

# THREAT ANALYSIS USING ARTIFICIAL NEURAL NETWORK

YEE CHAN PHENG

A project report submitted in partial fulfillment of the  
requirements for the award of the degree of  
Masters of Computer Science (Information Security)

Faculty of Computer Science and Information Systems  
Universiti Teknologi Malaysia, Kuala Lumpur

APRIL 2009

## ABSTRACT

The purpose of this study is to explore the use of an Artificial Neural Network threat analysis tools for analyzing threats in healthcare system. The research method used a feed forward neural network which consisted of 50 input variables and one output. The datasets used in neural network are provided by previous research conducted in one of the Government Supported Hospital. The neural network is trained with the datasets and performed prediction. In order to test the accuracy of ANN prediction, internal validation will be made. Six experiments conducted and the mean square error used as a scale to measure the accuracy of prediction. First three experiments which with 50 input variables and one output used 80%, 60% and 40% of data for training. While the last three experiments change the number of input variables to 15 and use 80%, 60% and 40% of data for training. The results between the six experiments were compared. It was discovered that when the size of trained data reduced, the MSE value increased. In contrast, while the size of trained data increased, the MSE value decreased. Lower MSE value means better prediction. Overall, the accuracy of prediction for artificial neural network is high. The changes in the number of input variables will not affect the power of ANN prediction. However, quantity of data is one of the important factors that affect ANN prediction result. With larger data size, the ANN prediction is more accurate.

## ABSTRAK

Kajian ini dilakukan bertujuan membina peralatan ancaman analisis dengan Jaringan Neural untuk sistem kesihatan. Kaedah penyelidikan ini menggunakan jaringan neural feed forward yang mengandungi 50 pembolehubah input dan 1 output. Data yang digunakan dalam jaringan neural ini disediakan oleh satu kajian yang dikendalikan dalam salah satu Hospital Awam. Jaringan neural dilatih dengan data dan membuat ramalan. Untuk menguji ketepatan ramalan Jaringan Neural, pengesahan dalaman akan dibuat. 6 eksperimen pengesahan dalaman akan dibuat dan min ralat kuasa dua yang diggunakan untuk mengukur ketepatan ramalan akan dikira. Eksperimen satu hingga tiga yang mempunyai 50 pembolehubah input dan 1 output masing-masing melatih dengan 80%, 60% dan 40% data. Manakala tiga eksperimen yang terakhir mengubah jumlah input pembolehubah ke 15 dan masing-masing melatih dengan 80%, 60% dan 40% data. Keputusan antara 6 eksperimen telah dibandingkan. Ia didapati bahawa bila saiz data terlatih turun, nilai min ralat kuasa dua meningkat. Manakala saiz data terlatih meningkat, nilai min ralat kuasa dua menurun. Nilai min ralat kuasa dua yang lebih rendah bermakna ramalan lebih tepat. Ketepatan ramalan bagi jaringan neural adalah tinggi. Perubahan dalam jumlah input pembolehubah tidak menjejaskan kuasa ramalan jaringan neural. Kuantiti data adalah satu faktor penting yang menjejaskan ramalan jaringan neural. Saiz data yang lebih besar, ramalan jaringan neural adalah lebih tepat.

**TABLE OF CONTENTS**

<b>CHAPTER</b>	<b>TITLE</b>	<b>PAGE</b>
	<b>DECLARATION</b>	ii
	<b>DEDICATION</b>	iii
	<b>ACKNOWLEDGEMENT</b>	iv
	<b>ABSTRACT</b>	v
	<b>ABSTRAK</b>	vi
	<b>TABLE OF CONTENTS</b>	vii
	<b>LIST OF TABLES</b>	xi
	<b>LIST OF FIGURES</b>	xii
	<b>LIST OF APPENDICES</b>	xiv
<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
	1.1 Background of Problem	2
	1.2 Problem Statement	2
	1.3 Project Aim	3
	1.4 Project Objectives	3
	1.5 Project Scope	4
	1.6 Summary	5

<b>2</b>	<b>LIRERATURE REVIEW</b>	<b>6</b>
2.1	Artificial Neural Network	6
2.1.1	The application of neural network	8
2.1.2	Neural Network Models	9
2.1.3	Neural Network Learning	10
2.1.4	Learning Paradigms	11
2.1.5	Learning algorithms	14
2.1.6	Types of artificial neural network	14
2.2	Feed forward neural network	15
2.2.1	Single-layer perceptron	16
2.2.2	Multi-layer perceptron	18
2.3	Existing Neural Network System	19
2.3.1	Simulators	20
2.3.1.1	Research simulators	21
2.3.1.2	Analysis and Prediction	25
	Simulators	
2.3.2	Comparison of the Artificial Neural Network System	29
2.4	Related Research Discussion	29
2.4.1	ANN Improve the Accuracy of Cancer Survival Prediction	29
2.4.2	Application of ANN-Based Survival Analysis on Cancer	30
2.5	Summary	31

