A Situation Awareness Model of System Survivability Based on Variable Fuzzy Set

Jinhui Zhao*[1], Yu Zhou[2], Liangxun Shuo[1]
[1]Network Information Security Laboratory Shijiazhuang University of Economics
No.136, Huai'an East Road, Shijiazhuang, China, 0311-87207577
[2]School of Electric Power,North China University of Water Resources and Electric Power
No.36 Beihuan Road, Jinshui District, Zhengzhou, China, 15037128252
corresponding author, e-mail: zhaojh9977@sohu.com, zhouyu_beijing@126.com, slx7151@sohu.com

Abstract

Because of the complexities and diversities of security alerts in modern networks, it is extremely difficult to precisely analyze and evaluate the survival situation. After analyzed the research progress of survival situation awareness, a hierarchical model of survival situational awareness is proposed, based on fusion technology of variable fuzzy set. In order to improve the discrimination of features, the efficient light-computation fusion mechanisms are employed to preproccess the multi-source data. Subsequently, the situation assessment adopted fusion technology of variable fuzzy set theory to improve the accuracy and objectivity of evaluation results. Finally, time series of services or hosts are introduced to judge the survival status and predict the trends by fuzzy reasoning. Experiments indicate that proposed model is suitable for a real network environment, and the perception results more scientific and accurate, which is valuable for generalizations and applications.

***Keywords**: multi-feature fusion, variable fuzzy set, survival situation, situation awareness, situation prediction*

1. Introduction

In the fields of network security, many safety products have been developed to guarantee the security of information. Because of the openness of the network, the vulnerability of operating system, and the security risks in hardware and software, meanwhile network viruses and network attacks are constantly variants and upgrade, it is impossible to build an absolute security network system. The third generation technology of information security came into being, which take survivability as the core. Compared with the traditional security technology, the technology of survivability is an active defense technology. It highlights the continuous services, and to protect the security and integrity of the data under attacks, even if some parts or components fail, or a malicious attacker manipulates the system. In the field of survivability, survival situational awareness is the basis for the implementation of the survival emergency strategy, the ultimate goal of the survival study is initiative to implement the appropriate strategy for survival. Timely and objective survival situational awareness can provide scientific basis and accurate judgment for the application of survival strategies.

At present, the study in survivability situational awareness is very limited. Some scholars have studied the framework and system model [1-3] for the survivability situational awareness, and presented a variety of detection and evaluation algorithm for the specific applications [4-6]. In this paper, a hierarchical model of survival situational awareness is presented, based on variable fuzzy set theory. The theory of variable fuzzy set [7-9] introduces the dialectical materialism based on traditional fuzzy theory, and considers the process from quantitative changes to qualitative changes is gradual and continuous, which is widely applied to various engineering fields. Combined with the ideas and methods of information fusion technology, objective and reliable evidences have been provided through a set of data acquisition and processing mechanisms for the choices of survival emergency policy. Case studies have shown that the model is effective and scientific with small amount of calculation.

Received September 27, 2012; Revised November 14, 2012; Accepted November 23, 2012

2. The Conceptual Model of Situation Awareness

The model of situation awareness describes the function of each component, the relationship of components and the relationship between the components and the environment, which is the premise and foundation of research. The traditional model includes three modules, namely: situation abstraction module, situation assessment module, situation forecast module, which is shown in Figure 1.

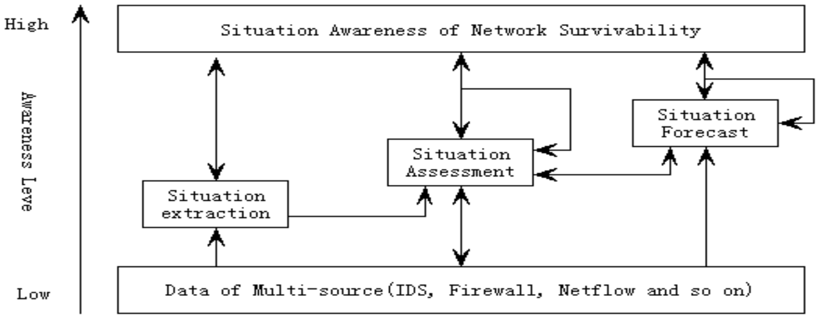
High
Awareness Leve
Low
Situation Awareness of Network Survivability
Situation Forecast
Situation Assessment
Situation extraction
Data of Multi-source(IDS, Firewall, Netflow and so on)

