the excitatory connection causes the summing mechanism to add unlike the inhibitory connection which causes the mechanism to subtract.

The neurons are usually organized in at least three layers: an input layer, one or more hidden layers, and one output layer. The connections between neurons can be across the layers

which is referred to as inter-layer connections and within a layer which is known as intralayer connectivity. There are various types of inter-layers and intra-layers connections.

The connections between layers can either be **fully** connected or partially connected. A fully-connected inter-layer connection refers to having every neuron in one layer connected to every neuron in the other layer. In a partially connected inter-layer connection, not all neurons in a layer have to be connected to all neurons in another layer. Another characteristic of inter-layer connections is whether the output of the second layer is fed back to the input of the first layer. In a feed forward network, the neurons of the first layer send their output to the neurons in the subsequent layer, but the first layer neurons do not receive any input back from the neurons on the subsequent layer. Unlike the feed forward network, in a bi-directional network, the subsequent layer does provide input back to the first layer. Feed forward networks as well as bi-directional connected networks can be either fully connected or partially connected.

Communications between layers can reflect a hierarchical structure in which the neurons in one layer can only communicate with neurons in the subsequent layer. Other times, communications may require bi-direction where messages are sent continuously between neurons until a certain condition is achieved. This format of bi-directional communications is called resonance.

More complex neural network have connections not only between layers but also within layers. The connections of neurons within a layer is referred to as intra-layer. The two types of intra-layer connections are recurrent and on-center/off surround. In the recurrent intra-layer, neurons receive inputs from another layer and will communicate their outputs

with other neurons in **the** same **layer** until some **condition is** achieved and signals **others to** communicate **their** outputs **of the layer to another layer.** In **the on-center/off** surround model, **a** neuron **has excitatory** connections **with itself** and **its** neighbors **and** inhibitory connections **with the other** non-neighboring neurons. **After** an **exchange** among **the** group **of neurons, the** output **is released.**

3.8 Learning

The learning **done in** the artificial neuron network **is** analogous **to the** learning **done in a** human **brain in that** learning **is** accomplished through **experience. Such** learning **is** accomplished **via the** adjustment **of the** weights **on the** connections. The weights **reflect the** strength **of the** connection between **the** neurons. **A** neural network's learning **ability is influenced** by **its** architecture and the training methodology.

The three types **of** training **methodologies that are associated with** neural networks are **as follows:**

- Unsupervised
- **"** Reinforcement
- **"** Back **propagation**

In **the** unsupervised learning, **the neurons in the hidden layer or layers do not need** any **outside** assistance **in** organizing **themselves.** This means **that no** sample **outputs are**

provided for a given set of inputs. The neurons do not have a sample to measure against. In the unsupervised training scenario, the neurons simply learn by doing.

The reinforcement learning method is a type of supervised learning. This learning method relies on reinforcement from the outside meaning that a teacher is involved. This teacher can either be a training set of data or an observer that grades the network. During the reinforcement training, the connections are continually reshuffled as the network is informed on its progress on solving the problem.

The unsupervised and reinforcement training are plagued by the inefficiency in obtaining the proper connection weights. The back propagation method attempts to address the inefficiencies of the other methodologies. In the back propagation method, the errors are filtered back, and the information about the errors is used to adjust the connections of the neurons between the layers. Back propagation is also considered a form of supervised learning. It has significant performance improvements over the other methods.

Aside from the three different, learning methodologies, learning is also characterized in one of two modes: on-line or offline. If a neural network is in the off-line mode, it is not in operating mode nor it is being used for decision-making. During the off-line mode, the neural network is adjusting the weights as it learns. Once the off-line neural network changes to operational, the weights are fixed. Most neural networks learn in off-line mode. As opposed to off-line networks which learn while the network is not operational, on-line neural networks are those that continually learn while it being used as a decision tool.

There are various learning algorithms. These algorithms are mathematically-based algorithms that are used to adjust the weights associated with the connections. The oldest learning algorithm is Hebb's Rule. Most of the subsequent learning algorithms are simply variations of Hebb's Rule. The current learning algorithms are simplifications of the complex process of learning. Although there is on-going research into learning algorithms, in this study the following the algorithms will be reviewed:

- Hebbian Learning Law
- Hopfield Law
- Delta Rule
- Kohonen's Learning Law

The Hebbian Learning Law was introduced **by** Donald Hebb in his book, *The Organization of Behavior,* in 1949. [24] The rule states that the weight of a connection between two neurons is strengthened if the two connected neuron mathematically have the same sign.

Building upon the Hebbian Learning Law, Hopfield Law denotes the magnitude of strength or weakness of connection. Hopfield Law states "if the desired output and the input are both active or both inactive, increment the connection weight by the learning rate, otherwise decrement the weight **by** the learning rate." [22] The learning rate is usually a learning constant between 0 and 1.

One **of the** most commonly **used** algorithms **is the** Delta Rule. **The** weights are adjusted **to reduce the delta or** difference between **the desired** output and **the** actual **output. This delta or error is** propagated **back** from **the** output **layer** through **the** hidden **layers.** With **this** algorithm **the** least **squared error is** minimized. **This** algorithm **is also called the** Windrow-Hoff **or the Least** Means **Square** Learning Rule. Neural networks **that utilize this** algorithm are **called feed** forward **back propagation** networks.

Another learning **rule is Kohonen's** Learning **Law.** This algorithm **searches for the neuron with** largest output. This **neuron is declared the** winner. **It will excite its** neighbors **while** inhibiting **its** competitors. **Since the Kohonen's** Learning **Law does not rely on the desired output, it can be used for unsupervised** learning.

3.10 Sensitivity Analysis

Determining **the** number **of inputs is critical to the modeling process. The number of inputs has a direct** bearing **on the** amount **of data that is needed** during training. **Since the** training process **as well as the data collection can be costly in** terms **of** money and time, **only** relevant **inputs should be used to avoid** redundant, irrelevant, **or** misleading **data.** The two types **of selection procedures to discern what are the** relevant **inputs are:**

- * Model-independent **approach**
- * Model-dependent **approach**

In **the** model-independent **approach,** the **statistical** merits **of the input** variables **affecting** outcomes are **directly evaluated** using **data.** In **the** model-dependent **approach, the inputs** are evaluated in conjunction with the learning algorithms. This means that the leaning algorithm is applied to the data and the subset of inputs that yields the best result for the algorithms is utilized in the modeling. Sensitivity analysis is an example of a modeldependent approach. Its goal is to remove any irrelevant or misleading inputs from a trained neural network.

Sensitivity analysis can be viewed as a procedure to determine how sensitive an output is to changes in an input. The analysis would check if a small change in the input can cause a significant change in the output. This would signal that the outcome is sensitive to the input. An input that has a significant affect on an outcome needs to be assessed very accurately. If the analysis shows that a large change in an output has little or no affect on an outcome, the input would be classified as having low sensitivity.

Sensitivity analysis can be applied to the entire set of data or to individual patterns. The following measures can be applied to the entire set:

- Delta Error Sensitivity
- Average Gradient Sensitivity
- Average Absolute Gradient Sensitivity
- Gradient Variance Sensitivity

The Delta Error Sensitivity, $S_{DE, i}$, defines a delta error measure in terms of the increase of the average square error of the network, TotalError_{ANN}, when an existing variable x_i (p) for any pattern p is replaced by its mean \bar{x}_i computed from all N_p patterns. This is given by:

$$
\overline{x}_i = \frac{1}{N_p} \sum_{p=1}^{N_p} x_i(p)
$$

The delta error sensitivity $S_{DE,i}$ of the input variable x_i is defined as:

$$
S_{DE,i} = \frac{1}{N_p} \sum_{p=1}^{N_p} S_{DE,i}(p)
$$

where

$$
S_{DE,i}(p) = TotalError_{ANN}(\overline{x_i}) - TotalError_{ANN}(x_i(p))
$$

The first summand of the above expression is the total square error when the training set is applied to the network.

If $S_{DE,i}$, is large, the input feature is sensible. In the case that $S_{DE,i}$ is small., the three additional sensitivity measures should be computed in order to determine if the input variable is useful.

The Average Gradient Sensitivity, $S_{AG, i}$, uses the average gradient of the artificial neural network's output, y, with respect to the specified input variable, x_i . This equation below denotes how *SAG,* i is calculated.

$$
S_{AG,i} = \frac{1}{N_p} \sum_{p=1}^{N_p} \frac{\partial y(p)}{\partial x_i}
$$

The Average Absolute Gradient Sensitivity, S_{AAG, i}, uses the absolute value of the average gradient of the ANN's output, y, with respect to the specified input, x_i . The Average Absolute Gradient Sensitivity can be calculated using the following equation:

$$
S_{AAG,i} = \frac{1}{N_p} \sum_{p=1}^{N_p} \left| \frac{\partial y(p)}{\partial x_i} \right|
$$

The Gradient Variance Sensitivity $S_{\text{GV},i}$, is calculated using the variance of the gradient of the output of the neural network to the specified input variable, x_i . This sensitivity measure is calculated using the following equation.

$$
S_{GV,i} = \sqrt{\frac{1}{N_p} \sum_{p=1}^{N_p} \left[\frac{\partial y(p)}{\partial x_i} \right]^2}
$$

If sensitivity analysis is done for an individual pattern as opposed to applying the analysis to an entire set, the recommended measures are:

- Delta Output Sensitivity
- * Output Gradient Sensitivity

If either measure is large for the given individual pattern, this shows that the output is greatly affected by changes in the input. In this case, the input variable plays an important role in the neural network's output.