<b>3</b>	<b>RESEARCH METHODOLOGY</b>	<b>33</b>
3.1	Operational Framework	33
3.1.1	Literature Review	35
3.1.2	Data Collection	35
3.1.3	Requirement Specification	35
3.1.4	Design the Mathematical Model	36
3.1.5	Testing	36
3.1.6	Prototype Design and Development	37
3.1.7	Validation	40
3.2	Summary	41
<b>4</b>	<b>IMPLEMENTATION</b>	<b>42</b>
4.1	Artificial Neural Network Structure	42
4.1.1	Input	43
4.1.2	Output	43
4.1.3	Artificial Neural Network Example	44
4.2	Program Interface and Functionality	45
4.2.1	Data Management	45
4.2.2	Analysis and Prediction	46
4.2.3	Validation	47
4.3	Summary	49
<b>5</b>	<b>RESULT</b>	<b>50</b>
5.1	Result	50
5.2	Summary	57

<b>6</b>	<b>DISCUSSION</b>	<b>59</b>
	6.1 Discussion	59
	6.2 Summary	60
<b>7</b>	<b>CONCLUSION</b>	<b>61</b>
	7.1 Conclusion	61
	7.2 Future Work	61
	<b>REFERENCES</b>	<b>63</b>
	APPENDICES A – G	66 - 73

## **CHAPTER 1**

### **INTRODUCTION**

Nowadays the information security threats grow larger and more healthcare systems are affected. They spend high costs on protecting the systems and system recovery. Although many security measures is adopting, it is impossible to prevent all threats. Therefore, an easy and quick IT system should be installed to help the healthcare professional to analyze threat factor in order to predict the survival of the system.

The Healthcare system needs an IT solution that could analyze the history of threats affecting data. One of the solution introduced in this research applied an artificial Neural Network. The method that is going to be developed attempted to introduce potential threats and this may help in planning for better system in healthcare.

This Introduction chapter will describe the problems faced in status quo by the healthcare and how the proposed system would help minimize the problems. This chapter begins by Background of Problems (section 1.1). Then it is followed by section 1.2 project problem statement in which the underlying problems or inconveniences within the current environment for healthcare is explained. Apart from that, Section 1.3 is project objectives, section 1.4 is project aims and Section 1.5 is project scope. Finally, a summary for this chapter will be in section 1.6.



## **1.1 Background of Problems**

Nowadays, due to the growing Internet connectivity of computers and networks and the corresponding user dependence on network-enabled services (such as email and Web-based transaction) have increased the number and sophistication of attack methods on healthcare systems. This trend puts healthcare systems at greater risk. Threats that always attack healthcare system including virus, spamming, adware, malware, hackers and so on [24]. Sometime, the threats even broke down the healthcare systems. As a result, healthcare is facing unpredictable identities and monetary lost.

## **1.2 Problem Statement**

There are a few underlying problems or inconveniences within the current environment for healthcare. They include:

### **1) Lack of user-friendly and economic Artificial Neural Network threat analysis tool for healthcare system**

Giant organization that applies sophisticated healthcare system spent more money to perform risk analysis. Whereby small clinics don't not even have a chance to purchase such tools to do risk analysis. In addition some of the tools do not have user friendly interface. The processes and interface is complex enough. It is inconvenient for the users to take a long time to learn and use the tools. Furthermore, the tools would slow down due to addition of more unusable features, which is inevitable as the history data size get larger.

## 2) **Lack of validation function that support line graph**

Although most of the Artificial Neural Network tools in the market provided validation function, line graph function is not included inside the validation function. In order to view the differences between the real result and the predicted result on line graph, the user needs to do it manually in external graph program.

## 3) **Limited number of input variables can be included in the analysis**

Although there are some Artificial Neural Network threat analysis tools in the market, the number of variables can be included in analysis is limited. As a result, if the number of variables is larger than the limited number of variables in the tool, the tool will not be suitable any more.

### 1.3 **Project Aim**

The aims of the research are to develop a threats analysis tools using ANN

### 1.4 **Objective**

The objectives of the research are:

#### 1) **To develop a user friendly and economic threat analysis tools for healthcare system**

Basically, the tool will have a simple and easy to learn interface. These will make the users convenient enough to use it. The tool is tailor-made for the healthcare, all the inadequate features and function are not included. So the tool may run faster

and efficiently. Also, the tool is economic enough. Small clinics will be able to afford the cost and have a chance expose to such threat analysis software.

