Neuro-Fuzzy based Software Risk Estimation Tool

Pooja Rani[1] and Dalwinder Singh Salaria[2]

[1] Lovely Professional University

Received: 11 December 2012 Accepted: 31 December 2012 Published: 15 January 2013

Abstract

To develop the secure software is one of the major concerns in the software industry. To make the easier task of finding and fixing the security flaws, software developers should integrate the security at all stages of Software Development Life Cycle (SDLC).In this paper, based on Neuro- Fuzzy approach software Risk Prediction tool is created. Firstly Fuzzy Inference system is created and then Neural Network based three different training algorithms: BR (Bayesian Regulation), BP (Back propagation) and LM (Levenberg-Marquardt) are used to train the neural network. From the results it is conclude that for the Software Risk Estimation, BR (Bayesian Regulation) performs better and also achieves the greater accuracy than other algorithms.

Index terms— software security, software threat, neural network, fuzzy logic, neuro-fuzzy.

1 Introduction

oftware systems are being used in every area to perform the different kind of activities all over the world. Due to the rapid growth of internet, technology advancement and the extensively usage of software systems results in security threats that are increasing day by day. So security becomes important concern to be considered. Threat can be any undesired event that is having potential to harm the system. Software Threat Modeling is an approach that deals with the identification, mitigation and prioritization of attacks that have to address. To predict the model for software threats, there are number of techniques like: Statistical techniques, Neural Network, Genetic Algorithm, Support Vector Machine, Fuzzy Logic and hybrid approaches: Neural Network with Genetic Algorithm, Neural Network with Support Vector Machine and Neuro-Fuzzy are being used .As it is fact that each technique has their own pros and cons. It cannot be say that one technique can overcome the limitations of all other techniques. But from the past research work, its find that the hybrid approaches provide more level of accuracy than the individual approaches.

Author ? : Student, Department of CSE Lovely Professional University Phagwara, Punjab -144411. E-mail : erpoojapuri88@gmail.com Author ? : Assistant Professor, CSE Dept Lovely Professional University Phagwara, Punjab -144411. E-mail : ds_salaria@yahoo.com II.

2 Review of Literature

For Software threat prediction, various statistical approaches as well as advanced approaches are introduced in different areas where Software systems are being used. For Cyber Threat, Cyber threat trend analysis model is proposed using Hidden Markov Model (HMM), to forecast the Cyber threat trend. HMM is a tool in which hidden state is determined .After comparison with existing techniques, the proposed model provides accurate results [1]. MERIT workshop and training programs are conducted for effective training about insider threat awareness. Insider threats are those undesired events that are performed by the legitimate users [2]. Threat Analysis and Modeling (TAM) tool is used to identify the threats and evaluate the risks. This process is useful in business applications [3]. To identify the most critical large system threats, Cyber Threat Tree is implemented as directed graph known as Multiple Valued Decision Diagram (MDD). Cyber Threat Markup Language (Cyma)

is used for cyber threat tree representation. Multiple Valued Logic function is used to represent the threat states and their interdepend -ence [4].

In the area of Software Security, to identify the security vulnerabilities in software systems and to show the sequential events that occur during an attack, Regular expression based attack patterns are created. Identification of vulnerabilities is done via matching sequence of components that trigger an event during an attack [5]. Threat Mitigation, Monitoring and Management Plan (TMMMP) approach is discussed to identify the threats, to monitor the remedial measures and to deal with management plans in case of failure of remedial measures. It uses Defense In Depth (DID) strategy for threat mitigation and risk management associated with threats [6]. To identify the security flaws at early stages of software development life cycle, Extended Model Driven Architecture (MDA) approach is introduced with quantitative security assessment model. It will provide the feedback at every stage of software development life cycle [7]. To prioritize the identified threats, Common Vulnerability Scoring system (CVSS) based Risk ranking Tool is used. This tool converts Yes/ No values into numerical values and then calculates the risk score using CVSS. It helps to software developer by answering the impact and exploitability of threats [8]. To overcome the limitation i.e. identification of effects by new security threats and to developing proper countermeasures, two kind of security patterns are introduced i.e. Software Requirement Patterns (SRPs) and Software Design Patterns (SDPs). To identify the threats Software Requirement Patterns (SRPs) are used. Software Design Patterns (SDPs) are used for the identification of remedial measures against identified threats [9].

