Journal of Advances in Computer Research
Quarterly pISSN: 2345-606x eISSN: 2345-6078
Sari Branch, Islamic Azad University, Sari, I.R.Iran
(Vol. 5, No. 4, November 2014), Pages: 79-90
www.jacr.iausari.ac.ir

Fuzzy Threat Assessment on Service Availability with Data Fusion Approach

Amin Sardeh Moghadam *, **Behzad Moshiri**

Control and Intelligent Processing Center of Excellence ECE, University of Tehran, Tehran, Iran

binaloodir@gmail.com; moshiri@ut.ac.ir

Received: 2014/01/06; Accepted: 2014/05/04

Abstract

Service Availability is important for any organization. This has become more important with the increase of DoS attacks. It is therefore essential to assess the threat on service availability. We have proposed a new model for threat assessment on service availability with a data fusion approach. We have selected three more important criteria for evaluating the threat on service availability and used anomaly detection algorithms to evaluate the network behavior. Anomaly of each parameter over time was measured based on its past behavior. The results of each algorithm were aggregated using the order weighted average (OWA) and finally using fuzzy inference system (FIS), threat has been calculated. We have evaluated our proposed model with data from a web server monitoring. The results show that it can provide network administrator with useful information about the status of service availability and help them to reduce threats and losses due to their actual activation.

Keywords: *Network security, information fusion, threat assessment, fuzzy logic*

1. Introduction

Today, computer networks are part of the infrastructure of any organization. Proper functioning of the infrastructure is important for any organization. Every organization has a mission, and certainly part of its mission is to respond to customers. So they try providing satisfaction for customers by providing them with proper care. Poor performance of computer networks can bring dissatisfaction to customers as well as causing financial losses to the organization. Risk management to reduce losses caused by human errors or cyber attacks is imperative. According to NIST Standard [1] threat assessment is an important part of risk management.

For instance, some organizations provide their services online via web wide web. Customers using the browser on their computer connect to the organization and benefit from their services. Disturbing or unsatisfactory performance in service delivery causes customer dissatisfaction. Customer dissatisfaction can reduce revenues and ultimately bankrupt of the organization. It is therefore essential to assess the availability of services in real-time.

Various solutions have been proposed in this field. The easiest way is to use network monitoring tools. These tools collect data from multiple network segments and provide network administrator with situation awareness of network. Network monitoring tools do not perform any processing on the data collected. Network administrator analyzes

data received based on the knowledge, experience and expertise to be aware of the status of service availability.
So the threat assessment using this method relies on the network administrator expertise which can cause damage due to human errors. So we can conclude that only use of network monitoring tools to assess the threats on network availability is not enough.
Threat modeling [2] is one of methods proposed for evaluating the threat. Threat modeling is the process helping the identifying, analyzing, documenting and assessing the vulnerabilities of the system. The system designer can prioritize security threats by Threat modeling and implementing countermeasure. Over time, however, it is proved threat modeling is not functional in threat assessment. Therefore, that various approaches have been proposed, in the following comes the description of some.
One approach to threat assessment is the use of Intrusion Detection Systems (IDS). Once the system detects an attack, it will produce an alert. Threat assessment methods are presented using network alerts and their severity.
Network behavior study is another threat assessment procedure. Various methods are used to evaluate the behavior of the network and anomaly detection is one of them. Detecting anomalies are used in intrusion detection systems to detect attacks. However, it can be used to assess the threat. Anomaly detection techniques have found a degree of anomaly for each test sample. This anomaly can be used to assess the threat.
Use alerts by intrusion detection systems as well as network behavior are among other solutions that have been proposed to assess the threat. These methods can be considered as subset of multi sensor information fusion.
In this paper, threat has been assessed only by the evaluation of network behavior. Related work is presented in section 2, and section 3 proposes a new model with a data fusion approach. Evaluation of proposed model is represented in the section 4 and finally section 5 concludes and shows the future work.