The Delta Output Sensitivity $S_{DO,i}$, calculates the sensitivity of the an input variable to the an individual pattern **by** simply taking **the difference or the delta of** the two. **The following equation denotes the calculation for the** Delta Output **Sensitivity:**

$$
S_{DO,i}(p) = y(x_1(p), x_2(p), ..., x_i(p), ..., x_n(p)) - y(x_1(p), x_2(p), ..., x_i(p), ..., x_n(p))
$$

The Output Gradient Sensitivity S_{OG, i}, uses the gradient of the individual pattern to the **specified input** variable, **xi This** measure **is** determined **using the following:**

$$
S_{OG,i}(p) = \frac{\partial y(p)}{\partial x_i}
$$

Since the goal of sensitivity analysis is to find the inputs that are relevant **to the output, it is** important **to ensure that** irrelevant **inputs** and misleading **are** removed **as well as** eliminating redundancy. **The sensitivity** measures **that have been discussed** earlier **address this well but are not suitable** in **identifying** redundant **inputs.** Redundancy **is present when there is a well defined relationship** between two **or** more **input** variables. **In such cases, only one of the inputs is** necessary. **Statistical methods such as the least** means **square** method **(LMS)** are **useful in finding if there is relationship** between two variables.

3.11 Implementing the Risk Prediction Model

Building **on the** knowledge **gained** from **studying the** artificial **neural** network fundamentals **and** reviewing **the sensitivity** analysis techniques, **the risk prediction** model **was** designed **based on the** methodology **that was discussed** earlier.

Based on the theorems for universal approximators, the risk predictor model is an artificial neural network. The number of inputs is 46 parameters and the number of outputs is two. The number of inputs is based key parameters identified **by** the security experts at Microsoft. The team of experts are Certified Information Systems Security Professionals that are part of the Microsoft Security Team, Microsoft Technical Systems Engineering Team, Microsoft Circle of Excellence in Security Team, and the Microsoft Premier Support Services Team. The team agreed that the parameters that affect security are not only based on technology but need to include people and process as well. The sensitivity analysis is designed to ensure that the number of input parameters are relevant and not redundant. The number of outputs is two. One output identifies the risk associated with the production server environment while the other output is the risk associated with the client machine environment.

In this scenario there is no feed back so the model was designed as a feed forward network. Feedback refers to situations where the output affects the input. Feedback is very suitable for control systems as in the case of closed loop control systems where the feedback changes the input of the plant. In feedback systems, the input is partly affected **by** the output. In this study of work, the model is a black box where the security risks are the outputs and where these outputs do not affect the inputs.

3.12 Training

Although the network inputs are not affected by the outputs, the errors in the outputs are propagated back to best establish the weights and the biases during the training phase. This

type **of training is the back propagation** method which **refers to the error being propagated back** through **the** network **to adjust** the weights and **biases.** In this study **the error is defined as (Target-Output)2.**

The training **process** relied **on k-fold cross validation** which **is** useful **in** estimating **the generalization error** based **on** *resampling.* **[25-30] K-fold cross validation consists of dividing the data into** *k* **subsets of approximately equal size. The** network **is the then trained** *k* times using *k-I* **subsets. The** weights and **biases** are obtained **after each** training. The **averages for both the** weights and **biases are calculated after the** network **has been trained** *k* times. **The subset that is excluded is used for computing the error criteria.** In this study, **the** number **of folds was five so the** network **was trained** *5* times **using** 4 **datasets each time.**

3.13 Data Collection

An **essential** part **of building** an **approximator is the data collection because it is used to build, test** and **train** the **approximator.** The **collection of data was** two **fold: one set consisted of data generated by TableLab, a data generation tool developed at the Center for** Advanced Technology and **Education at Florida** International University, **and the other set was collected using a** Microsoft **website. The** reasoning **for doing this is to first validate the** methodology **with generated data before applying it to the actual data.**

Using the TableLab application, the datasets were randomly generated, yet they were purposely **created to** exhibit **statistical dependencies** amongst certain columns within the matrix **of data. The** columns are comparable **to the** parameters **in the questions** and the number of rows would reflect customer **responses.** The number of rows that were generated

using TableLab was 1000. The generated data contained **the** same number **of inputs and** outputs as the **datasets that were** collected **via a** secure **website hosted by** Microsoft **for interacting** with customers. This **site is hosting** an anonymous security questionnaire, and **a copy of the** questionnaire **can be** found **in Appendix A.**

The survey **included questions regarding** an organization's environment **such as people, process** and technology **as well as if the** organization **was** impacted by the **Sasser** worm and **to what extent was the** organization impacted. **This refers to** impacts **to the production servers** and **the client** machines. This information was **obtained via an** anonymous questionnaire **that** Microsoft's customers filled **out.**

The reason that the Sasser worm was selected for this study was that it impacted a sufficient number **of** customers **as well as having been** in **the past** year **so** customers **could** more **easily recollect its** impact. In **addition to this, its** impact **was defined as critical** which means **a** vulnerability **whose exploitation could allow an** Internet worm **to** be **propagated without user action. The Sasser** worm **and its** variants **exploits the Local** Security **Authority** Subsystem **Service (LSASS). The** Sasser worm **has the potential of affecting the** Microsoft Windows XP **client** machine and **the** Microsoft Windows **2000** Server and **the** Microsoft Windows 2003 Server machines. LSASS is responsible for the security mechanisms in the Microsoft **Windows operating** systems. **The** service verifies **that the user logons to the** computer **or sever** are **valid. LSASS is specifically responsible for** generating **the process that authenticates users** for **the Winlogon** service. **The** worm **is capable of exploiting a** buffer overrun vulnerability. A buffer overrun refers to passing data to a memory buffer which **is** smaller than **the** amount **of data that are being passed to it. The data that can be**

passed to the buffer can be malicious code. A successful exploitation of a buffer run could allow an attacker to execute malicious code. In the case of the Sasser worm, the exploit of the buffer overrun vulnerability can be accomplished remotely and can compromise the system. Depending on the version of the operating system, there is a difference on who can exploit the vulnerability. For example, the vulnerability can be exploited **by** an anonymous user on the Microsoft Windows 2000 and XP operating systems. If the operating system is Microsoft Server 2003 or Microsoft Windows XP 64-bit Edition 2003, the vulnerability can be exploited by local, authenticated users. Microsoft released a patch to address the vulnerability on April 13, 2004.

The survey contained questions pertaining to key security parameters such as people, process and technology which correspond to the inputs of the network. The survey also queried customers on the effects that the Sasser worm had on the production servers and on client machines. The responses to these questions corresponded to the outputs. The number of columns in the matrix data are 48 with 46 columns reflecting inputs and the last two reflected the outputs. The matrix of the raw data showing the actual responses can be found in Appendix B. Although there are more than 48 columns, some columns were actually compounded questions that were combined. They were merely separated to ease the querying of the customers. The data exhibited in Appendix B is the actual raw data whereas Appendix C displays the normalized data. The data were normalized so that all data would be between 0 and 1, inclusive. The number of columns in the matrix corresponded to the questions in the survey. The rows in the matrix of data *refer* to the companies' responses to the questionnaire.

CHAPTER 4

Implementation of the Prediction **Model**

In this chapter, the implementation of the prediction model will be discussed. The focus of this chapter is the implementation of the prediction model utilizing the dataset generated by the TableLab application. The design and development of a neural network can be complex and is often iterative until a satisfactory design is developed. The following steps are typical of any artificial neural network implementation and were utilized in this study:

- Determining how many layers are necessary
- Arranging the neurons in the various layers
- Selecting the type of connections between the layers as well as within a layer
- Deciding the strength of each connection by permitting the network to learn what are the suitable weights for each connection
- Determining how the neurons receive inputs and yield the outputs
- Selecting what is the most suitable training method

This chapter will detail how the artificial neural network was built, implemented, trained and tested using the generated data with TableLab. All models were built, trained and tested using the Neural Studio application. The analysis and comparisons were charted using Microsoft Excel. The discussion begins with the initial ANN configuration and it evolves to how the final and optimal configuration was achieved.