In the Networked organizations, to enhance the security by prioritizing threats and vulnerabilities, a new methodology is proposed that integrates threat modeling with formal threat analysis. This method is divided into three phases: Threat modeling, asset mapping and mitigation plan that enable the system to identify, quantify the threats and vulnerabilities [10]. For identification of threats in networked organizations, a new approach is introduced that provides reliability statistics to defense analyst to identify the top node in the network. It is useful to identify the top threats in networked organizations [11]. Now a day's modern technique Neural Network is emerged. It is also used to model the software threats. For an intrusion detection system, user behavior modeling approach is introduced that use the neural algorithm and provides better results than existing results [12]. With the use of hybrid approach i.e. Neural network and support vector machine, Intrusion detection system is constructed. It is observed that the performance of this hybrid approach is superior and deliver accurate results [13]. As we know new intrusions are introduced day by day, so there is need to update the new rules to intrusion detection systems. To meet this requirement, a new intrusion detection system is presented with Genetic algorithm approach [14].

To model the real world risk scenarios, risk analysis modeling is introduced that uses fuzzy logic technique. Fuzzy logic model the vagueness in natural way. Thus it provides the accurate recommendations [15]. For electronic commerce development, web based Fuzzy Decision Support System (FDSS) is introduced. This will help to identify electronic ecommerce risk factors [16]. With the use of fuzzy logic secure software system (SSS) approach is introduced. It will help to avert the failed state of the system [17].

For the development of marketing strategy, hybrid intelligent system is developed with the combined approach of Neural Network, Fuzzy Logic and expert system. For the settlement of marketing strategy, this hybrid system is useful to produce intelligent advice [18].

Neural fuzzy scheme is proposed for the development of Direction of arrival (DOA) estimation algorithm by Self-constructing Neural fuzzy Inference Network (SONFIN). The performance of this newly developed algorithm is superior than RBFN [19]. To calibrate the conversion ratios for backfiring technique, accuracy is achieved for software size estimation [20]. To make the decision about Distributed Intrusion Prediction and Prevention system (DIPPS) , a model named Hierarchical Neuro-Fuzzy Online Risk Assessment(HiNFRA) using Neuro-fuzzy approach is introduced. This model by using Neuro-fuzzy approach results in more robustness and better performance [21].

3 III. Neuro-Fuzzy Risk Prediction Model

For the prediction of risk, Neuro-Fuzzy approach is used in this paper. Because the combination of Neural Network and Fuzzy Logic results in such hybrid intelligent system that is having learning ability to optimize its parameters with the use of neural network and to represent the knowledge in an interpretable manner, with the use of Fuzzy System. The hybrid Neuro-Fuzzy technique is well suitable to those areas or applications, where the interpretation and interaction of user is required. Neuro-Fuzzy approach provides more accurate results than other existing hybrid techniques.

4 a) Fuzzy Inference System

Fuzzy Inference System is based on the concept of Fuzzy set, If Then Rules and Defuzzification. In this paper, MATLAB Fuzzy toolbox that is Graphical User Interface tool used to build the Fuzzy Inference System. To determine how Neuro-Fuzzy approach can be applied to evaluate the Software risk, some of the software factors that affect the security vulnerability are considered. These risk factors are abstracted from [22] [23] [24]. Regarding these input attributes, Corresponding security vulnerability output in the form of Low, Medium, High, Very Low and Very High are obtained from Software industry experts in from of surveys. The total 17

input risk attributes includes the following. C calibrated model is generated by using neuro-fuzzy approach. From this model, it is concluded that higher 17. Wrong Functions, Properties and UI(User Interface) Development.

i. Fuzzification Fuzzification is the process to describe the input parameters through linguistic variables with meaning like 'Low','High','Medium','Very Low' and 'Very High'. Fuzzy sets are representation of input parameters. These sets are represented by Membership Functions. Input parameters are represents by Zmf (Z-shaped built-in membership function). Similarly, Output parameters are represented by Gauss (Gaussian curve built-in membership function).

5 ii. Rule Evaluation

The total 137 if-then rules are generated after the creation of input output fuzzy sets and Membership functions. In the rules 'T' means "True" and representing value 1 and 'F' means "False" and representing value 0. The rules created in rule base of Fuzzy Inference System (FIS) are represented in the following format:

If(Fault/Changing Requirements is 'T') and (Lack of user Co-operation is 'F') and (Poor Project Planning is 'T') and (Poor Project management and Resource Estimation is 'F') and (Undefined Project Milestones is 'T') and (Personnel Shortfalls is 'F') and (Insufficiently Trained Team Members is 'T') and (Lack of Specialization is 'F') and (Inexperienced Project Manager is 'T') and (Schedule variation is 'F') and (Budget variation is 'T') and (Deviation From Software Requirements is 'F') and (Shortfalls in Externally Furnished Components is 'T') and(Shortfalls in Externally Performed Tasks is 'F') and (Limitations on Real Time Performance Activities or Tasks is 'T') and (Computer Science Difficulties is 'F') and (Wrong Functions, Properties and UI Development is 'T'). Back propagation (triangdx) is a learning algorithm means it learns from many inputs for desired output. It is very simple. It does not require any specialization. But the Limitation of this algorithm is that its having low prediction capabilities. Due to low prediction capabilities, it does not provide accurate results.

