Fuzzy Preference Approach for Computer Network Attack Detection

Milos Manic, Bogdan Wilamowski

University of Idaho,
College of Engineering, Center Boise,
800 Park Blvd., Suite#200, Boise, ID 83712
{misko, wilam}@ieee.org

Abstract

Irrepressible growth of complex interconnectedness of information systems besides its obvious benefits unfortunately brought up the questions of their vulnerability. Practically universal access to computers has enabled hackers and would-be terrorists to attack information systems and critical infrastructures worldwide. Fuzzy preference relation, based on fuzzy satisfaction function is applied to comparison of attack signatures. Fuzzy signatures (their gamma resolution sets) are combined by fuzzy operators. Therefore, qualitative, fuzzy decision fuzzy decision system is achieved. Different fuzzy set operators used in construction fuzzy satisfaction function, as also as different fuzzy preference relations have been tested. Proposed method provided smoother results that one obtained by traditional methods. Experiments demonstrated that final outcome dependence on correct determination of fuzzy values out of signature attacks, as also as on adequate choice of fuzzy set operator.

1 Introduction

With increased requirements for interconnection in today's enterprise networks comes the increased vulnerability to abuse and misuse of computer systems both from within and outside those networks [1].

An acceptable information protection model of intrusion detection should accomplish three broad goals.

First, it should accommodate mechanisms that protect information assets from compromise, abuse, damage or destruction [2].

Second, it should recognize that compromise is inevitable and that measures must be taken in advance of the compromise to facilitate a means for recovery. While some investigators would disagree with this statement, most "front line" practitioners point to empirical evidence that such is, actually, the case. The obvious theoretical explanation is that no defense is perfect. The practical explanation from front line experience is that the state of the intruder's art, whether our egos permit such an admission or not, is more often than not ahead of the state of the defender's art. To believe otherwise appears naive. Recovery, in this instance, means investigating and recovering, managing and protecting the evidence of the compromise for future use in legal proceedings.

Finally, the model should provide feedback that can speed response to a compromise and generate information that can be used to prevent similar compromises in the future. It is implicit in such a model that recovery is of greater urgency than prosecution [1].

Quantitative nature of statistical and other traditional approaches to comparison of attack system's signatures [2] did not leave much room for survivability and adaptability of such real world problems: "hard" operators deal too "crisp with comparison of variable nature of signatures, different shapes and static nature of idle system signature could incur false alarms, etc. [3,4,5,6].

This paper proposes fuzzy preference approach for capturing the degree of deviation of attacked system signature from the idle system one. This approach gives the freedom of depicting signatures by comprising gradual values (including zero) forming arbitrary, fuzzy sets, eliminating discretization errors, associated with classical approaches.

Malicious act research is usually represented by 3R's. **Resistance** understands hardening the system against hacker attacks. **Recognition** aims in intrusion detection and **Recovery** deals with ways of surviving malicious acts. Survivability is considered to be a combination of recognition and recovery steps (where recognition can be omitted in case of solely fault masking objective) [7].

Attack signatures could be acquired or recorded in many ways. Commercial intrusion detection software suites come with a number of them, with the ability for customers to add or modify their own signatures (*Internet Security Systems*). They come verified, QA tested and digitally signed for

0-7803-7044-9/01/$10.00 ©2001 IEEE

authentication [8]. In this paper, authors concentrate on profile based signatures based on Markov models. They are represented by frequency spectrum of the functionalities in system. Deviation of the signature of attacked, monitored system and "safe" one indicates possible intrusion.

Problems that arise while comparing such signatures are exactly the problems this paper tends to solve. Those are difficulties with differentiating normal signatures from system under attack one and deciding upon a threshold of differentiation between these two classes. This paper aims to propose one possible approach to a process of differentiation between signatures of normal and attacked functioning of a monitored system, and is inspired by the concept proposed in [9].

Therefore, paper aims to encompass as well as computer/network survivability, as fault-tolerant systems, that in spite of their probabilistic, malicious attack features have the same goal.