4.1 Determining **of** the Starting Configuration

Although the design process is generally based on trial and error, rules of thumb were considered. When designing the model, the two givens were the number of inputs and the number of outputs. The number of neurons in the input and output layers were equated to the number of input parameters and the number of output parameters, respectively. In this study, there were 46 inputs and 2 outputs.

The first effort was determining how many layers are needed. In this case, the number of layers was set to three: an input layer, one hidden layer, and an output layer. Since there were over 40 input parameters, the number of layers was kept to the minimum. Having such a large number of input parameters was already adding a level of computational complexity.

The rules of thumb were applied with respect to the recommended number of neurons in the hidden layer. The rules of thumb as offered by Hecht-Nielsen recommended using $2N+1$ neurons where N is the number of neurons in the input layer [17]. Applying the rules of the thumbs and using the fact that number of neurons in the input layer is forty-six, the recommended number of neurons in the hidden layer would be 93 neurons. This yielded a topology of 46-93-2.

The use of the rules of thumb was considered a good starting point, but the application of the rules of thumb needs to be verified and validation for this specific scenario and data. The rules of thumbs simply offer guidance, and there are counterexamples. [17] That is why there was much effort spent on verifying and testing the topologies to determine which offered the optimal solution. The tests that were performed in this investigation revolved

around the speed with which model can process and the speed with which the model yielded a testing error of less than 5%. The reason for selecting 5% as the error level is that 5% is a statistically accepted error value. In this case, the speed of processing refers to speed of training. This factor gave an measure of the computational complexity the model encountered. Throughout this study, the error is defined by the following equation:

$$
Error = (Target - Output)^2
$$

One thing to note is that the Neural Studio application provides a straight error value, but they can be treated as percentages since the datasets have been normalized.

When the 46-93-2 topology was tested for its ability to converge in a timely manner, it failed to do so. In fact during the testing phase, the topology was given over 90 minutes to converge, and it never yielded an *error* less than 90%. The testing consisted of letting the model run until an error of less than 5% was achieved or giving the model at least 90 minutes of running time to see how fast it would converge. In the 90+ minutes it had processed 191 iterations, yet the error was reduced from 113.95% to 90.92%. The other test was to check its processing speed. The model was checked to see how fast it can process 15 iterations. The 46-93-2 topology processed 15 iterations in **6** minutes and 17 seconds. Since this model did not produced a timely convergence, the number of neurons in the hidden layer were reduced to 80. The same testing criteria was applied to the new topology. Like its predecessor, the 46-80-2 topology did not produce an *error* of less than 5%. This topology yielded errors that were worse than its predecessor. After 1 hour and 42 minutes and 325 iterations, the error was at **223.25%.** The reduction in the number of neurons did lower the time it took to process 15 iterations from over six minutes to under five minutes. As part of the iterative process to determine the optimal topology, the number of neurons in the hidden layer was reduced by ten, thereby yielding a 46-70-2 topology. With this topology, an error of less than 5% was achieved in approximately 9.5 minutes and after 38 iterations. The 46-70-2 topology was able to process 15 iterations in 4 minutes. The number were again reduced in the hope of improving the speed with which the model would converge without impacting its ability to deliver within the error guidelines of less than **5%.** After the reduction in the number of neurons in the hidden layer, the resulting topology was 46-60-2. This one achieved the targeted error range in 7 minutes and 22 seconds using 36 iterations. The 46-60-2 topology was able to process 15 iterations in 3 minutes and 16 seconds. The process of removing 10 neurons from the hidden layer continued until there were only 20 neurons in the hidden layer. The reduction of neurons did result in improvements in the speed with which the model would process 15 iterations. For example, the 46-50-2 topology was capable of processing 15 iterations in 2 minutes and 29 seconds as opposed to the 46-20-2 topology that processed the same number of iterations in 54 seconds. When the rate with which the topology was able to achieve the targeted error range was analyzed, it showed that the 46-30-2 topology reached the error rate better than the 46-40-2 or 46-20-2 topology. The topology with 30 neurons produced less than 5% error after 16 iterations whereas the 46-20-2 topology needed 29 iterations. The 46-20-2 topology did offer the ability to process an iteration faster. The hidden layer was altered by ⁵neurons resulting in a 46-25-2 topology. The analysis of this topology showed that it was the optimal one. This topology achieved the error goal of less than 5% in 1 iteration. In addition to its ability to achieve the necessary error rate, it also demonstrated its ability to process 15 iterations in 66 seconds. The analysis with which the various topologies

achieved the targeted error range is captured in Table 4.1 and graphically displayed in Figure 4.1. Table 4.2 shows the speed with which each topology was able to process 15 iterations. Figure 4.2 provides a graphical representation comparing how fast each topology processed the 15 iterations. Figure 4.3 showcases the optimal artificial neural network topology that resulted.

				Error Difference	Delta Time
		Start Error	Stop Error	Stop-Start Error	Stop - Start Time
Topology	Iterations	$(\%)$	(%)	$(\%)$	(sec)
$46 - 93 - 2$	191	113.95	90.92	23.03	5733
$46 - 80 - 2$	325	237.27	223.25	14.01	6139
$46 - 70 - 2$	38	246.48	4.95	241.53	576
$46 - 60 - 2$	36	107.84	4.96	102.88	442
$46 - 50 - 2$	25	11.48	4.96	6.51	238
$46 - 40 - 2$	19	6.89	4.95	1.93	136
$46 - 30 - 2$	16	12.14	4.93	7.21	79
$46 - 25 - 2$		4.84	4.84	Ω	4
$46 - 20 - 2$	29	22.97	4.98	17.99	91

Table 4.1 Iterations needed to achieve an error $\leq 5\%$

Processing needed to achieve an error < 5%

Figure 4.1 Iterations needed to achieve an error $\leq 5\%$

Table 4.2 Processing speed of the different topologies

Time to process 15 iterations

Figure 4.2 Processing speed of the different topologies

 λ

Figure 4.3 Optimal Artificial Neural Network Topology

Once the topology was defined, the next step is determining what type of activation function needs to be used at the different layers and what the are the parameters associated with each of the activation functions. A dynamic approach was created to quickly narrow in on the optimal transfer function configuration and its parameters. The reason why the dynamic approach was developed was due to the number of combinations that would need to be reviewed to determine the optimal transfer function and its parameters. For instance, a Gaussian transfer function has two parameters that need to be evaluated. If the hidden and

output **layers both utilize the** Gaussian **activation** function and the **evaluation process** needed **to evaluate 10 options of each** parameter, **then 1000** combinations **would need to be evaluated for one** configuration **alone. There** are **four** types **of** configurations **that need to be evaluated. This approach first determines what is the** optimal configuration **followed by** analyzing **what** parameter **settings** are best **suited, thereby** substantially **reducing the** number **of** combinations **that need to be reviewed.**

The activation function **is the** transfer function **that processes the** cumulative sum **of the product of input values to a node and the** connections weights; **thereby, its correct selection is critical.** Analyzing **the input layer, the linear function is selected because** the **hidden layer should receive the input values without** any filtering. **The linear** function **best** matches **the output to its input as depicted in** Figure **4.4. The equation that defines the** linear function is $y= ax + b$. In Figure 4.4, a and b are 1 and 0 which yields $y=x$. [31]

Figure 4.4 Linear function where $y = x$

For the hidden **layer** and **the** output **layer, the desired** feature **for** the **transfer** function **is an ability to differentiate.** The reasoning **for selecting** an **activation** function **that is differential is to** fully satisfy **the** complexity **of the dataset. Activation** functions **such as the logistical sigmoidal** and **the** Gaussian functions are **considered good differentiable functions.** In **fact, the logistic** sigmoid **is often used in non-input layers for** networks **that rely on**

backpropagation. [32] An important consideration with respect to the output layer is selecting an activation function that provides smoothing. The differentiable functions such as the sigmoid have the ability to do smoothing.

The logistic sigmoidal function is defined by the following equation and depicted in Figure 4.5: [31]

Figure **4.5** Logistic **sigmoidal** function

A graph of the Gaussian function can be found in Figure 4.6. [31] This differentiable function is defined by the following equation:

$$
y = e^{-\frac{(x-b)^2}{2a^2}}
$$

1.0
0.5
0.0

Figure 4.6 Gaussian function

The next step **in the process is to** determine which **activation** function configuration **is best suited for** the **risk** predictor model. **It is** necessary **to** determine **what** differentiable function **is optimal for the hidden layer** and **for** output **layer. The** configurations **for the** input hidden **-** output **layers that were** analyzed are:

- **" Linear** Gaussian Gaussian
- **" Linear** Gaussian **-** Logistic **Sigmoidal**
- * **Linear -** Logistic **Sigmoidal -** Gaussian
- **" Linear** Logistic **Sigmoidal- Logistic Sigmoidal**

The analysis associated **with the selection of the** transfer function **consists of studying the errors after each iteration for the first ten iterations. The error** after **100 iterations was also reviewed.** In **essence, the** model's **ability to converge using the four different** configuration **was studied. Each** configuration **that was** analyzed had **the** same **topology, 46-25-2, that** was defined earlier **as** optimal.