Bayesian Regulation (Trainbr) is advanced method. This algorithm is more suitable for those

6 Results and Comparison

Neural Network is trained with three different algorithms: BR, BPA & LM and outputs are obtained. From the table 1. The comparison among three different algorithms can be seen. In the table 17 inputs parameters are used and corresponding Security vulnerability output is computed for BR, BP and LM algorithms. The comparison shows that BR provides the better results than BP and LM algorithms. The results provides by BR are accurate where as BP and LM are over fitting the values for the same dataset.

The table1: Summarizes the results achieved by these three different algorithms over the same dataset. Some short terms are used in the table for input parameters are as follows.

7 Conlcusion

Software Risk Prediction is one of the most important tasks for the development of secure and reliable system. It should be preferred that during the early stages of software development life cycle to find and fix the security flaws. Neuro-fuzzy approach based risk prediction tool is developed using MATLAB. After creation of Fuzzy Inference System, Neural Network is trained with three different algorithms using 'trianbr', 'traingdx' and 'trainlm'. From the results it is concluded that BR algorithm performs better than other algorithms. With the use of BR algorithm better accuracy level is achieved then other algorithms. [1]

Figure 1: C

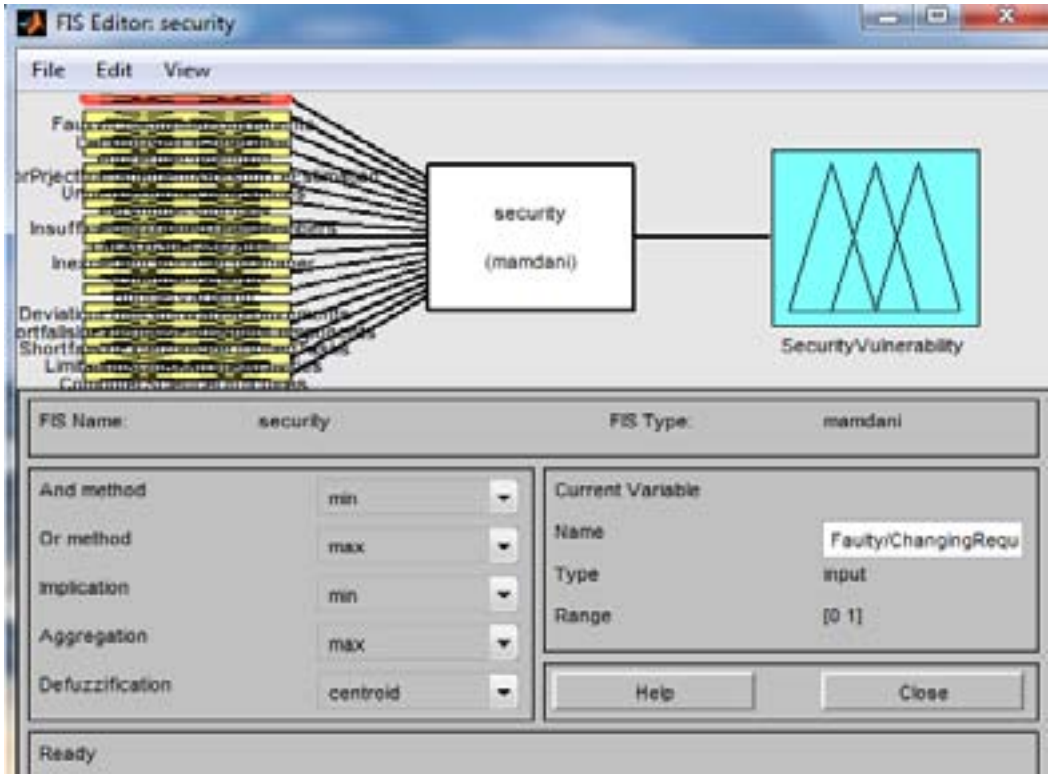


Figure 2:

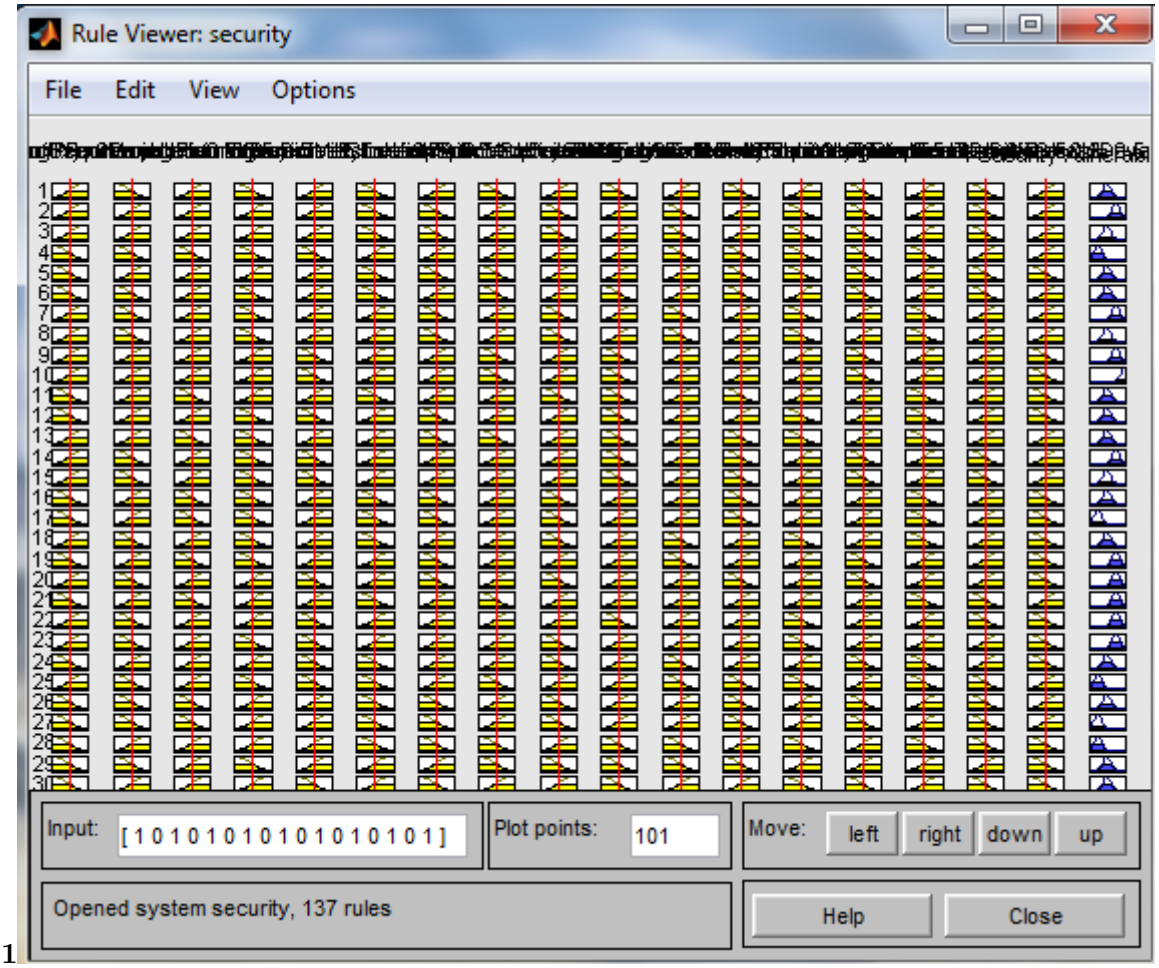
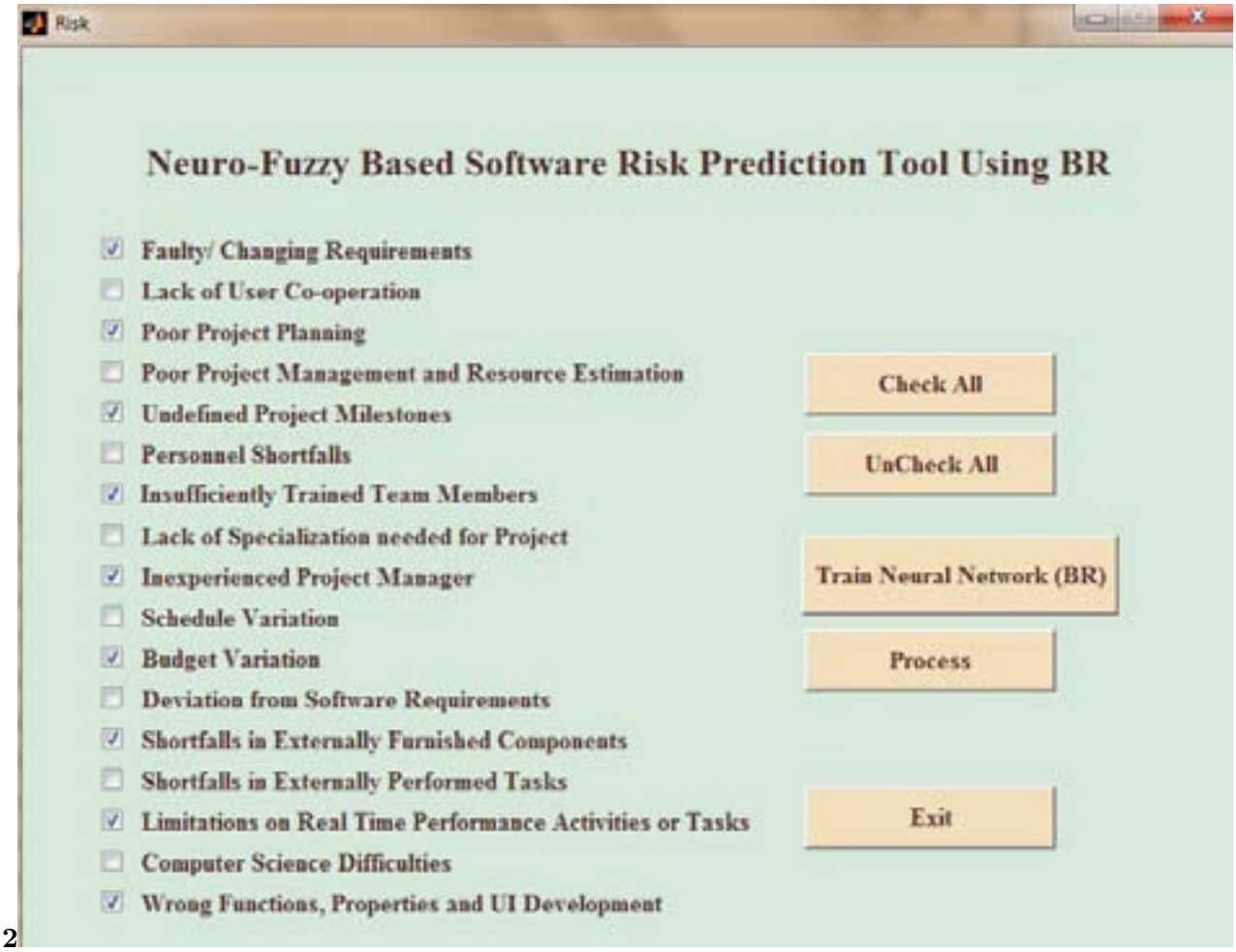