2. Related Works

As noted above, one way to assess the threats is to use IDS alerts. In the paper [3], IDS alerts after preprocessing and normalization are being verificated using NASL script Language. Then their Severity is determined based on CVSS, and finally severity of alerts is multiplied by success rate to calculate threat. Paper [4] first proposes online algorithm for the alert fusion. This algorithm is similar to alert correlation methods. Then threat priority is calculated based on D-S theory. Paper [5] chooses a number of metrics that show the availability of network and uses Dempster Shafer evidence reasoning theory to combine them. Another paper [6] has provided an information fusion framework to assess the threat. This paper examines the information aggregation models and finally suggests a fusion architecture in which Bayesian belief networks are used as a mechanism to assess the threat. “Network Security Situation Awareness Based on Multi Sensor Information fusion”, is another paper’s title [7] where the concept of a hierarchical architecture is proposed. In this paper attributes are extracted using NetFlow and Snort, and being considered as input for Multi-level feed-forward network algorithm while its output is criterion showing network security status. The paper [8] is similar to the previous article except that the support vector machine is used to assess the threat. The article [9] firstly discusses security operation center, then based on information fusion theory proposes a multilevel model for data collection, refinement,

normalization and assessment. Radial Basis Function Neural Network algorithm is used in this paper. Finally the proposed model is evaluated based on a simulated network. Paper [10] uses intrusion detection system alert and network performance parameters for assessing the threat. In this paper, two parameters, priority accuracy and services availability are introduced as the threat of evaluation criteria. Markov model and D-S intuitive reasoning algorithms are used to measure the two criteria, respectively. Paper [11] proposed a four level model in which the first level determines the target. Then desired parameters are selected based on the target. Natural language is used in the second level of fuzzification process. In this level, results of each parameter are being integrated using the proposed algorithm, EFWA. The third level is the defuzzification process. At this point, the Euclidean distance formula is used to return the results to natural language words. The fourth level linguistically represents the level of threat. This paper model is evaluated based on hospital data. Finally Paper [12] has used the previous research model, and computer network information of Malaysian Government agencies is used in assessments.

3. The Proposed Method

Network behavior assessment is a way to the evaluation of threats in computer networks environment. The correct behavior of the network indicates proper functioning of the network. Various criteria have been proposed for the performance of the network such as bandwidth, CPU usage and memory usage. But, what the network correct behavior is and how incorrect networks behavior identified. In this paper, in the assessment of the correct behavior of the network at each time step, we have compared it with past behavior. This comparison is done using anomaly detection methods. Anomaly detection and various methods are described below.

In this paper, we have proposed a multi-level modeling to assess the threat. This model is shown in the figure 1. At the lowest level of the model are devices that should be monitored by sensors. In this model, there is only one sensor that is responsible for the data collection task. The sensor can be a tool for network monitoring. The sensor collects all the data associated with each device, and gives them to the feature selection unit. Features are selected by administrator based on the threat target in the feature selection unit. If the target is specified, feature selection algorithms can be used in this section.

After feature selection, data related to each feature is given to network behavior assessment unit (Anomaly Detection Unit). This unit uses anomaly detection algorithms to detect anomaly rates. There are various algorithms to detect anomalies and each of which can be used for this purpose. The use of several anomaly detection algorithms can increase the accuracy of anomaly detection. Type and number of algorithms depends on many factors such as data type and amount of available resources. Since threat is evaluated in a real-time fashion in this model, using various algorithms may reduce efficacy.