After **10** and **100 iterations, the first** configuration, Linear **-** Gaussian - Gaussian, **exhibited** an **error of 91.84% down** from **92.40% after just** two **iterations. After 100 iterations the error was still at 91.84%. The error associated with the Linear -** Gaussian - Logistic **Sigmoidal** configuration **stayed relatively the** same. **After 1 iteration the error was 65.94%, but it increased** to **65.96% after 2 iterations.** The **error** remained **at 65.96% for the** remainder **of the 100 iterations. The** configuration, **Linear - Logistic** Sigmoidal **- Gaussian, never** converged. Throughout **the 100 iterations the error was 1218.01%,** and **it did not**

change. **The last** configuration, **Linear** - Logistic **Sigmoidal- Logistic Sigmoidal,** exhibited **the** best **convergence. The error was 4.84%** after **1 iteration; by the end of the test the error was** down **to** 3.38%. **The** configuration, **Linear -** Logistic **Sigmoidal- Logistic** Sigmoidal, **was** deemed **the** optimum **activation** function configuration **for the** input-hidden-output **layers. Table** 4.3 **lists the errors associated with the different activation** function configurations. Figure **4.7** graphically **depicts how well the** models **with the different** configurations **converged.**

		Activation Functions					
		linear-gaussian- gaussian	linear-gaussian- logsig	linear-logsig- gaussian	linear-logsig- logsig		
		92.40%	65.94%	1218.01%	4.84%		
	$\mathfrak{2}$	89.91%	65.96%	1218.01%	4.81%		
	3	91.84%	65.96%	1218.01%	4.78%		
	4	91.84%	65.96%	1218.01%	4.75%		
	5	91.84%	65.96%	1218.01%	4.73%		
Iterations	6	91.84%	65.96%	1218.01%	4.70%		
	7	91.84%	65.96%	1218.01%	4.67%		
	8	91.84%	65.96%	1218.01%	4.64%		
	9	91.84%	65.96%	1218.01%	4.62%		
	10	91.84%	65.96%	1218.01%	4.59%		
	100	91.84%	65.96%	1218.01%	3.38%		

Table **4.3** Errors associated with the different activation function configurations

figure 4.7 Errors associated with the different activation function configurations. The error is defined as $(T \text{arget-Output})^2$

Once the optimal transfer function configuration has been obtained, the next step is to select what are the parameters that are optimal for this study. Reviewing the equation associated with the logistic sigmoid, there are two parameters, a *and* b, that are configurable. Upon closer inspection, it can be noted that the parameter, b , only provides translation across the x-axis. Such translation is compensated via the weights and the biases. This means that parameter, b , can be set to 0 and that parameter, a , is the parameter that needs to be under investigation. In this study, the parameter was varied from 0.5 to 5 in increments of 0.5 . If the value is geater than 5, the function practically becomes a step function and is not differentiating as well. Table 4.4 showcases the errors, that were measured as the parameter is varied, after the marked iterations. A graphical depiction of the full set of measurements can be found in Figure 4.8. Evaluating the table and graph leads to the selection of parameter, $a = 2.0$. This setting yields the lowest error for the already optimal configuration.

Table 4.4 Comparison of Errors **in %** for activation function parameter selection

Figure 4.8 Comparison of errors for activation function parameter selection

4.2 Determination of the Optimum Predictor Configuration

The optimum predictor configuration is a three layer network consisting of 46 inputs, 25. neurons in the hidden layer and 2 outputs. The activation functions for the input layer, the hidden layer and the output layer are linear, logistic sigmoid and logistic sigmoid respectively. The key parameter in the activation function is set at 2.0.

After the optimum predictor configuration was obtained, the model was validated using 5fold cross validation. Cross validation refers to dividing the dataset into k subsets and using the $k-1$ subsets to train the model k times. The reason why k-fold cross validation was selected for this study is that it is designed to avoid over training and improve

generalization. One drawback with over training a model is that it fails to generalize. In overtraining scenarios, the model is essentially memorizing the patterns.

In this study, the dataset contained 1000 rows and 48 columns. The first 46 columns represent the inputs whereas the last two are the outputs. With the application of *5-fold,* cross validation, the 1000 rows dataset was subdivided into *5* subsets with each containing 200 rows. With 5-fold cross validation, 4 subsets are used for training for each of the five training sessions. This means that each training session uses 800 data rows, and the remaining 200 rows are used for testing.

4.3 Discussion **of** the Training Errors

The testing error and the training error were constantly monitored and compared in order to avoid over training. The training was stopped when the testing error was at a minimum. Each session has a set of biases and weights that are evaluated and recalculated throughout the training. When the training was stopped at each session, the weights and biases were recorded. This resulted in *5* sets of biases and weights which were then inputted into Microsoft Excel, and the averages for both the biases and weights were calculated and recorded. A listing of the weights and biases can be found in Appendix D. The average weights and average biases were then inputted into the Neural Studio application.

Table 4.5 denotes the training and testing *errors* that were recorded for each fold during the cross validation testing. The table clearly indicates that the training error was always below the targeted maximum of 5% and that the testing error was well below the desired maximum

value. As noted from inspecting Table 4.5, the errors indicate that the model was capable of generalizing and did not have the over fitting problem.

	Training Error	Testing Error	
Training Set	$\mathcal{O}(6)$	$(\%)$	Iterations
	4.641	0.673	
	4.585	0.800	
	4.385	0.828	
	3.980	0.763	
	3.649	0.569	

Table 4.5 Errors associated with Cross Validation

CHAPTER 5

Implementation of the Practical Predictor Model

This chapter is dedicated to the analysis and discussion on how the predictor model was implemented using real data. The chapter discusses how the real data was obtained, how the normalization procedures were applied to the datasets, as well how the predictor model was implemented, trained and validated.

The knowledge that was gained from building a model based on the data generated from TableLab was leveraged in this effort. Key insights that were gained from this implementation of the model, that was built using the generated data, were applied in the practical application. Some of these key insights include:

the number of layers were set to three: one input layer, one hidden layer, and one output layer

- the use of a linear function for the input layer's transfer function
- the use of differentiable functions in the hidden and output layers' activation functions
- the dynamic approach for discerning the activation functions' parameters
- \bullet the application of k-fold cross validation
- the need for a generalization test

5.1 Data Collection **and** Processing

Since collection of the data is an instrumental part of implementing a practical prediction model, the first endeavor was to determine how to best collect data. In this study, it was necessary to provide a mechanism that would let companies candidly share their experiences with security breaches. Although there is much in the media about security breaches such as the computer viruses and worms that have attacked the information systems of businesses globally, there is no formal collection of the data that showcases the extent and impact of the attacks or even profiles the information technology infrastructure of the impacted businesses. The survey which can be found in Appendix A was devised for this study and is the first step in collecting such information. The creation of the survey was the result of a multitude of activities which ranged from reviewing forms that queried the technical controls that customer implement in their environments to interviewing security professionals as to what they saw were key factors to having a secure information system. The survey was hosted on a Microsoft web site and was designed to be anonymous so customers could comfortably fumish information without divulging any association with their specific company. Another thing to note when inspecting the survey is that the questions regarding impacts associated with a security breach only focused on the impact of the Sasser worm. The Sasser worm was specifically picked because of its critical nature and its relevant recent activity making it easier on the customer filling out the questionnaire. The responses were posted to a database and a spreadsheet was used to access the database. Appendix B details the raw responses that the customers furnished. When the responses of the survey are reviewed, intra subject inconsistency was observed. For instance, a customer replied having an elevated maturity level in the security personnel, yet the company did not

73

have security personnel **with** any **certifications or provided** security training **to its** employees.

After **reviewing the** survey **questions, it can be noted that** some **questions** exhibited **certain** dependencies **on** other **questions. For instance, the question** which queried **whether the** customer performed software **code reviews is** dependent **upon the question that** determines **if the** customer **developed** any in-house **applications.** Questions **such** as these **were** treated **as** compounds **questions. The following topics were treated as** compound **questions which** are **dependent upon the** company developing in-house **applications:**

- **" the** requirement **of** software **code reviews**
- * **the** testing **of** software **by a quality** assurance department
- the strict validation of input
- * **the** handling **of error** messages
- the use of threat modeling

Another compound parameter **is whether the** anti-virus signatures **are up to date. This** parameter **is dependent upon the use of** anti-virus software. **Likewise the questions that** pertain **to having the firewalls** which **block** unnecessary **ports are dependent upon the use of firewalls.**

This compounding **affected the** normalization procedure and **the** number **of** parameters **in the** input **layer of the** predictor model. After compounding **was** applied, **essentially, the**

74

number of questions in the survey pertaining to the information system technical infrastructure, people, and process corresponded to the number of inputs in the model and the number of questions pertaining to the extent of the impact that the Sasser worm effected in terms of percentage denoted the number of outputs in the model. The final result was 46 inputs and 2 outputs.