Figure 3: Figure 1 :



2

Figure 4: Figure 2 :

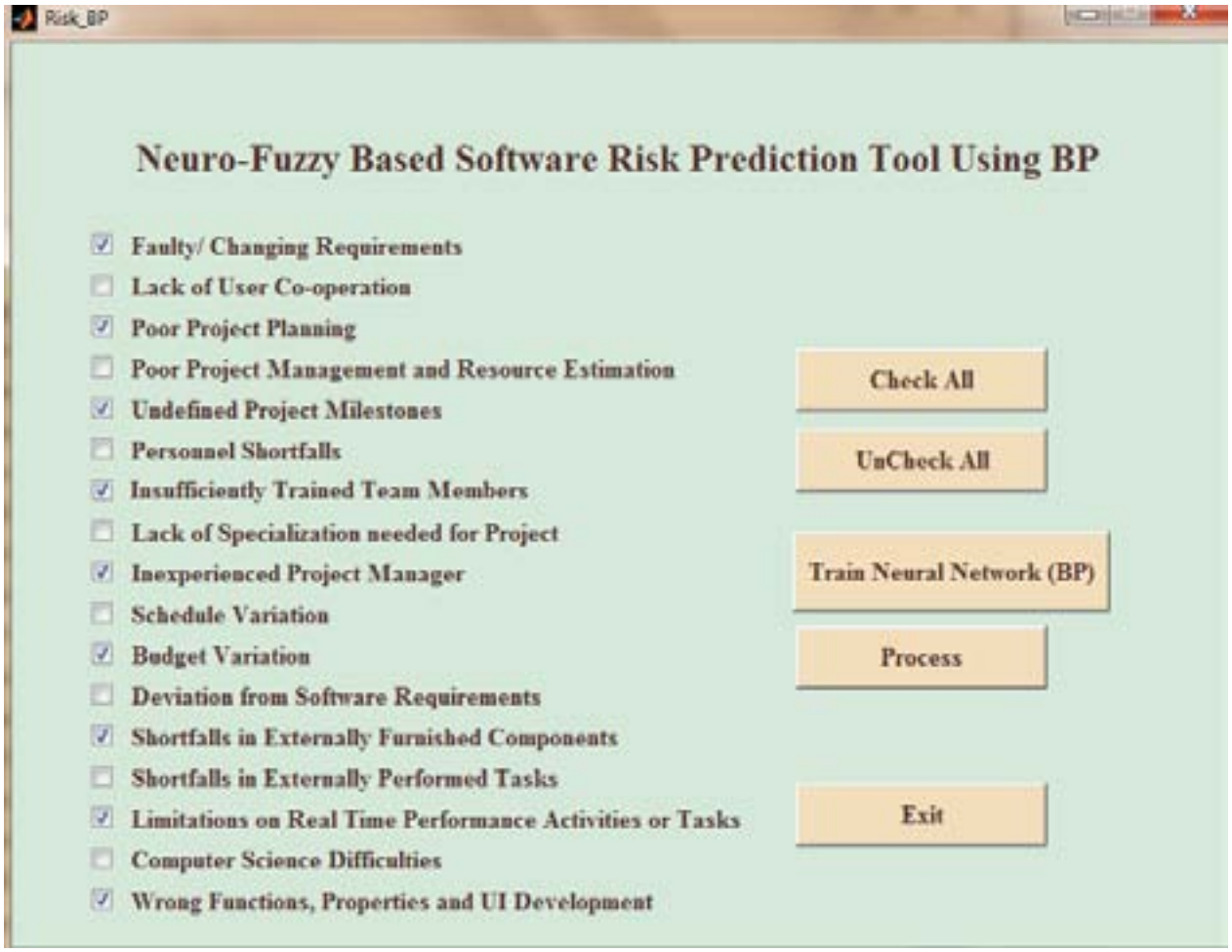
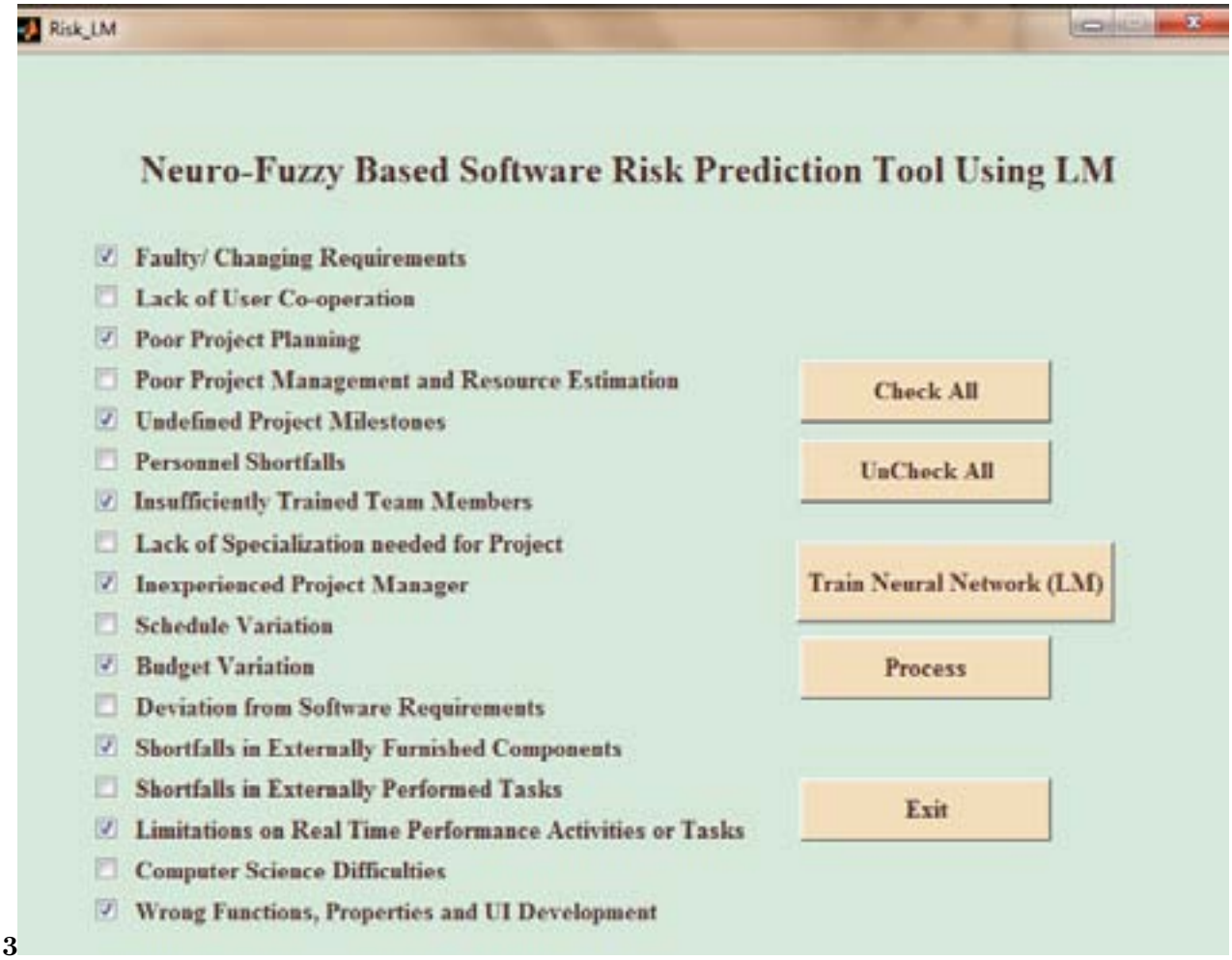