Figure 1. Conceptual Model of Survival Situation Awareness

The main function of The situation extraction module is to gain the useful parameter information from the mass information, and standardize them. All of those are the basis for the situational assessment.

The situation assessment module, which is the core part of the situation awareness, analyze the information of survival events, understand the interrelationship about them, and aware the survival situation. According to the historical data, the situation forecast module predicts the trend of situation in the near future. The decision-makers can made the rational decision by this forecast. These modules reflect the information processing functions at different levels.

3. Hierarchical Model of Survival Situational Assessment Based on Multi-feature Fusion

3.1.Model Structure of Survival Situation Assessment

The actual network system is a hierarchical structure, which includes the three levels: system, host, service. Survival situation assessment analyzes the exceptions and the impacts at all levels of the system. So, hierarchical model of survival situational assessment is designed based on multi-feature fusion, as show in Figure 2.

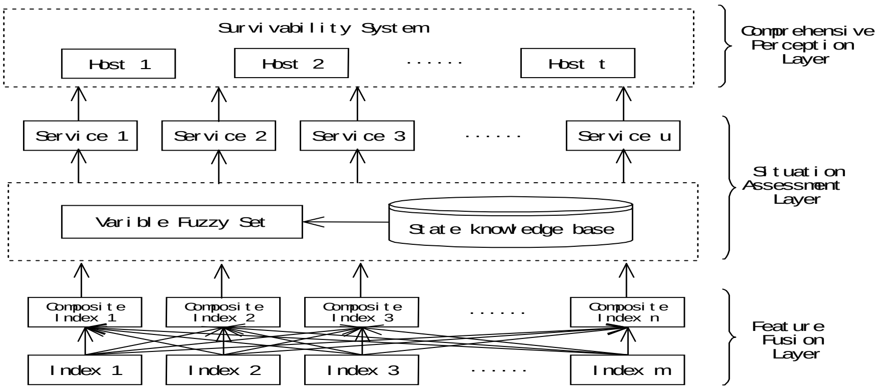

Figure 2. Structure of Survivability Awareness

The model also includes three layers: feature fusion layer, situation assessment layer and comprehensive assessment layer. The feature fusion layer is mainly to collect and pretreat the status information to provide the objective and high distinguish eigenvectors of evaluation index for the situation assessment layer. The situation assessment layer, namely fuzzy evaluation layer, evaluates the status of the basic unit in system environment according to variable fuzzy set. The comprehensive assessment layer analyzes the survival situation of services in host or hosts, or survival situation evolution process of whole system.

3.2.Feature Fusion

The survivability is a multifaceted, multi-level description of system. In generally, an incident has multiple characteristics, which have different performance at different time, and some incidents are dynamic or hidden. In acquisition system, it is very important to improve the accuracy of judgment by feature fusion, and fully describe the change of incident by least indexes. Therefore, we give the definitions of essential feature and shared feature.

Definition 1: Assume u_i is a behavior in discussed domain U, $u_i \in \mathrm{U}$. There is a related set of features $R_i = (r_{i1}, r_{i2}, \cdots\cdots, r_{ij})$ and a behavior recognition set $\mathrm{U}^{\mathrm{u}_\mathrm{i}}$. For any u_k, $u_i \neq u_k$, if $r_{ij} \neq r_{ik}$, the r_{ij} is the essential feature of u_i; if $r_{ij} = r_{ik}$, the r_{ij} is the shared feature of u_i and u_k.