Fixed, crisp thresholds possibly result in either false alarms or low sensitivity to actual ones. Adaptive thresholds, on the other hand, could enable slow changes in system and therefore unnoticed intrusion. Quantitative numerical methods applied traditionally could not offer satisfying results. Soft computing techniques, namely fuzzy logic, offer more qualitative depiction of data by its inherent linguistic manner of data compression.

The main advantage proposed in this paper is the ability of comparison of two attack signatures, regardless of their shape. The method takes into consideration both the shape and mutual position of two signatures. Not only that, the approach proposes applicable adaptability, that would enable sensitivity and avoid false alarms at the same time, while eluding from danger of slow shift to malicious behavior. Method gives the comparison of points of gravities, i.e. takes equally into a consideration all frequencies from the attack signature.

2 Motivation

Dealing with the survivability and fault tolerance through attack signatures has following reasons. Top-down approach, focusing on the identification of critical functionalities, proved to be more cost efficient opposite to approach of encompassing complete systems. Nevertheless, even more optimized solution can be achieved by focusing on critical functionalities only through the attack signatures, i.e. bottom-up approach [9]. Attack signatures generated in an off-line generation process, can be used in a bottom-up fashion to improve the survivability of the system, exploiting the restrictive distributed fault-tolerant systems' principles for redundancy. This way, critical functions' survivability is supported, even in the presence of intrusion. Attack recognition is carried out by intrusion detection.

Intrusion Detection System can be broken down into the following categories: **N**etwork **I**ntrusion **D**etection **S**ystems (*NIDS*) that monitors packets on the network, **S**ystem **I**ntegrity **V**erifiers (*SIV*) that monitors system files, **L**og **F**ile **M**onitors (*LFM*) that monitors log files generated by network services, and **D**eception **S**ystems (decoys, honeypots) which contain pseudo services which emulate known holes in order to entrap hackers.

Stallings categorizes **intrusion detection techniques** in two domains, based on the **detection method**. **Misuse or knowledge base** is an attempt to recognize the well known flaws or vulnerabilities of software or computer system. It can detect the general attack signatures that stem from the known holes such as exploiting a software bug. Anomaly or behavior based detection, can identify intrusions by unusual behavior of operations [10].

From the **audit source location aspect**, there are two categories of IDS. The **host based IDS** monitors a single host machine employing the audit trails of a host operating system as a main source of input. **Network based IDS** monitors hosts in one network segment, therefore auditing multiple hosts and network traffic to identify intrusion signatures. Unlike host based, network based IDS can detect attacks such as doorknob rattling, masquerading, diversionary attacks, multipronged attacks, chaining, loopback, etc.

The main advantage proposed in this paper is the ability of comparison of two attack signatures, regardless of their shape, by introducing a new ***measure of differentiation***. The method takes into consideration both the shape and mutual position of two signatures. Not only that, the approach proposes applicable adaptability, that would enable sensitivity and avoid false alarms at the same time, while eluding from danger of slow shift to malicious behavior. Method gives the comparison of points of gravities, i.e. takes equally into a consideration all frequencies from the attack signature.

The rest of this paper is structured as follows. In section 3 and 4, proposed soft computing approach is elaborated. Fifth section provides the test example results and exploited method of fuzzification of signature attacks. Sixth section concludes this paper with directives for future work.

3 Fuzzy extension of classical preference structures

Fuzzy preference structure on a set of alternatives **A** is a triplet $\Pi_{\varnothing} = (P, I, R)$ can be characterized by means of a unique binary relation **S** in A, called characteristic preference relation **S(a,b)** which represents the degree in which alternative **a** is at least as good as alternative **b**.