The normalization of the data was key in the development of the model. Since the desired outputs are risks, and risks are best denoted as percentages, the input data was also normalized to fall between 0 and 1, inclusively. This normalization process was designed to simplify any internal mathematical processing as in the case of applying the transfer functions. When inspecting the customer responses in Appendix B, it can be noted that certain responses were in days as in the case of how often are passwords reused. This resulted in utilizing the following normalization equations:

$$
1 - \frac{x}{worst}
$$
 if x \le worst
0 if x> worst

where x is the customer response and worst is the industry accepted worst case value for the security parameter. For instance, patches to virus are often released 6 months prior to the release of a virus that exploits the vulnerability. If for some reason, the company exceeds the industry worse case value, the result is capped at 0 and is not allowed to go less than 0. In this study, a value of θ in the input parameter signifies the worst case scenario and a value of 1 indicates the best possible response. In this study, the yes responses were given a value of 1 during the normalization process, and the *no* responses were assigned a value of 0. If the customer fail to answer a question associated with an input parameter, the response to the

question was treated as a θ . With respect to questions that reflect the impact of the virus, the customer would only furnish percentages and did not give specifics on the actual number of machines that were hit. The percentages was used in the analysis and not the other number. It was asked to ascertain the size of the company from an infrastructure perspective.

Parameters related to the pay scales of the security personnel and the IT professionals were affected during the normalization and compounding. For instance, customers that classified the scale as according to industry standard were issued a value of 0.5. If the customer defined the pay scale as below industry standard, a value of 0 was assigned as opposed to issuing a value of 1 to companies paying above the industry standard. Another parameter that was effected **by** the compounding and normalization procedures was the desktop manageability. Two customers did provide a response to having a managed desktop environment, but they failed to provide a provide a percentage indicating how many desktops are managed. In this case where there is a compounding effect, those two customers were each assigned a value of 0.50 for management desktop since they answered half of the two compounded questions.

During the normalization process, if a customer responded that they had deploy a technology as in the case of digital rights management but were in the midst of testing, a value of **0.5** was assigned. These questions expected a yes or no answer which translates to a 1 or 0 so a testing response was given to signify the middle value. This same process was applied to responses pertaining to people and process as in the question on whether the passwords are changed on domain accounts when an IT administrator is terminated. One customer responded sometimes so a value of 0.5 was given.

76

The analysis that took effect as a result of the compounding process and the results of the normalization process can be found in Appendix C.

5.2 Development **of** Practical Predictor

The methodology that was validated with the generated data was then utilized to build the practical predictor model using the customer responses after the normalization procedures were applied to the dataset collection. In this scenario, certain givens are noted. These givens are the number of neurons in the input layer being 46 and the output layer having 2 outputs. The number of layers were kept to three meaning that there is only one hidden layer. Once these givens have been identified, the methodology focuses on determining the optimum topology by determining how many neurons should be in the hidden layer.

In search of the topology that best suits this specific scenario, the rule of thumb concerning the number of neurons in the hidden layer was first applied. Such an application resulted in the 46-93-2 topology.

The process of verifying the topology consisted of analyzing how fast the topology would achieve an error less than or equal to *5%* and how fast the topology can process 15 iterations of data. The latter measure denotes how fast the model is trained, yet it can be used as an indication of the computational complexity of the model. Table *5.1* showcases how quickly the topologies achieved the desired *error* level, and the Figure *5.1* provides a graphical view of this measure. To compare the different topologies in terms on how fast the topologies are capable of processing data, Table *5.2* and its graphical counterpart, Figure *5.2,* are provided.

Inspecting the **values** associated with **the 46-93-2** topology, **it can** be **noted that it** only **took one** second **to** achieve **the** desired **error level, yet the** processing speed **is 2.14 iterations/sec.** Since **the** number **of** inputs and outputs are **given values, the only factor that can** be changed **is the** number **of** neurons **in the** hidden **layer.** Reducing **the** number **of neurons in the** hidden **layer to 80 resulted** in **the 46-80-2** topology. This **topology did not converge. Even after 1000 iterations, the error reduced only** marginally from **22.50% to 22.27%. The processing associated** with **the topology did** improve **to 2.50 iterations/sec. The topology was again** changed **to 46-70-2. This** configuration **was able to converge, but it took 4 iterations and 30 seconds to** achieved this as **opposed to the 46-90-2 topology which required only 2 iterations** and **1 second. The processing speed of the new** topology **increased to 3 iterations/sec. A new topology having 60** neurons **in the hidden layer was** analyzed. **This one required 7 iterations to reach the targeted error level.** The **processing speed** improved by 25%. The 46- 50-2 topology **displayed** more **reduction in** terms **of the** number **of iterations needed to** meet **the desired error level, yet its processing speed stayed on par with 46-60-2 topology. The** 46-40-2 **topology required only one iteration to achieve the 5% error level,** and its processing speed jumnped to 5 iterations/sec. Reducing the number of neurons in the hidden **layer to 30 resulted in** an **increase in the** number **of iterations needed to** meet **the targeted error level.** This time **it took 2 iterations to get to** an **error of 3.53%. The associated processing speed is 7.50 iterations/sec. The reduction of 10** neurons **in the** hidden **layer yielded the 46-20-2** topology. This **one required 3 iterations to reach the desired error level. The** processing **speed of 15 iterations/sec is 100% better** than **its** predecessor. A 46-25-2 configuration **was created** and analyzed **to see how** well **it** addressed each metric. **This** configuration required **only 1 iteration to** achieve an **error of 2.80%,** and **at the** same time **its**

processing speed was 15 iterations/sec. Since this topology excelled at each metric, it was deemed the optimal one. The optimal model can be seen in Figure 5.3. One thing to note is that this configuration had an error of only 2.80%, yet the 46-40-2 configuration had a 2.04% error.

		Start Error	Stop Error	Error Difference Stop-Start Error	Delta Time Stop – Start Time
Topology	Iterations	(%)	$(\%)$	$(\%)$	(sec)
46-93-2		18.26	3.05	15.21	
$46 - 80 - 2$	1000	22.50	22.27	0.23	422
$46 - 70 - 2$	4	46.85	3.18	43.67	30
$46 - 60 - 2$		46.00	3.21	42.79	θ
$46 - 50 - 2$	$\overline{2}$	5.16	2.88	2.28	θ
$46 - 40 - 2$		2.04	2.04	0.00	Ω
$46 - 30 - 2$	ົ	10.10	3.53	6.56	
$46 - 25 - 2$		2.80	2.80	0.00	Ω
$46 - 20 - 2$	3	19.10	2.71	16.38	

Table 5.1 Iterations needed to achieve an error $\leq 5\%$

Processing needed to achieve an error *<* **5%**

Figure 5.1 Iterations needed to achieve an error $\leq 5\%$

Table 5.2 Processing speed of the different topologies to iterate through 15 iterations

			Error	Delta Time	
			Difference	$Stop - Start$	Processing Speed
	Start Error	Stop Error	Stop-Start Error	Time	15/Delta Time
Topology	$(\%)$	$(\%)$	(%)	(sec)	(Iterations/sec)
$46 - 93 - 2$	113.95	93.29	20.66	377	2.14
$46 - 80 - 2$	237.27	224.70	12.56	291	2.50
$46 - 70 - 2$	246.48	9.42	237.06	240	3.00
$46 - 60 - 2$	107.84	7.43	100.41	196	3.75
$46 - 50 - 2$	11.48	5.81	5.67	149	3.75
$46 - 40 - 2$	6.89	5.28	1.61	113	5.00
$46 - 30 - 2$	12.14	5.01	7.13	87	7.50
$46 - 25 - 2$	4.84	4.50	0.34	66	15.00
$46 - 20 - 2$	22.97	5.85	17.13	54	15.00

Time to process 15 iterations

Figure 5.2 Processing speed of the different topologies to iterate through 15 iterations.