Definition 2: Assume $r_1, r_2, \cdots r_i \cdots, r_n$ is a set of essential feature. $f_i(i = 1,2,\cdots\cdots n)$ is the collection function for each feature. The basic feature fusion methods define as:

Plus fusion: $r=r_1+r_2+\ldots+r_n$, $fr(r)=f_1(r_1)+ f_2(r_2)+\ldots+ f_n(r_n)$

Multiply fusion: $r=r_1*r_2*\ldots*r_n$, $fr(r)=f_1(r_1)* f_2(r_2)*\ldots*f_n(r_n)$

And Fusion: $r=r_1\&r_2\&\ldots\&r_n$, $fr(r)=f_1(r_{1)} \cap f_2(r_{2)} \cap \ldots \cap f_n(r_n)$

Or Fusion: $r=r_1 \mid r_2 \mid \ldots \mid r_n$, $fr(r)=f_1(r_1) \cup f_2(r_2) \cup \ldots \cup f_n(r_n)$

Above are some basic fusion methods. These methods can be further combinations, which are called composite fusion methods, such as r1=x1/(x1+x4), weighted fusion, and so on. The features, after composite fusion, are composite features. Composite features can reduce the number of essential feature and improve the discrimination by lightweight computing, which can enhance the accuracy of survival situation awareness.

3.3.Variable Fuzzy Model of Survival Situation Assessment

For the paces, the basis of variable fuzzy set don't describe in detail, which can refer reference [7-9]. The core of process of variable fuzzy recognition is to calculate the relative membership degree of each feature index, and the key is to select the fuzzy evaluation function.

According to the regulations of information security and the experience of the experts, we divide survival situation into five levels in order to reflect the exact status of services. Each index includes five levels too, and the range of each level is obtained in accordance with the performance of services in each level, which shows in table 1.

Table 1. The classification of service survival situation

Level	Description	Decline of performance
1	Good	<5%
2	Normal	5%~15%
3	Inferior	15%~35%
4	Worse	35%~60%
5	Worst	>60%

From the Table 1, we can establish the standard interval matrix (I_{ab}) in model of variable fuzzy recognition.

$$I_{ab} = \begin{bmatrix} [a_{11}, b_{11}] & [a_{12}, b_{12}] & \cdots & [a_{1c}, b_{1c}] \\ [a_{21}, b_{21}] & [a_{22}, b_{22}] & \cdots & [a_{2c}, b_{2c}] \\ \vdots & \vdots & \vdots & \vdots \\ [a_{m1}, b_{m1}] & [a_{m2}, b_{m2}] & \cdots & [a_{mc}, b_{mc}] \end{bmatrix} = (a_{ih}, b_{ih}) \tag{1}$$

Where h=1, 2, 3, 4, 5 is the level.

According with the bound of adjacent interval in I_{ab}, the scope interval matrix of variable fuzzy set is

$$I_{cd} = \begin{bmatrix} [c_{11}, d_{11}] & [c_{12}, d_{12}] & \cdots & [c_{1c}, d_{1c}] \\ [c_{21}, d_{21}] & [c_{22}, d_{22}] & \cdots & [c_{2c}, d_{2c}] \\ \vdots & \vdots & \vdots & \vdots \\ [c_{m1}, d_{m1}] & [c_{m2}, d_{m2}] & \cdots & [c_{mc}, d_{mc}] \end{bmatrix} = (c_{ih}, d_{ih}) \qquad (2)$$

Indexes in information system include tow types: efficiency index and cost index. The value of efficiency index is higher and the benefit is greater. The smaller value in cost index may be more beautiful. If $R_1 \in \mathrm{R}$， $\mathrm{R}_2 \in \mathrm{R}$, then $R_1 \cap \mathrm{R}_2 = \Phi$, where R_1 is the set of efficiency index and R_2 is the set of cost index. According to the classification of indexes and the actual situation, the points are determined where the degree of relative membership equal to one, $\mu_{\underset{\sim}{A}}(u) = 1$. So the M matrix of i index and h level is built.