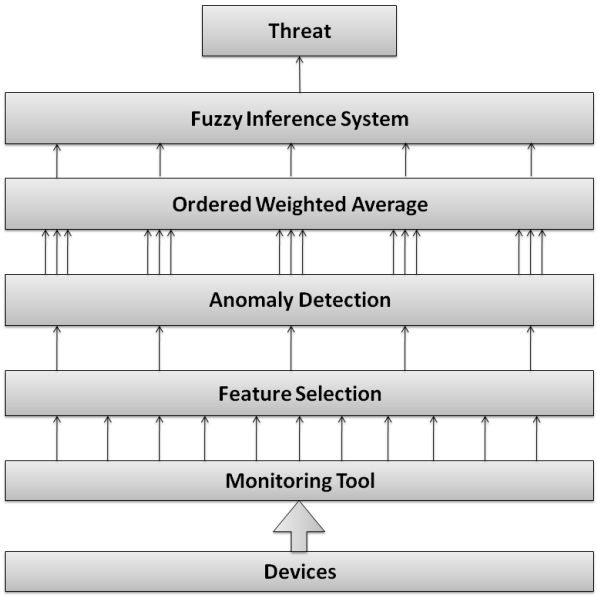

Figure 1. Fuzzy threat assessment model

Anomaly detection refers to the problem of finding patterns in the data that are not expected behavior [13]. These rule out patterns often titled as Anomalies, Outliers, Exceptions, Discordant Observations, Aberrations, Surprises, Peculiarities or Contaminants. Among these titles, anomalies, and Outliers are two words that are sometimes used interchangeably in the field of anomaly detection. Anomaly detection is widely used in various fields such as fraud in credit cards, insurance, cyber security intrusion detection, fraud detection in safety critical systems, and the military monitoring system used in military operations.
Anomaly is a data model which is not fully consistent with the definition of normal behavior. The data anomaly may be occurred due to a variety of reasons, including malicious activities such as credit card fraud, cyber influence, terrorist activities or disable critical systems. Large anomaly in different parameters of network may reflect attack in the network. Networks with high anomaly rate can also be prone to attacks. So having knowledge of the occurrence of anomalies in network parameters may prevent attacks and provide necessary information to perform an appropriate action. Various techniques for anomaly detection are provided:

- Techniques based on Classification
- Techniques based on Nearest Neighbor
- Techniques based on Clustering
- Statistical Techniques
- Techniques of Information Theory
- Spectral Techniques

Different anomaly detection algorithms can provide quotations as output in different intervals. It is suggested to normalize each algorithm in [0, 1] interval. There are several ways to normalize the subject matter out of this article scope.

The output of the network behavior assessment is several anomaly rates for each attribute. So it requires a method that can be used to aggregate various anomaly rates for each aggregation of attributes. This method should be such that even lower or higher rates had an influence on the final rate anomaly. It is therefore suggested to use the order weighted average [14] (OWA) for this purpose.
OWA output unit is an anomaly rates for each attribute in interval [0, 1]. To assess the threat posed by the different rates of anomalies in different attributes, using fuzzy inference system [15] (FIS) is proposed. Membership functions corresponding to each attribute is selected by the network administrator. The output of this stage is the threat which is indicative of the level of service availability.

4. Evaluation of the Proposed Method

To evaluate the proposed method we monitored a web server with PRTG monitoring tool [16]. We have chosen three most important features on service availability. They are CPU load, available memory and consumption bandwidth. Graphs of each of these attributes are shown in the following figures:

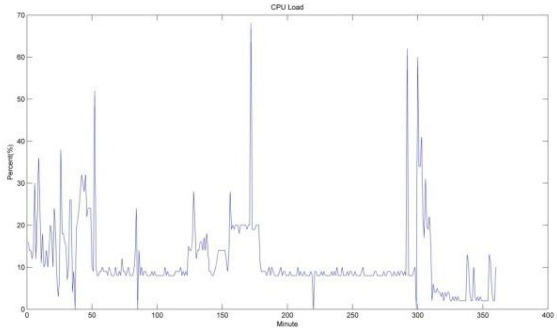

Figure 2. CPU Load

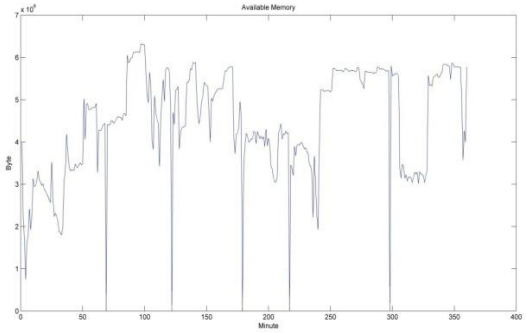

Figure 3. Available Memory

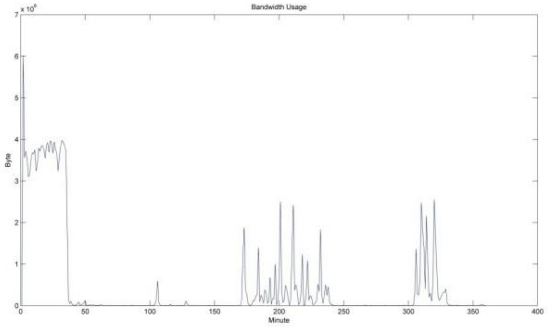

Figure 4. Bandwidth Usage

Behavior evaluation unit is provided with this data. In this paper, two anomaly detection methods are used to evaluate the behavior of the web. The first is local outlier factor [17]. Local outlier factor is based on local density concept which is obtained by the K-Nearest Neighbor (KNN). These neighbors are used in order to estimate the density. For any given sample data, the average LOF rate equals to average local density of k-

nearest neighbor and the local density of the data. To find a local density for a given data sample, the authors first find the radius of smallest hyper-sphere centered at the data instance that contains its k nearest neighbor. The local density is calculated by dividing k by the volume of this hyper-sphere. The local density for a normal data that has been compressed in an area is of its neighbors, while in an anomalous sample, local density is lower than of the nearest neighbors. Anomalous samples will result in higher-rated LOF.