Figure 5.3 Optimal Artificial Neural Network Topology

After the optimal topology was found through empirical analysis, the next step was to determine what activation function is optimally suited for each layer. One given in this respect is the use of the linear function for the input layer. The reasoning for this is that the use of a linear function in the input layer provides the entire input to the hidden layer. The configurations in question are what type of differentiable function should be used in the hidden and output layers. The reasoning for using a differentiable function is that they are **best suited for** models **that rely on back** propagation **for** training. **The configurations for** the input **-** hidden - output **layers that were** analyzed **are:**

- * **Linear -** Gaussian Gaussian
- * **Linear -** Gaussian **Logistic Sigmoidal**
- * Linear **Logistic Sigmoidal -** Gaussian
- **" Linear Logistic Sigmoidal- Logistic Sigmoidal**

The analysis **consisted on** monitoring **each of the first ten iterations to** determine **how well the** model **was capable of reducing the error. The** model **was again check after 100 iterations to see what the error was. This is** charted **in Table 5.3** and graphed **in** Figure **5.4. The table** and graph **denote that the use of the logistic sigmoidal** function **in the both the hidden** and output **layers yielded the best results. The error was reduced** from 1.71% to 0.37% **over the** course **of 100 iterations. The** linear-logistic **sigmoidal-gaussian** configuration **had** a **lower error,** 1.36%, **at the** start **of the test, yet** this **error did not reduce** through **the** course **of** the **test. The other** two configurations **did** indicate **reductions** in **error, but the errors were larger at the** beginning and **at the end of the test when** comparing **with** linear-logistic **sigmoidal-logistic sigmodal** configuration.

Table 5.3 Errors associated with the different activation function configurations

Figure 5.5 Errors associated with the different activation function configurations.
The error is defined as $(Target-Output)^2$

Once the optimal topology and transfer function configuration were determined, the methodology called upon the dynamic approach to determine what was the value of the activation function parameter that best suited the dataset. The dynamic approach consisted of determining the activation function first and then narrowing in on the value of the activation parameter without having to review thousands on configurations. **By** doing so the combinations in this case is reduced to 10 different values of the activation parameter.

Recalling that the logistic sigmoidal function is defined using the equation:

$$
y = \frac{1}{1 + e^{-a(x-b)}}
$$

In this study, the parameter of interest is a since the parameter, b, simply provides a translation across the horizontal axis. Such translation can be accounted by the biases and weights.

Table 5.4 and Figure 5.5 showcase the effects in terms of error that model has as the activation function parameter, a , is changed This parameter is changed from 0.5 to 5.0 in increments of 0.5. Such a range was suitable in that values of a below 0.5 yielded very little smoothing and appeared almost linear and values greater than 5.0 resulted in the logistic sigmoidal function nearing a step function. This means that for values outside this range, the sigmoidal did not exhibit its differentiating qualities as necessitated by this study. The test consisted of monitoring the error after each of the first ten iterations then monitoring the error after every 10 iterations until 100 iterations were performed.

Reviewing Table *5.4* and **its corresponding** graph, Figure *5.5,* **it can** be **noted that the** model having an **activation** parameter **equal to 3.5** exhibit **the** smallest **error** throughout **the test. The error at the** start **of** the **test was 0.13%** and **at the end it was** measured **at 0.07%.**

		Activation Function Parameter, a									
		0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
	$\mathbf{1}$	4.13	1.35	0.61	0.37	0.30	0.18	0.13	0.45	0.51	0.56
	$\overline{2}$	3.51	1.11	0.57	0.37	0.29	0.16	0.13	0.41	0.48	0.54
	3	3.04	0.99	0.55	0.36	0.28	0.15	0.13	0.36	0.45	0.52
	4	2.67	0.92	0.54	0.36	0.27	0.15	0.12	0.32	0.41	0.50
	5	2.39	0.88	0.53	0.35	0.26	0.15	0.12	0.28	0.38	0.47
	6	2.17	0.85	0.52	0.35	0.26	0.14	0.12	0.25	0.33	0.45
	$\overline{7}$	2.00	0.83	0.52	0.35	0.25	0.14	0.12	0.23	0.29	0.42
	8	1.86	0.82	0.51	0.34	0.25	0.14	0.12	0.22	0.27	0.38
	9	1.74	0.81	0.51	0.34	0.25	0.14	0.12	0.21	0.25	0.33
Iterations	10	1.65	0.80	0.50	0.34	0.24	0.14	0.12	0.21	0.24	0.30
	20	1.21	0.75	0.47	0.30	0.22	0.12	0.11	0.18	0.20	0.23
	30	1.10	0.72	0.44	0.27	0.19	0.11	0.11	0.17	0.18	0.21
	40	1.06	0.70	0.41	0.25	0.18	0.11	0.10	0.16	0.17	0.18
	50	1.04	0.67	0.38	0.23	0.16	0.10	0.09	0.15	0.16	0.17
	60	1.02	0.65	0.36	0.21	0.15	0.09	0.09	0.15	0.15	0.16
	70	1.01	0.63	0.34	0.19	0.14	0.09	0.08	0.14	0.15	0.15
	80	1.01	0.62	0.32	0.18	0.14	0.09	0.08	0.13	0.14	0.15
	90	1.00	0.60	0.30	0.17	0.13	0.08	0.08	0.13	0.13	0.14
	100	1.00	0.58	0.28	0.16	0.12	0.08	0.07	0.12	0.13	0.13

Table 5.4 Comparison **of** errors in % **for** activation function parameter selection

Figure 5.5 Comparison of errors for activation function parameter selection

.3 Determining the Optimum Configuration

The resulting configuration of the development phase is a model that consists of three layers:

- an input layer containing 46 neurons and using a linear function, $y=x$ for its transfer Ġ function
- a hidden layer containing 25 neurons and relying on the logistic sigmoidal function ۰ for the activation function where $a=3.5$
- and an output layer containing 2 neurons and utilizing the logistic sigmoid for a transfer function where $a=3.5$

The methodology calls for a training phase. In this study, k-fold cross-validation was utilized for obtaining the optimum weights and biases. This validation is useful when there are a small sets of data and when back propagation is used as a training method. There are 21 patterns, and the weights and biases are adjusted according to the errors that are generated and propagated back through the layers. These conditions lend itself to the use of k-fold cross validation. The fact that cross-validation proved to be useful when the model was trained using generated data was another reason k-fold cross validation was used. In this case, the number of folds is five. Four of the subsets have a size of 4 pattems and the last subset contains five pattems.

Recalling that k-fold validation refers to training with $k-1$ subsets k times. The subset not used during the training is reserved for testing. The testing and training errors are compared. In this test, the training continued as the training errors reduced when testing error would increase, the training was stopped. The reason for stopping at this point was to avoid over training a network which results in the network unable to generalize and tends to simply memorize the pattems. Table 5.5 contains the testing and training errors associated with each of the 5 times that the network was trained. When the training was stopped, the biases and weights were recorded. The biases and weights from the five sessions were averaged. The average of the biases and weights can be found in Appendix **E.**

	Training Error	Testing Error	
Training Set	$\%$	$\frac{6}{6}$	Iterations
	0.13	0.03	
	0.13	0.36	
	0.42	0.18	
	0.37	0.20	
	0.43	0.28	

Table 5.5 Errors associated **with** cross validation

CHAPTER **6**

Discussion of Results

The focus of this chapter is a detailed discussion of the results. In this dissertation two models were built. The first model was to built to validated the methodology and built with generated data, and the latter model was built upon the validation of the methodology using the datasets that were collected using the Microsoft hosted website. This chapter is dedicated to the discussion of both implementations.

6.1 Validation of **the Prediction Model**

A generalization test was then run to determine how capable was the final model of generalizing. The sum of the errors as determined **by** the generalization test are 1.91 for the risk associated with the production server environment and 2.26 for the risk pertaining to client machine environment. The averages of the these errors, as calculated **by** the sum of the errors divide **by** the total number of patterns in the testing test which is 1000, are 0.0019 and 0.0023, respectively. The minimum of the errors and the maximum of the errors across all outputs is 0.00 and 0.10. The average error across outputs was calculated to be 0.004.

To validate the predictor, different patterns was randomly selected from the dataset and applied to the **ANN.** For each pattern, the prediction of the associated risks for each one of the categories was computed as the output of the two last nodes of the **ANN.** Those outputs were compared with the targets. The analysis with respect to the risk associated with the production environment are shown in Table 6.1, and the analysis of the errors associated with the risk to client machine environment can be found in Table 6.2.

Table **6.1** Errors associated with the production environment risk analysis

Table 6.2 Errors **associated** with the client machine environment risk analysis

6,2 Validation of the Practical Predictor

A model was built using the average of the weights and biases that were obtained during the k-fold cross validation. As part of the validation of this newly built model, a generalization test was performed. This generalization test was designed to ensure that the model was capable of generalizing and that it had not been overtrained. The sum of the errors as calculated during the generalization test was 12.68% for the risk associated with the production environment and the 17.34% for the risk pertaining to the client machines. The

average of the errors was also obtained by dividing the sum of the *errors* by **the** number **of** patterns, **21.** The **averages of the errors** are 6.04% **for the** production environment **related risk** and **8.26% for the risk associated with the client** environment. **The average of the errors across all** outputs **was** found **to be** 1.43%. The minimum **error across all of the** outputs **was 0.00% as opposed to the** maximum **which was calculated to be** 7.94%.