The satisfaction degree $S_g\left(A_i > A_j\right)$ of the comparison fuzzy numbers $A_i > A_j$ can be regarded as the preference degree of A_l to A_j, and let $\mathscr{A} = \{A_1, A_2, ..., A_n\}$ be a set of fuzzy numbers which might be the evaluations for the competing alternatives, then a **fuzzy preference relation** R_s, with respect to the satisfaction function S_g, can be defined as follows:

$$R_s : \mathscr{A} \times \mathscr{A} \to [0,1] \qquad \text{....(1)}$$

The value fuzzy preference relation $R_s\left(A_i, A_j\right)$ indicates the degree to which fuzzy number A_i dominates to fuzzy number A_j.

Different approaches to fuzzy preference relation construction can be found in literature [11]. One of the earliest, used for further developing of other approaches, is Orlovsky's fuzzy preference relation :

$$R_s\left(A_i, A_j\right) = \begin{cases} S_g\left(A_i > A_j\right) - & \text{when } S_g\left(A_i > A_j\right) \geq \\ -S_g\left(A_j > A_i\right), & \geq S_g\left(A_j > A_i\right) \\ 0, & \text{otherwise} \end{cases} \qquad (2)$$

Lee with associates [12] has proposed fuzzy preference relation:

$$R_s\left(A_i, A_j\right) = S_g\left(A_i > A_j\right) + \tfrac{1}{2} S_g\left(A_i = A_j\right) \qquad (3)$$

They have also proved next two frequently used properties of fuzzy preference relations:

$$R_s\left(A_i, A_j\right) + R_s\left(A_j, A_i\right) = 1, \quad \forall A_i, A_j \in \mathscr{A} \qquad (4)$$

$$R_s\left(A_i, A_i\right) = 0.5, \quad \forall A_i \in \mathscr{A} \qquad (5)$$

4 Fuzzy satisfaction function

The satisfaction degree of an arithmetic comparison relation of two fuzzy numbers is exploited in constructing of fuzzy preference relation. This degree is calculated by using a fuzzy satisfaction function [12].

In order to find satisfaction degree, it is essential to discrete fuzzy alternative's performances expressed by fuzzy numbers. **Fuzzy number A** is a fuzzy set defined in the real domain R and its membership function has to fulfill conditions of convexity, normality and continuity on the universes of discourse. In this paper, performances of alternatives are considered to be in a continual form.

The satisfaction function $S_g\left(A_i < A_j\right)$ is defined as:

$$S_g\left(A_i < A_j\right) = \frac{\sum_{x=I_{\min}}^{I_{\max}} \sum_{y=J_{\min}}^{\min\{x-g, J_{\max}\}} m_{A_i}(x) \Theta m_{A_j}(y)}{\sum_{x=I_{\min}}^{I_{\max}} \sum_{y=J_{\min}}^{J_{\max}} m_{A_i}(x) \Theta m_{A_j}(y)} \qquad (6)$$

while the equality comparison $S_g\left(A_i = A_j\right)$ has the following form:

$$S_g\left(A_i = A_j\right) = \frac{\sum_{x=\max\{I_{\min}, J_{\min}\}}^{\min\{I_{\max}, J_{\max}\}} m_{Ai}(x) \Theta m_{A_j}(y)}{\sum_{x=I_{\min}}^{I_{\max}} \sum_{y=J_{\min}}^{J_{\max}} m_{A_i}(x) \Theta m_{A_j}(y)} \qquad (7)$$

5 Experimental results

Test examples have included cases of various mutual position of two fuzzified attack signatures. Naturally, signatures in real comparison would overlap 100%, i.e. at least left (zero) point coincide. But, for the sake of demonstration, the following test was chosen.

Signatures represented by fuzzy values A_1 and A_2, are investigated with overlapping factor from 0% over maximal% to 0% (Figure 1). Therefore, fuzzy values moving from connected in unique spot, over max% overlapping to connected in one spot again, but from the other side, are studied. A case of separated fuzzy values is not considered because of its simplicity (values of fuzzy preference relation are obviously either 1 or -1).

Smooth moving of fuzzy preference relation values for intersection operations and approximately continual changes for overlapping of 0% over max% to 0% are obvious. Union operators result in discontinued changes in fuzzy preference relation values.