K-distance (A) is the distance between object A and k nearest neighbor, the overall k of nearest neighbor in this distance is shown as N_k (A) and the availability distance between two objects will be determined by the following function:

$$\begin{aligned} &\text{reachability} - \text{distance}_k(A,B) \\ &\qquad = \max\{k - \text{distance}(B), d(A,B)\} \end{aligned} \tag{1}$$

Where d (A, B) is the distance between two objects A and B. Also, the local reach ability density of an object A is determined by the following equation:

$$\text{lrd}(A) = 1/\left(\frac{\sum_{B\in N_k(A)} \text{reachability} - \text{distance}_k(A,B)}{|N_k(A)|}\right) \tag{2}$$

Finally local outlier factors obtained using the following equation:

$$\text{LOF}_k(A) = \frac{\sum_{B\in N_k(A)} \frac{\text{lrd}(B)}{\text{lrd}(A)}}{|N_k(A)|} \tag{3}$$

To normalize results in the [0, 1] interval, we have used the following algorithm. Since the normalization is not possible for a sample, the following normalization algorithm has been applied using 14 previous samples.

$$x_{norm} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{4}$$

This is done for each sample. x_i in the formula indicates the sample data.

The results of using the LOF after normalization are shown in the following figures:

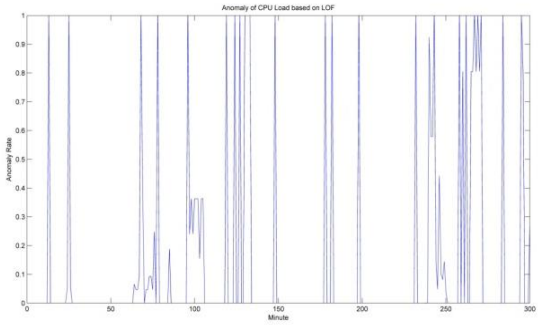

Figure 5. Anomaly of CPU load based on LOF

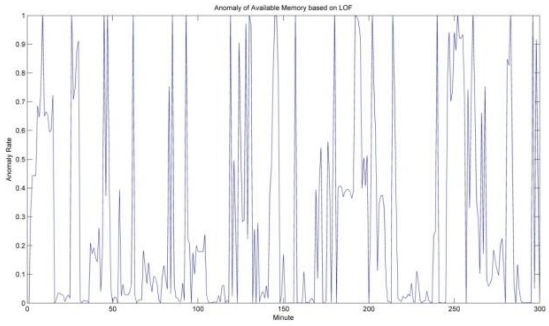

Figure 6. Anomaly of Available Memory based on LOF

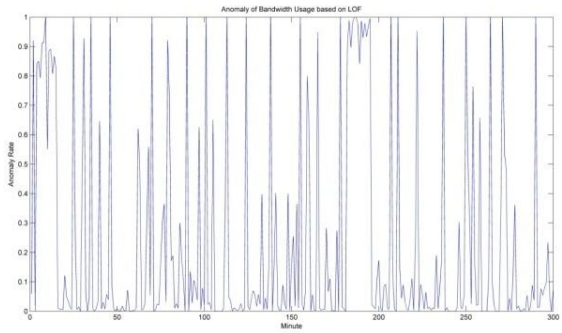

Figure 7. Anomaly of Bandwidth Usage based on LOF

But there is another way to measure the anomaly, using Parzen window [18] which is a subset of statistical methods. This technique is built upon the probability density function (PDF). This technique does not assume any unknown PDF form and lets this criterion be determined fully by using data without the obligation to choose centers' locations.

PDF is estimated by placing a well-defined kernel function on each data point of training. Gaussian kernel is mainly used in the operation. A new instance which lies in the low probability area of this is declared to be anomalous.