$$I_M = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1c} \\ M_{21} & M_{22} & \cdots & M_{2c} \\ \vdots & \vdots & \vdots & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mc} \end{bmatrix} = (M_{ih}) \qquad (3)$$

3.4.Process of Variable Fuzzy Recognition

After the establishment of a viable variable fuzzy model, we apply the method of variable fuzzy recognition to perceive the system's survival trend. We firstly build the matrix of original data.

$$V = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1m} \\ v_{21} & v_{22} & \cdots & v_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ v_{n1} & v_{n2} & \cdots & v_{nm} \end{bmatrix} \qquad (4)$$

Its relative difference degrees ($D_{\underset{\sim}{A}}(u)$) can be calculated by relative difference function, which presents

$$\begin{cases} D_{\underset{\sim}{A}}(u) = \left[\dfrac{r-a}{M-a} \right]^{\beta}; r \in [a, M] \\ D_{\underset{\sim}{A}}(u) = -\left[\dfrac{r-a}{c-a} \right]^{\beta}; r \in [c, a] \end{cases} \qquad (5)$$

$$\begin{cases} D_{\underset{\sim}{A}}(u) = \left[\dfrac{r-b}{M-b} \right]^{\beta}; r \in [M, b] \\ D_{\underset{\sim}{A}}(u) = -\left[\dfrac{r-b}{d-b} \right]^{\beta}; r \in [b, d] \end{cases} \qquad (6)$$

According to the relative membership function

$$\mu_{\underset{\sim}{A}}(u) = (1 + D_{\underset{\sim}{A}}(u)) / 2 \qquad (7)$$

We get the relative membership degree of i index and h level. Therefore, the matrix of relative membership degree is $[\mathrm{U_h}] = \mu_{\underset{\sim}{A}}(u)_h$.

Applying variable fuzzy recognition model to calculate the optimal membership degree:

$$ {}_j u'_h = 1 \Big/ 1 + \left\{ \frac{\sum_{i=1}^{m} [w_i(1-\mu_{\underset{\sim}{A}}(u)_h)]^p}{\sum_{i=1}^{m} (w_i \mu_{\underset{\sim}{A}}(u)_h)} \right\}^{a/p} \tag{8} $$

Where u' refers optimal membership degree, w_i to index weight, m to recognizing indexes number, a to optimal model parameter, p to distance parameter and p=1 is Hamming distance, and p=2 is Euclidean distance.

The relative membership degree distribution function should satisfy the normalization.

$$ \sum_{h=1}^{c} \mu_{\underset{-h}{A}}(u) = 1 \tag{9} $$

The evaluation level of each part will be classified with

$$ H(u) = (u'_h / \sum_{h=1}^{c} u'_h) h \tag{10} $$

$H(u)$ reports overall relative characteristics between h and $\mu_{\underset{-h}{A}}(u)$, which comprehensively and objectively describe the attribution of the sample.

3.5.Comprehensive Assessment

After analyzed the fuzzy assessment of elements, the comprehensive analysis is required to the services or hosts in each layer or the entire network. In this part, the weighted average method is employed.

$$ L = \vec{H} \cdot \vec{W} = [H_1, H_2, \cdots, H_m] \cdot [W_1, W_2, \cdots, W_m] \tag{11} $$

Where: L is the survival situation of the integrated unit, H_i is the assessment result of basis unit, W_i is the weight of basis unit that is alculated by AHP according to the importance of service resources.

4. Survival Situation Forecast Based on Fuzzy Reasoning

From above analysis, we can aware the survival situation at each level. In order to select and develop more effective survival strategies and change from passive to active, we need to predict the survival trends of the system environment and the services. Because the change of survival situation is random and irregular, it is difficult to describe the changes by function or mathematical model. Fuzzy reasoning method can effectively describe the fuzzy relationship between the forecast factors and predictors, so in this part situation forecast based on fuzzy reasoning is employed to predict the trends.