	Production Risk Target	Production Risk Computed	Production Risk Error
Pattern Number	$(\%)$	$(\%)$	$(\%)$
	5.00	0.00	0.63
$\overline{2}$	0.00	0.00	0.04
3	0.00	0.00	0.49
4	0.00	0.00	0.89
5	0.00	0.00	0.20
6	0.00	0.00	0.03
$\overline{7}$	0.00	0.00	0.13
8	0.00	0.00	0.86
9	0.00	0.00	0.75
10	10.00	0.00	1.48
11	0.00	0.00	0.06
12	0.00	0.00	0.12
13	1.00	0.00	0.19
14	0.00	0.00	0.07
15	0.00	0.00	0.11
16	15.00	000	3.49
17	5.00	0.00	0.47
18	10.00	0.00	72.54
19	0.00	0.00	0.04
20	0.00	0.00	0.04
21	0.00	0.00	0.05

Table **6.3** Errors associated **with** the production environment risk analysis

Another part **of** the **validation of the** predictor **is to see how well the** model would **predict the risk to the production** environment and **to the client** machine environment. **The** patterns **used in the** training were **again** submitted. **The** outputs **obtained** from **the** ANN **were** compared **against the targets as defined in the datasets. Table 6.3 contains the risk to the production** environment as computed **by the** model and **furnished by the** pattern. Likewise

risk to the **client** environment was computed and compared **with the** target **risk value** furnished by **client.** This comparison **is** detailed **in Table** 6.4.

	Production Risk Target	Production Risk Computed	Production Risk Error
Pattern Number	$(\%)$	$(\%)$	$(\%)$
	5.00	0.00	0.50
\overline{c}	1.00	0.00	0.10
3	1.00	0.00	0.04
4	0.00	0.00	0.59
5	0.00	0.00	0.07
6	0.00	0.00	0.04
7	0.00	0.00	0.58
8	0.00	0.00	0.54
9	0.00	0.00	0.46
10	10.00	0.00	1.50
11	3.00	0.00	0.12
12	0.00	0.35	0.00
13	1.00	0.00	0.31
14	7.00	0.00	0.95
15	1.00	0.00	0.97
16	80.00	51.82	7.94
17	5.00	0.00	0.48
18	80.00	2.54	0.56
19	0.00	0.00	0.04
20	10.00	0.00	1.48
21	0.00	0.00	0.06

Table 6.4 Errors associated **with** the client machine environment risk analysis

6.3 Risk Predictor Tool

After analyzing the results of the risk predictor model **in** terms **of how well did the** model **predict the risk to the client** and **production** companies, **a risk predictor was built based on the** final **artificial** neural network model. This **tool is** graphical **user** interface (GUI) that allows **users to input characteristics of their** production and **client** environments. Based **on these inputs, the tool** will normalize **the** inputted **data to values** between 0 and **1** and provide two **results, the risk to the production** environment and **the risk to the client** environment. The following figure **is a screen shot of the tool.**

Figure 6.1 Risk Predictor Tool

CHAPTER **7**

Conclusion

The contribution of this work has significant impact in the area of security. It demonstrates that obtaining an actuarial risk can be easily done and does not have be time consuming. Traditionally, risk assessment tools simply rank and order risk and do not provide an actual risk. Obtaining an actual risk has been customarily a time consuming and costly effort because it requires a team of experts to review the information systems and attempt to calculate the cost of not having these systems available. Such an assessment has typically been so difficult that it is often not performed. The risk assessment tool that was developed as part of this study provides an actuarial risk in a very timely manner. In fact, the predictor tool simply requires a few minutes of time to input information about information systems' environment. As part of the risk assessment, the tool distinguishes between the risk to a environment that is a production ready mission critical environment from the client machine environment.

During this study, a holistic approach encompassing people, process, and technology to security was undertaken. Historically, security issues related to information systems have been focused only on technology. In this study, the technology was only factor. This *work* accounted for people and process in addition to technology. Although this work used data from customers using Microsoft technology, the model can assess risk to the environments regardless of the platform vendor. In fact, the customers were queried with respect to technology but at no time was a vendor specific query made.

Another significant result of this work is the methodology that was provided to build the security risk predictive model. The methodology provided a step **by** step approach on how to build, implement, and validate a predictor. As part of the methodology, much emphasis was placed on the collection of the data. Since the predictor was designed for assessing security risks, there was great apprehension with the obtaining data. Customers were repeatedly assured of the confidentially of the data., and an anonymous website was created for data capture.

As the model was built, this work at first relied on the recommendations that the rules of thumb provided. The rules of thumb are a good starting point, but they do not take into account the number of training sets or the complexity of the data. Rules of thumb tend to focus on the simplicity and the quantitative accuracy of the data. In this case, the rules of thumb recommended using a 46-93-2 topology, but this model was not suitable in this work. This study empirically revealed that a 46-25-2 topology was optimal.

When selecting the appropriate activation function for the hidden and output layers in this study, the empirical process revealed that the logistic sigmoid was a better fit than the Guassian function even though they are both differentiable functions. Gaussian functions are traditionally suited for clusters of data and classification scenarios. Clusters are mostly used when trying to classify data. In this study of work, the goal was to build an approximator, and it did not involve looking at clusters of data. This is why it was not surprising that the Gaussian function was proven that it was not the optimal activation function for the datasets. In this study, the Gaussian function was not a good choice for either of the two approximators that were built as evidenced **by** the results. The logistic sigmoidal function proved to be a better choice for the models that were built.

Although this work has made great strides in taking a proactive stance towards security, there are future enhancements that can be made. For instance, much of the responses to the survey were yes or no answers. At times, customers did respond with *sometimes*, *almost*, mostly instead of a blanket yes or no. These answers were mapped to values, yet the application of fuzzy logic could be used to better capture such responses. In the survey, the format for responding was free format instead of a simple binary selection in order to give customers the flexibility to provide such answers. The reasoning for this was to best capture the customer information without provide any filtering.
References

- 1. "ILOVEYOU virus", http://searchsecurity.techtarget.com/sDefinition/0,sid14 gci21_4542,00.html, May 10, 2000.
- 2. Knowles, **D.** Perriot, F. and Szor, P., "W32.Blaster.Worm", http://security response.symantec.com/avcenter/venc/data/w32.blaster.worm.html, Symantec, February **26,** 2004.
- 3. Shimonski, Robert J.,"Trojan Horse Primer", http://www.windowsecurity.com /articles/Trojan Horse Primer.html, July 23, 2004.
- 4: Richardson, Robert, "2003 **CSI/FBI** Computer Crime and Security Survey", http://i.cmpnet.com/gocsi/db area/pdfs/fbi/FBI2003.pdf, Computer Security Institute, July 2003.
- *5.* Norah, Laurence, "Security Attacks Cost Business Billions", http://itvibe.com /default.aspx?NewsID=2494, I.T. Vibe, April 28, 2004.
- 6. Lazarus, David, "Wells in a World of Woe after Theft", http://www.sfgate.com/ cgi-bin/article.cgi?file=/chronicle/archive/2003/11/21/MNGLT37MH71.DTL, San Francisco Chronicle, November 21,2003.
- 7. Debar, H., Becker, M., and Siboni, D., "A Neural Network Component for an Intrusion Detection System", *Proceedings of the IEEE Symposium on Research in Security and Privacy,* Oakland, CA, pp. 240-258, May 1992..
- 8. Andress, Amanda, "Surviving Security: How to Integrate People, Process, and Technology", Second Edition. Auerbach Publications, December 2003.
- 9. McDowell, Mindi, "Avoid Social Engineering and Phishing Attacks" http://www.us-cert.gov/cas/tips/ST04-014.html, United States Computer Emergency Readiness Team, July 28, 2004.
- 10. CISSP book **C.** M., "Discrete Fourier Transforms when the Number of Data Samples Is Prime," *Proceedings of the IEEE,* Vol. *56,* June 1968, pp. 1107-1108.
- 11. Patterns and Practices. "Security Operations for Microsoft Windows Server". Redmond, Washington: MS Press, 2002.
- 12. DeZulueta, M., Adjouadi, M., "Using Threat Modeling When Architecting a Healthcare System", Proceedings of the 8th World Multi-Conference on Systematics, Cybernetics and Informatics (Orlando, Florida, July 18-21, 2004).
- 13. Howard, Michael and LeBlanc, David. "Writing Secure Code", Redmond, Washington: MS Press, 2003.
- 14. Meier, J. **D.** and etal, "Building Secure Microsoft **ASP.NET** Applications". Redmond, Washington: MS Press, 2003.
- 15. Funahashi, K.. "On the approximate realization of continuous mappings **by** neural networks". Neural Networks, **2:183-192, 1989.**
- 16. Hecht-Nielsen, R., "Kolmogorov's mapping neural network existence theorem". *Proceedings of the International Conference on Neural Networks,* Volume 3, pages 11-14, New York, 1987. IEEE Press..
- 17. Hecht-Nielsen, R.,. "Theory of the back-propagation neural network". *Proceedings of the International Joint Conference on Neural Networks,* Volume 1, pages 593-608, New York, 1989. IEEE Press.
- 18. Homik, K., "Approximation capabilities of multilayer feedforward networks". *Neural Networks,* 4:251-257, 1991.
- 19. Kurkova, V., "Kolmogorov's Theorem and multilayer neural networks". Neural Networks, 5:501-506, 1992.
- 20. Homik, K., Stinchcombe, M., and White, H.,. "Multi-layer feedforward networks are universal approximators". *Neural Networks,* 2:359--366, 1989.
- 21. Sprecher, **D.** A., "One the structure of continuous functions of several variables". *Transactions of the American Mathematical Society,* 115:340- *355,* 1965.
- 22. Shalvi, Doron, "An Introduction to Artificial Neural Networks", http://www.ee.umd.edu/medlab/papers/nnsumm/nnsumm.html, September 30, 1997.
- 23. Fausett, Laurene. "Fundamentals of Neural Networks. Architectures, Algorithms, and Applications". Englewood Cliffs New Jersey: Prentice-Hall Inc., 1994.
- 24. Hebb, Donald, "The Organization of Behavior". John Wiley *&* Sons, Inc., 1949.
- *25.* Weiss, **S.M.** and Kulikowski, C.A., "Computer Systems That Learn", Morgan Kaufmann. 1991.
- 26. Efron, B. and Tibshirani, R.J. "An Introduction to the Bootstrap", London: *Chapman &* Hall. 1993.
- 27. Efron, B. and Tibshirani, R.J. (1997), "Improvements on cross-validation: The .632+ bootstrap method", *Journal of the American Statistical Association, 92,* 548-560.
- 28. Shao, J., "Linear model selection **by** cross-validation," *Journal of the American Statistical Association,* **88,** 486-494. 1993.
- **29.** Cowan, **J.D.,** Tesauro, **G.,** and Alspector, J. (eds.) "Advances in Neural Information Processing Systems", San Mateo, CA: Morgan Kaufman, pp. **391- 398.** 1994.
- **30.** Hjorth, J.S.U., "Computer Intensive Statistical Methods Validation, Model Selection, and Bootstrap", London: Chapman & Hall. 1994
- 31. The Math Works, Inc. "Neural Network Toolbox", http://www.mathworks.com /access /helpdesk/help/toolbox/nnet/, 1994 2004.
- 32. Spencer-Smith, Richard, "Neural Networks", http://www.mdx.ac.uk/www/ psychology/cog/psy3250/Index.htm, 1996.