Figure 1 shows values of a fuzzy preference relation moving from 0 to 1, for overlapping from 0% over max% to 0%. These values are obtained for test example depicted by Figure 2.

Overlapping factor of fuzzified attack signatures A_i and A_j is calculated by formula:

$$\frac{A_i \cap A_j}{A_i \cup A_j} \cdot 100\% \qquad (8)$$

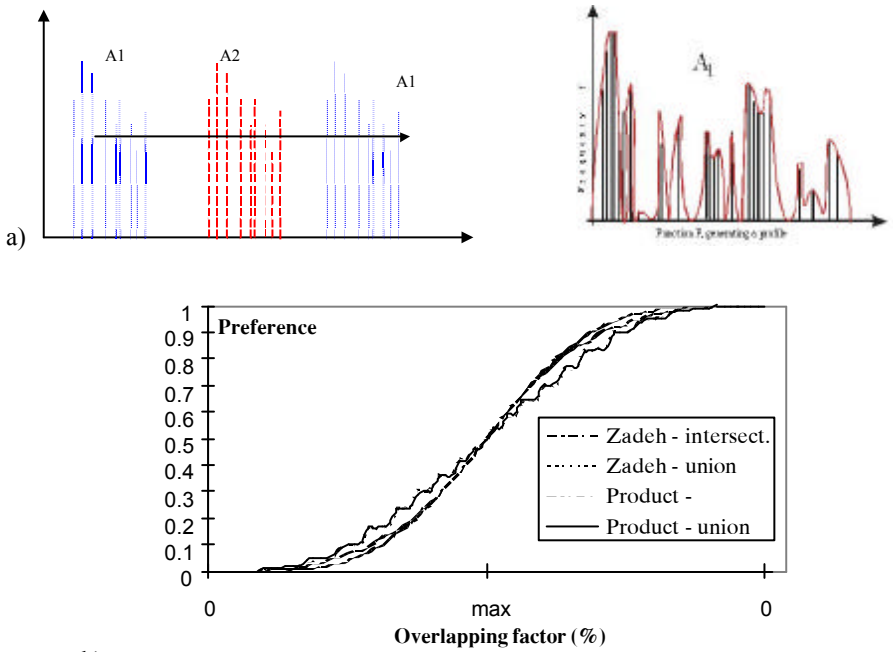

Figure 1: a) Two attack signatures, and their correspondent fuzzy value b) Preference relation depending on an overlapping extent and shape of attack signatures, for this test example

where $\cap$ and $\cup$ are defined as in [8]:

$$F \cup G = \int_U \left(m_F(u) \vee m_G(u) \right) / u \ , \quad (9)$$

Since the partial attack signature is known, simple interpolation (curve fitting) could be used to obtain certain fuzzy value (set) out of partial signature Si. This topic is outside of this paper and will not be further discussed. However, the precision of this conversion has a crucial influence on final results.

Many researchers have shown necessity of using different operators for aggregating fuzzy sets. The primary consideration in defining fuzzy subsets operators is, that they must reduce to the ordinary classic set operators when the subsets are ordinary crisp sets.

These operators can be generated by simple arithmetic transformations (Table 1) or by more complex functional transformations - **functional compensatory classes**. In this paper, authors have considered the influence of general algebraic operators (Zadeh, Mean, Mean^2, Product, Bounded Sum) in generating fuzzy satisfaction functions. Different applications of functional operators can be found in literature.

Table 1: Algebraic Intersection and Union Compensatory Operators

	Intersection	Union
Zadeh	min (m_a(x), m_b(y))	max (m_a(x), m_b(y))
Mean	(m_a(x) + m_b(y)) / 2	(min (m_a(x), m_b(y)) + 2* (max (m_a(x), m_b(y)))) / 3
Mean^2	mean(int)^2	mean(un)^2
$\sqrt{}$ Mean	mean(int)$^{1/2}$	mean(un)$^{1/2}$
Product	(m_a(x) * m_b(y))	(m_a(x) + m_b(y)) - (m_a(x) * m_b(y))
Bounded Diff/Sum	max (0, m_a(x) + m_b(y) -1)	min (1, m_a(x) + m_b(y))