## **2) To do validation with line graph**

The validation function is supported with line graph export function. After finishing the validation, the tool may enable the user to export the result to Microsoft excel 2007 and create line graph automatically.

## **3) To develop an Artificial Neural Network threat analysis tool with 50 numbers of input variables for healthcare**

The tool is tailor-made for the HealthCare. So the tool is able to analyze up to 50 variables data.

### **1.5 Project Scope**

- **Use Artificial Neural Network to analyze threats and perform system survival prediction based on previous data and results.**

This tool is tailor-made for the healthcare. By inputting the certain situation data, and with the help of previous data and result (previous threat affected history and the system survival result), the tool may predict whether the system going to survive.

## **1.6 Summary**

This research is going to implement neural network in analyzing threats for healthcare system. A tailor-made Artificial Neural Network threat analysis tool for healthcare will be developed

Chapter one reported the background of problem, introduction, problem of statement, objective, aim and project scope. The following chapter will describe more details on introducing Artificial Neural Network and discuss some previous research related to this topic.

## REFERENCES

1. Chih-Lin Chia, W. Nick Street b, William H. Wolbergc. Application of Artificial Neural Network-Based Survival Analysis on Two Breast Cancer Datasets, 2007.
2. R. Bendaña, G. Astray. COMPUTER SYSTEM FOR RURAL LAND VALUATION BASED ON NEURAL NETWORKS TECHNOLOGY , 2007 , ISSN: 1573-4377
3. Steffen Nissen. Implementation of a Fast Artificial Neural Network Library (fann) , October 31, 2003
4. Juan Ramon Rabunal, Julian Dorrado , Artificial Neural Networks in Real-life Applications , Published by Idea Group Inc (IGI) , pages 168-169 ; 2006
5. T.J.A. Jansen , Player Adaptive Cooperative Artificial Intelligence for RTS Games , June 13, 2007
6. Shiguo Lian , Image Authentication Based on Neural Networks ,31 Jul 2007
7. M. Govindarajan, and R. M. Chandrasekaran. Classifier Based Text Mining for Neural Network , 21 MAY 2007
8. Jan Drchal .Evolution of Recurrent Neural Networks. Diploma. CZECH TECHNICAL UNIVERSITY in PRAGUE ; January 2006
9. ANDREW G. BARTO and THOMAS G. DIETTERICH. Reinforcement Learning and it Relationship to Supervised Learning. 2006

10. Peralta, Richard C. Intelligent space tube optimizer. United States Patent Application 20070168328. 2007
11. Prof. Subrata Kumar Mitra. IMPROVING ACCURACY OF OPTION PRICE ESTIMATION USING ARTIFICIAL NEURAL NETWORKS. 2005
12. U. Muller. Artificial intelligence—applications in high energy and nuclear physics. 2003
13. Harry B. Burke, Philip H. Goodman and David B. Rosen. Artificial Neural Networks Improve the Accuracy of Cancer Survival Prediction. 1997
14. Gurney, K. An Introduction of Neural Networks. 1997
15. De Laurentiis, M. and Ravdin, P. M. .A technique for using neural network analysis to perform survival analysis of censored data. 1990
16. Haykin, S. Neural Networks: A Comprehensive Foundation. 1999
17. Mandic, D. & Chambers, J. Recurrent Neural Networks for Prediction: Architectures, Learning algorithms and Stability. Wiley. 2001
18. Gurney, K. An Introduction to Neural Networks London .1997.
19. Maass, W., & Markram, H.. 2002.
20. Wilkes, A.L. & Wade, N.J. . Bain on Neural Networks". 1997.
21. Ness, Erik. On the use of artificial neural networks in species taxonomy. 2005
22. Bhadeshia H. K. D. H. "Neural Networks in Materials Science".1999.
23. Bhagat, P.M. Pattern Recognition in Industry, Elsevier. 2005
24. Stanley Azen . Comparison of Predictive Accuracy of Neural Network Methods and Cox Regression for Censored Survival Data. 2000

25. Yee, P. V. "Regularized radial basis function networks: Theory and applications to probability estimation, classification, and time series prediction," Ph.D. Thesis, McMaster University, Hamilton, Ontario. 1998
26. Wan, E. A. Time series prediction by using a connectionist network with internal delay lines. 1994
27. Sandberg, I.W., and Xu, L. "Uniform approximation and gamma networks," Neural Networks, vol. 10, pp. 781–784. 1997
28. Anderson, J. A. Introduction to Neural Networks. 1995
29. Broomhead, D. Sand D. Lowe. Multivariable functional interpolation and adaptive networks Complex Systems. 1998
30. Poggio, T., and F. Girosi. Networks for approximation and learning .1990