If p(x) is the density function to be estimated, the estimation Parzen window p(x) based on n sample is as follows:

$$\hat{p}(x) = \frac{1}{n}\sum_{i=1}^{n}\delta_i(x - x_i) \qquad (5)$$

Where is the kernel function. Thus the above formula using the Gaussian kernel will be as follows:

$$\hat{p}(x) = \frac{1}{n(2\pi)^{d/2}\sigma^d}\sum_{i=1}^{n}\exp\left\{-\frac{||x - x_i||^2}{2\sigma^2}\right\} \qquad (6)$$

So that d is the dimension of the feature space. The result of studying behavior, by using the Parzen Window after normalization is shown with the above algorithm in the following figures:

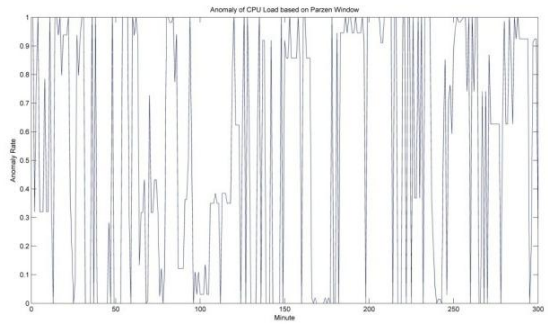

Figure 8. Anomaly of CPU Load based on Parzen Window

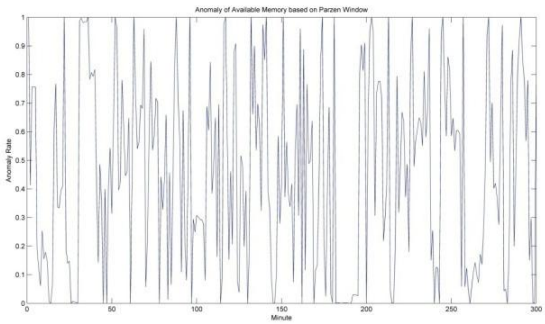

Figure 9. Anomaly of Available Memory based on Parzen Window

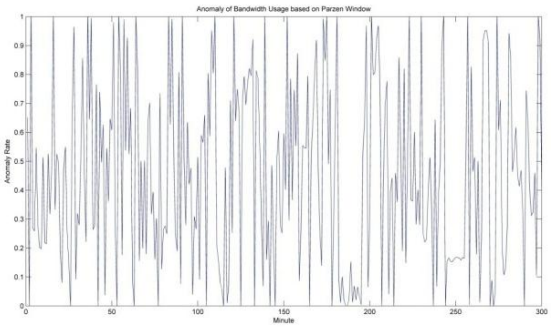

Figure 10. Anomaly of Bandwidth Usage based on Parzen Window

The above results are delivered to OWA unit to aggregate the results. Various algorithms have been proposed to determine the weights. To obtain the weights, fsolve function in MATLAB was used. The number of inputs in the assessment was 2 and alpha value is considered 0.7. Based on the above algorithm and data, the weights are 0.5038 and 0.4962, respectively. Anomaly rates after aggregation of the results using OWA is shown in the following figures:

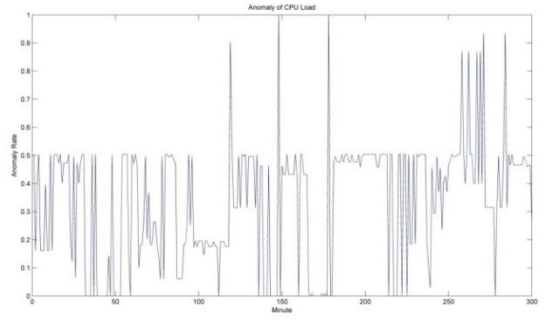

Figure 11. Anomaly of CPU Load after aggregation

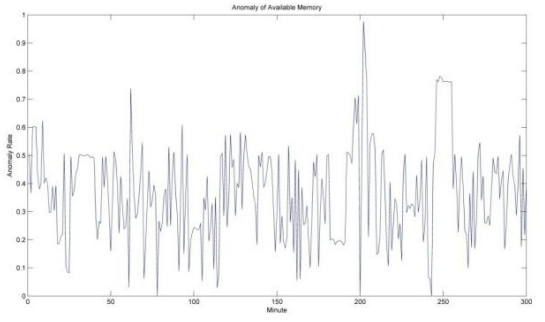

Figure 12. Anomaly of Available Memory after aggregation

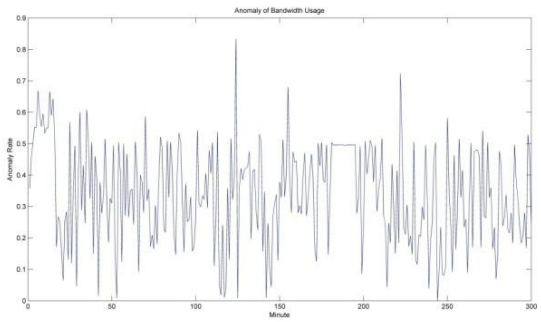

Figure 13. Anomaly of Bandwidth Usage after aggregation

Output of OWA is the input for FIS unit. Each feature is considered as an FIS input. FIS rules defined as the following:

Table 1. FIS rules

CPU Load	Available Memory	Bandwidth	Threat Rate
Low	Low	Low	Low
Low	Low	Medium	Low
Low	Low	High	Low
Low	Medium	Low	Low
Low	Medium	Medium	Low
Low	Medium	High	Medium
Low	High	Low	Low
Low	High	Medium	Medium
Low	High	High	High
Medium	Low	Low	Low
Medium	Low	Medium	Medium
Medium	Low	High	Medium
Medium	Medium	Low	Low
Medium	Medium	Medium	Medium
Medium	Medium	High	Medium
Medium	High	Low	Medium
Medium	High	Medium	High
Medium	High	High	High
High	Low	Low	Low
High	Low	Medium	Medium
High	Low	High	High
High	Medium	Low	Medium
High	Medium	Medium	High
High	Medium	High	High
High	High	Low	High
High	High	Medium	High
High	High	High	High

The membership functions defined for each feature is of trimf type with the following characteristics:

Table 2. Membership function of FIS

	Low	Medium	High
CPU Load	[0 0 0.4]	[0.2 0.4 0.6]	[0.5 1 1]
Available Memory	[0 0 0.3]	[0.2 0.4 0.6]	[0.5 1 1]
Bandwidth	[0 0 0.3]	[0.2 0.4 0.6]	[0.5 1 1]
F(u)	0.1	0.5	0.9

Results of applying these features to FIS are shown in the figure 14.

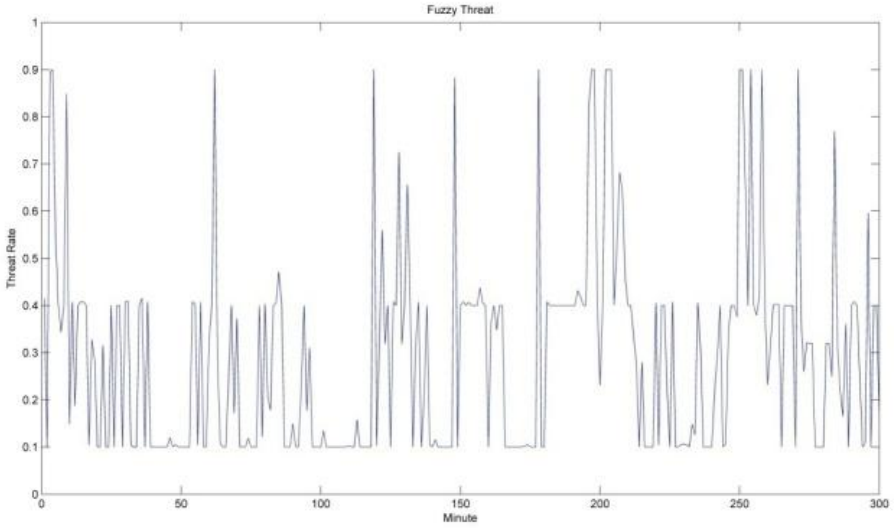

Figure 14. Result of Fuzzy Threat Assessment

If the interval [0, 0.4] is considered low threat, [0.4, 0.7] is considered medium threat and [0.7, 1] is considered high threat, we can see that the overall threat level has been low, while in some moments, threat level has been high. It can be concluded that in normal function of network, threat level is generally low and high threat level is momentary. Only if the threat level is consistently high and continuous, it would be necessary to take measures to eliminate the threat.