Table 2: Fuzzy satisfaction function and Lee's fuzzy preference relation for Zadeh's intersection operator (test example)

Overlap. fact. (%)	$S_g(A_i > A_j)$	$S_g(A_i = A_j)$	$S_g(A_i < A_j)$	R_s
0.00%	0.00	0.00	1.00	0.00
21.35%	0.11	0.07	0.83	0.12
100.00%	0.46	0.08	0.46	0.50
56.56%	0.66	0.06	0.28	0.69
0.00%	1.00	0.00	0.00	1.00

Table 2 shows the results of fuzzy satisfaction function and Lee's preference relation (3) for Zadeh's intersection operator (first example). Obtained values fulfill satisfaction function properties as well as fuzzy preference relation properties given by equations (4,5).

Attack signature corresponding to A_i (atomic attack, smallest attack technology unit) is given at Figure 1. An example of approximating fuzzy value upon given determined partial attack signature, is also given on same figure.

The results of Lee's fuzzy preference relation given by equation (3) are presented for the above example in Figure 1 as also numerically in Table 2.

6 Conclusion

Numerous attempts to cope with intrusion detection, due to their quantitative nature, lack the flexibility of compressed data representation and methods for dealing with those.

Authors have proposed a new methodology based on fuzzy preference relation in order to capture the difference between signature attacks, regardless of their shape.

Finally, different fuzzy preference relations did not show significant influence. Experimental results are represented by tables and figures.

Further work authors plan to concentrate would include survivable architectures based on proposed method, signatures adaptability, etc.

7 References

[1] T.Lonjgstaff, , C.Chittister, R.Pethia *"Are we forgetting the risks of information technology"*, *IEEE Computer*, Dec. 2000, pp. 43-51.

[2] P.R. Stephenson, "*Intrusion Management Paper: A Top Level Model for Securing Information Assets in an Enterprise Environment",* Enterprise Networking Systems, Inc., 2000. http://www.imfgroup.com/docs/wpapers/IntrusionManagementPaper.pdf

[3] A.L. Zadeh, "Appendix", Proc. of the U.S.-Japan Seminar on Fuzzy Sets and Their Application. Berkeley, Ca., July 1974, pp.27-39.

[4] Cox, E. *The Fuzzy Systems Handbook*. Academic Press, Inc., UK, 1994.

[5] M. Manic, **Alarm systems for monitoring driven by fuzzy logic**, Proceedings, ***P***reventive ***E***ngineering & ***I***nformation ***T***echnologies, Nis, December 1994, pp.30.1-30.4.

[6] M. Manic, **Fuzzy-Operators Weight Refinements**, Proceedings, Annual Reliability & Maintainability Symposium, RAMS'99, from ***IEEE Reliability*** *Society*, Washington, DC USA, January 18-21 1999, pp.245-251.

[7] Ellison, el.al., "Survivable Network Systems: An Emerging Discipline, CMU, Software Engineering Institute, Thec Report CMU/SEI-97-TR-013, revised May 1999.

[8] Internet Security Systems http://www.iss.net/

[9] Krings, A.W., S. Harrison, J.Dickinson, and M. McQueen, `` Survivability of Computers and Networks based on Attack Signatures", *Proc. 3rd Information Survivability Workshop, (ISW-2000)* , Boston, Massachusetts, October 24-26, 2000, pp. 91-94,

[10] Stallings., W. *Network security essentials: applications and standards*, Prentice-Hall, Inc., Upper Saddle River, New Jersey, 2000.

[11] S.Zahariev. "On Orlovsky's definition of nondomination". *Fuzzy Sets and Systems* 42, 1991, pp.229-235.

[12] K.M.Lee, C.H.Cho, H.L.Kwang. "Ranking fuzzy values with satisfaction function*". Fuzzy Sets and Systems* 64, 1994, pp.295-309.