APPENDICES

Appendix A

Security Questionnaire

This questionnaire is designed to provide data to train a neural network that predicts the level of risk to virus attack.

1. Was your organization impacted by the Sasser worm? Yes or No

2. How many servers were hit? Percentage of server hit to total number of servers

3. How many clients were hit? Percentage of clients hit to total number of clients

4. How long was your organization working to address the impact of the virus after detecting being hit by the virus? Number of days spent patching, testing, etc.

Environment questions

1. Does your organization have a managed desktop environment? Yes or No If yes, what percentage

2. Do your organization have a centralized management/administration of Information Technology services? Yes or No

3. Are patches tested before being deployed on a server? Yes or No

4. Time to roll out a patch to server? Days between a patch is announced and it is deployed. This may be an average number of days

5. Are patches tested before being deployed on a client? Yes or No

6. Time to roll out a patch to client? Days between a patch is announced and it is deployed. This may be an average number of days

7. Is there security awareness training for employees? Yes or No If yes, is training done periodically? Yes or No

8. Is there specialized security training for security professionals? Yes or No If yes, is the training done periodically? Yes or No

9. The percentage of security professionals that hold a security related certification?

10. Is there a security policy? Yes or No If yes, is it enforced?

11. Do you have security officer? Yes or No

12. Is the security personnel 24/7? Yes or No

13. What is the average maturity level (proficiency of security procedures) of the security personnel? Rate 1 to 10 with 10 being the highest level of maturity level

14.Is there a clear separation of duties among the employees? Yes or No

15. When an employee is terminated, what is delay between the termination and the locking out of the employee's access to system resources? Time in days or fraction of day

16. When an employee is terminated, is the employee escorted out by security? Yes or No

17.The percentage of employees that run applications with non-administrative privileges?

18.1s there secured physical access to the production servers? Yes or No

19. Is there a firewall separating the internet from the web servers? Yes or No If yes, are all ports other than internet ports (HTTP and HTTPS) blocked? Yes or No

20. Is there a firewall separating the web servers from the backend servers? Yes or No If yes, are all ports other than necessary ports blocked? Yes or No

21.Does your organization use a anti-virus software? Yes or No If yes, how often are the anti-virus signatures updated? Interval in days (eg daily $= 1$, weekly $= 7$, etc.)

22. Does your organization have a password policy? Yes or No If yes, are client computers tested to ensure that they meet password policy? Yes or No

23. Does your organization require strong passwords? Yes or No If yes, is this tested? Yes or No

24. Does your company allow the reuse of passwords? Yes or No If yes, what is the period before a password or portion of a password can be reused? Period in days

25. Does your organization utilizes directory services for authentication? Yes or No

26.Does your organization make use of group policy? Yes or No

27. Does your organization use access control lists? Yes or No

28. Does your organization use the encrypted file system? Yes or No If yes, what percentage of the field machines use encrypted file system?

29.Does your organization have an Intrusion Detection System? Yes or No

30. Are there IP restrictions? Yes or No

31. Are there DNS restrictions? Yes or No

32.Does your organization use 2 token authentication (eg, biometric and password, smart card and password)? Yes or No

33. Is encrypted login used? Yes or No

34. Is Digital Rights Management utilized? Yes or No

- **35.** Is Instant Messaging allowed? Yes or No
- 36. Are email attachments that are executables or scripts blocked? Yes or No
- *37.* Are there email quotas? Yes or No
- 38. Does your organization utilize digital certificates? Yes or No
- 39. Are domain accounts changed when an IT pro is terminated? Yes or No
- 40. Does your organization use IPSEC? Yes or No
- 41. Does your organization develop any applications? Yes or No
- a. If yes, are software code reviews required? Yes or No
- b. If yes to 41, is software tested by a quality department? Yes or No
- c. If yes to 41, is there a strict validation of input? Yes or No
- d. If yes to 41, are error messages filtered? Yes or No
- d. If yes to 41, is threat modeling performed? Yes or No
- 42. Are servers routinely backed up? Yes or No If yes, what is period in days?

43. Does your organization pay its security professionals according to industry standard? Yes or No If no, is it above or below? Above or Below

44. Does your organization pay its information technology professionals according to industry standard? Yes or No If no, is it above or below? Above or Below

Appendix B

Customer Responses

 \sim \sim

 \sim

 \mathcal{A}^{max}

T.

 $\mathcal{A}^{\mathcal{A}}$

 ~ 10

Appendix C

Normalized Customer Responses

Appendix D

Averaged Weights and Biases Associated the Predictor Model Built using Generated Data

Layers:

Biases:

Weights:

 $\sim 10^6$

Appendix E

Averaged Weights and Biases Associated the Predictor Model Built using Customer Data

Layers:

Biases:

Weights:

VITA

MONICA DEZULUETA

PRESENTATIONS AND PUBLICATIONS

DeZulueta, M., **Adjouadi,** M., "Using Threat Modeling When **Architecting a** Healthcare System", Proceedings of the 8th World Multi-Conference on Systematics, Cybernetics **and** Informatics (Orlando, Florida, **July** 18-21, 2004).

DeZulueta, M., **HIPAA** Accelerator **for BizTalk** Server **2004** Overview **Webcast,** May 2004

DeZulueta, M., Developing **Solutions** for BizTalk Server **2004** FL and AL, April - July 2004

DeZulueta, M., Managing **and** Deploying **BizTalk** Server **Solutions, July 2004**

DeZulueta, M., Building **a** Custom Pipeline **for** BizTalk Server **2004** Microsoft Global Briefing, Atlanta, **GA,** July 2004

DeZulueta, M., Technical **Webcasts on** Commerce Server **2002 and** Content Management Server **2002**

DeZulueta, M., Orchestrating Web Services, Lockheed Mission Critical Enterprise Symposium **Conference,** Orlando, **FL,** October **2003**

DeZulueta, M., Information Worker Infrastructure, Windows Server **2003** Launch, FL and **AL,** May 2003

CONTRIBUTIONS

Meier, J. D. and etal, "Building Secure Microsoft **ASPNET** Applications". Redmond, Washington: **MS** Press, **2003.**

"XML Data **Islands,** Updategrams, **Stored** Procedures, **and More",** MSDN Magazine, May 2002

CERTIFICATIONS

Certified Information Systems Security Professional

DISCLOSURE

DeZulueta, M., Adjouadi, M., and Ayala M., Methodology **and Tool for** Assessing Risks **in** Information Systems, Oct **2004.**