Risk_BP
Neuro-Fuzzy Based Software Risk Prediction Tool Using BP
Faulty/ Changing Requirements
Lack of User Co-operation
Poor Project Planning
Poor Project Management and Resource Estimation
Undefined Project Milestones
Personnel Shortfalls
Insufficiently Trained Team Members
Lack of Specialization needed for Project
Inexperienced Project Manager
Schedule Variation
Budget Variation
Deviation from Software Requirements
Shortfalls in Externally Furnished Components
Shortfalls in Externally Performed Tasks
Limitations on Real Time Performance Activities or Tasks
Computer Science Difficulties
Wrong Functions, Properties and UI Development
Check All
UnCheck All
Train Neural Network (BP)
Process
Exit

Figure 5: C



3

Figure 6: Figure 3 :

FR	LUC	PPP	PPMRE	UPM	PS	ITM	LOS	IPM	SV	BV	DFSR	SEFC	SEPT	LRTPA	CSD	WFUID	SV BR	SV BP	SV LM
T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	94.49%	122.00%	107.50%
F	T	T	T	T	F	T	T	T	T	T	T	T	T	T	T	T	92.90%	126.42%	116.44%
T	F	T	T	F	T	T	T	T	T	T	T	T	T	T	T	F	89.29%	62.04%	79.91%
T	T	T	T	T	T	F	T	F	T	T	F	T	T	F	T	T	84.77%	145.45%	96.73%
T	T	F	T	T	T	T	T	T	F	T	T	F	T	F	F	T	79.81%	-20.94%	73.39%
F	T	T	F	T	F	T	F	T	T	F	T	T	F	T	T	T	72.06%	90.22%	73.43%
T	F	T	F	F	T	F	F	T	T	F	T	T	T	T	T	F	68.96%	36.64%	38.46%
F	T	T	T	F	T	F	T	F	T	F	T	T	T	F	T	F	65.51%	75.38%	47.59%
T	F	F	T	F	T	F	T	T	F	T	F	T	F	T	F	T	59.30%	34.81%	61.16%
T	T	F	F	T	F	F	T	T	F	T	F	F	T	F	T	T	53.03%	47.23%	57.09%
F	F	F	T	F	T	F	F	F	T	T	F	F	T	T	T	T	47.27%	100.01%	68.02%
T	F	T	F	T	F	T	F	F	T	F	T	T	F	T	F	F	43.63%	3.58%	23.40%
T	T	T	F	F	F	T	F	F	T	F	T	T	F	F	T	F	37.41%	54.94%	29.29%
F	T	F	F	F	F	F	T	T	F	F	F	T	T	T	F	T	33.26%	-14.34%	18.33%
F	F	F	T	F	T	T	F	T	F	T	T	F	T	F	F	F	28.80%	19.98%	27.09%
F	F	F	F	F	T	T	F	T	F	T	T	F	T	F	F	F	21.32%	20.63%	31.53%
T	F	F	F	T	T	T	F	F	F	T	F	F	F	F	F	F	17.35%	26.50%	23.10%
F	F	F	F	F	F	F	F	F	T	F	T	T	F	T	F	F	12.64%	35.06%	12.13%
F	F	F	F	F	F	F	F	F	F	F	F	T	F	T	F	F	5.96%	21.22%	2.85%
F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	2.23%	10.60%	4.48%

45

Figure 7: Figure 4 :Figure 5 :

[Debar et al. (1992)] 'A neural network component for an intrusion detection system'. H Debar , M Becker , D Siboni . *Research in Security and Privacy, 1992. Proceedings. 1992 IEEE Computer Society Symposium on*, 1992. May. IEEE. p. .

[Gong et al. (2005)] 'A software implementation of a genetic algorithm based approach to network intrusion detection'. R H Gong , M Zulkernine , P Abolmaesumi . *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, 2005 and First ACIS International Workshop on Self-Assembling Wireless Networks*, 2005. May.

[Gandotra et al. (2010)] 'A step towards secure software system using fuzzy logic'. V Gandotra , A Singhal , P Bedi . *Computer Engineering and Technology (ICCET), 2010 2nd International Conference on*, 2010. April. IEEE. 1 p. .

[Stango et al. (2009)] 'A threat analysis methodology for security evaluation and enhancement planning'. A Stango , N R Prasad , D M Kyriazanos . *Emerging Security Information, Systems and Technologies, 2009. SECURWARE'09. Third International Conference on*, 2009. June. IEEE. p. .

[Hu et al. (2010)] 'A unified intelligent model for software project risk analysis and planning'. Y Hu , X Zhang , X Sun , J Zhang , J Du , J Zhao . *Information Management, Innovation Management and Industrial Engineering (ICIII), 2010 International Conference on*, 2010. November. IEEE. 4 p. .

[ARTCom'09. International Conference on] *ARTCom'09. International Conference on*, IEEE. p. .