5. Conclusion and Future Works

In this paper, we presented a method to assess the threat on service availability and different algorithms were used. Firstly, we chose three main features for service availability and obtained the amount of their anomaly applying two anomaly detection methods. The anomaly rate in each time interval has been calculated based on feature behaviors in an hour and normalized in [0,1] interval. The results were aggregated using OWA. Anomaly rate of each attribute was considered as FIS input, and the threat was assessed based on membership functions and defined rules. In future works, other feature related to service availability can be considered, or they can be chosen using feature selection algorithms such as genetic selection algorithm. Other anomaly detection algorithms and OWA could be subjects of future works.

6. References

[1] International Standard ISO/IEC 27001. Information technology, Security techniques, Information security management systems, Requirements, 2005.

[2] E. Jangam, "Threat Modeling and its usage in mitigation security threats in an application," http://isea.nitk.ac.in/publications/ThreatModeling.pdf; Thesis Submitted in partial fulfillment of the requirements for the degree of master of technology, 2009.

[3] R. Xi, X. Yun, S. Jin, Y. Zhang, "Network Threat Assessment Based on Alert Verification," 12th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT), 2011.

[4] J. Ma, Z. Li, H. Zhang, "An Fusion Model for Network Threat Identification and Risk Assessment," International Conference on Artificial Intelligence and Computational Intelligence, AICI '09, 2009.

[5] X. Chen, S. Li, J. Ma, J. Li, "Quantitative threat assessment of denial of service attacks on service availability," International Conference on Computer Science and Automation Engineering, CSAE 2011.

[6] J.M. Beaver, R.A. Kerekes, J.N. Treadwell, "An information fusion framework for threat assessment," Information Fusion, 12th International Conference on Information Fusion, FUSION '09, 2009.

[7] L. Xiaowu, Y. Jiguo, W. MaoLi, "Network Security Situation Generation and Evaluation Based on Heterogeneous Sensor Fusion," 5th International Conference on Wireless Communications, Networking and Mobile Computing, WiCom '09, 2009.

[8] X. Liu, H. Wang, J. Lai, Y. Liang, C. Yang, "Multiclass Support Vector Machines Theory and Its Data Fusion Application in Network Security Situation Awareness," International Conference on Wireless Communications, Networking and Mobile Computing, WiCom 2007.

[9] N. Yi, Z. Qi-lun, P. Hong, "Network security management based on data fusion technology," 7th International Conference on Computer-Aided Industrial Design and Conceptual Design, CAIDCD '06, 2006.

[10] X. Chen, Q. Zheng, X. Guan, C. Lin, J. Sun, "Multiple behavior information fusion based quantitative threat evaluation," International journal of computer & security, Volume 24, Issue 3, May 2005, Pages 218–231.

[11] N.M. Zain, G.N. Samy, R. Ahmad, Z. Ismail, A.A. Manaf, "Fuzzy Based Threat Analysis in Total Hospital Information System," In: Proceedings of AST/UCMA/ISA/ACN 2010 Conferences, Miyazaki, Japan, Volume 6059, pp 1-14, 2010.

[12] F.H.M. Ali, W.M.N.H.W. Ismail, "Network Security Threat Assessment Model Based on Fuzzy Algorithm," IEEE International Conference on Computer Science and Automation Engineering (CSAE), 2011.

[13] V. Chandola, A. Banerjee, V. Kumar, "Anomaly Detection : A Survey," in: ACM Computing Surveys, 2009.

[14] Z. Xu, "An Overview of Methods for Determining OWA Weights," International Journal of Intelligent Systems, Vol. 20, 843–865, 2005.

[15] H.J. Zimmermann, "Fuzzy Sets, Decision Making and Expert Systems," Kluwer Academic Publishers, USA, 1987.

[16] PAESSLER, the network monitoring company; http://www.paessler.com/prtg [accessed 10.20.13]

[17] M. Breunig , H.P. Kriegel , T. Raymond, J. Sander, “LOF: Identifying Density-Based Local Outliers”, Proceeding of the 2000 ACM Sigmod International Conference on Management of Data.

[18] C. Chow, D. Yeung, "Parzen-window network intrusion detectors" In Proceedings of the 16th International Conference on Pattern Recognition, 2002, Vol. 4. IEEE Computer Society, Washington, DC, USA, 40385.