[Wong et al. ()] 'Calibrating function point backfiring conversion ratios using neuro-fuzzy technique'. J Wong , D Ho , L F Capretz . *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 2008. 16 (06) p. .

[Greitzer et al. ()] 'Combating the insider cyber threat'. Frank L Greitzer , P Andrew , Dawn M Moore , Dee H Cappelli , Lynn A Andrews , Thomas D Carroll , Hull . *Security & Privacy, IEEE* 2008. 6 (1) p. .

[Ongsakorn et al. (2010)] 'Cyber threat trees for large system threat cataloging and analysis'. P Ongsakorn , K Turney , M Thornton , S Nair , S Szygenda , T Manikas . *Systems Conference*, 2010. April. 2010. IEEE. p. . (4th Annual IEEE)

[Dhillon ()] 'Developer-Driven Threat Modeling: Lessons Learned in the Trenches'. Danny Dhillon . *Security & Privacy, IEEE* 2011. 9 (4) p. .

[Shieh and Lin ()] 'Direction of arrival estimation based on phase differences using neural fuzzy network. Antennas and Propagation'. C S Shieh , C T Lin . *IEEE Transactions on* 2000. 48 (7) p. .

[Okubo et al. (2011)] 'Effective security impact analysis with patterns for software enhancement'. T Okubo , H Kaiya , N Yoshioka . *Availability, Reliability and Security (ARES)*, 2011. August. 2011. IEEE. p. . (Sixth International Conference on)

[Tang and Shen (2009)] 'Extending Model Driven Architecture with Software Security Assessment'. X Tang , B Shen . *Secure Software Integration and Reliability Improvement*, 2009. July. 2009. 2009. IEEE. p. . (Third IEEE International Conference on)

[Ngai and Wat ()] 'Fuzzy decision support system for risk analysis in e-commerce development'. E W T Ngai , F K T Wat . *Decision support systems* 2005. 40 (2) p. .

[Bragina and Tabunshchyk (2011)] 'Fuzzy model for the software projects design risk analysis'. T Bragina , G Tabunshchyk . *CAD Systems in Microelectronics (CADSM), 2011 11th International Conference. The Experience of Designing and Application of*, 2011. February. IEEE. p. .

[Haslum et al. (2008)] 'Hinfra: Hierarchical neuro-fuzzy learning for online risk assessment'. K Haslum , A Abraham , S Knapskog . *Modeling & Simulation, 2008. AICMS 08. Second Asia International Conference on*, 2008. May. IEEE. p. .

[Haslum et al. (2008)] 'Hinfra: Hierarchical neuro-fuzzy learning for online risk assessment'. K Haslum , A Abraham , S Knapskog . *Modeling & Simulation, 2008. AICMS 08. Second Asia International Conference on*, 2008. May. ieee. p. .

[Frantz and Carley (2009)] 'Information assurances and threat identification in networked organizations'. T L Frantz , K M Carley . *Computational Intelligence for Security and Defense Applications*, 2009. July. 2009. IEEE. p. . (CISDA 2009. IEEE Symposium on)

[Mukkamala et al. ()] 'Intrusion detection using neural networks and support vector machines'. S Mukkamala , G Janoski , A Sung . *Neural Networks, 2002. IJCNN'02. Proceedings of the 2002 International Joint Conference on*, 2002. IEEE. 2 p. .

[Lee et al. (2007)] T Lee , S O Jung , H P In , H J Lee . *Cyber Threat Trend Analysis Model Using HMM. In Information Assurance and Security*, 2007. August. 2007. 2007. IEEE. p. . (Third International Symposium on)

[Gegick and Williams ()] 'Matching attack patterns to security vulnerabilities in softwareintensive system designs'. Michael Gegick , Laurie Williams . *ACM SIGSOFT Software Engineering Notes*, 2005. ACM. 30 p. .

197 [Snpd/Sawn ()] Snpd/Sawn . *Sixth International Conference on*, 2005. IEEE. p. .

198 [Boehm ()] *Software risk management: principles and practices. Software*, B W Boehm . 1991. IEEE. 8 p. .

199 [Li ()] 'The development of a hybrid intelligent system for developing marketing strategy'. S Li . *Decision Support*
200 *Systems* 2000. 27 (4) p. .

201 [Gandotra et al. (2009)] 'Threat mitigation, monitoring and management plan-A new approach in risk manage-
202 ment'. V Gandotra , A Singhal , P Bedi . *Advances in Recent Technologies in Communication and Computing*,
203 2009. October. 2009.

204 [Ingalsbe et al. ()] 'Threat modeling: diving into the deep end'. Jeffrey A Ingalsbe , Louis Kunimatsu , Tim
205 Baeten , Nancy R Mead . *Software, IEEE* 2008. 25 (1) p. .