4.1.Ideas and steps of fuzzy reasoning

The ideas of fuzzy reasoning takes known data of m former cycles as input, predicts the survival trend at next time unit by approximate reasoning of characteristic expansion. Its steps are as follows.

(1) Taking the interval between the maximum value and minimum value as a discuss domain {U}, according to statistics about the actual awareness data of the services or the hosts.

(2) The value of x at U is discretized, and divided into five level too. Each level is a fuzzy set $\{Y_k\}$, $Y_k \in U$ k=1, 2, ……, l. The distribution function of the membership can use simple symmetrical linear distribution function, such as triangular distribution function and so on.

(3) Taking the fuzzy information of m former time units as predicted input and fuzzy reasoning model as the fuzzy controller, to obtain the fuzzy output of the next unit by the corresponding fuzzy transform. (N-m) × m-dimensional composite fuzzy conditional statements are given according to the n former time units.

$$\begin{cases} if \quad Y_{1,1}^k \quad and \quad Y_{1,2}^k \quad \cdots and \quad Y_{1,j}^k \quad \cdots and \quad Y_{1,m}^k \quad then \quad Y_{1,m+1}^k \\ if \quad Y_{2,1}^k \quad and \quad Y_{2,2}^k \quad \cdots and \quad Y_{2,j}^k \quad \cdots and \quad Y_{2,m}^k \quad then \quad Y_{2,m+1}^k \\ \qquad\qquad\qquad \cdots \qquad\qquad\qquad\qquad \cdots \\ if \quad Y_{i,1}^k \quad and \quad Y_{i,2}^k \quad \cdots and \quad Y_{i,j}^k \quad \cdots and \quad Y_{i,m}^k \quad then \quad Y_{i,m+1}^k \\ \qquad\qquad\qquad \cdots \qquad\qquad\qquad\qquad \cdots \\ if \quad Y_{n-m,1}^k \quad and \quad Y_{n-m,2}^k \quad \cdots and \quad Y_{n-m,j}^k \quad \cdots and \quad Y_{n-m,m}^k \quad then \quad Y_{n-m,m+1}^k \end{cases} \tag{12}$$

Where: $\{Y_{ik}\} \in U(x)$, i=1, 2, ⋯, n-m; j=1, 2, ⋯, m+1; k=1, 2, ⋯, l.

It is a m × (n-m) dimension fuzzy controller, which describes the fuzzy relationship between the m+1^{-th} factor and the m input factors. Formula (12) can simplify as:

$$R = \bigcup_{i=1}^{n-m} \left(\coprod_{j=1}^{m+1} Y_{i,j}^k \right) \tag{13}$$

Where: $\{Y_{ik}\} \in U(x)$, i=1, ⋯, n-m; j=1, ⋯, m+1, k=1, ⋯, l.

4.2. Fuzzy Reasoning Prediction Based on the Level of Eigenvalue

(1) The discuss domain is determined according to census the eigenvalues of n former continuous time units. The eigenvalues of level are gained by survival situation assessment

(2) in the fuzzy controller, the eigenvalues of m former continuous time units $\{H(\underset{\sim}{A}_{i,j})\}$ are taken as input data, the eigenvalue of m+1^{-th} time unit $\{H(\underset{\sim}{B}_i)\}$ is the output.

$$\underset{\sim}{R} = \bigcup_{i=1}^{n-m} ((\prod_{j=1}^{m} H(\underset{\sim}{A}_{i,j}) \times H(\underset{\sim}{B}_i)) \tag{14}$$

In the process of dynamic prediction, the new eigenvalue constantly produce. Because the future trends closely related to near eigenvalues and have poor relationship with longer historical data, the eigenvalues of n former continuous time units are constantly updated.

$$H^{(0)}(T) = \{h^{(0)}(t-n+1), h^{(0)}(\mathrm{t}-n+2), \cdots, \;\; h^{(0)}(t)\} \tag{15}$$

(3) Taking the eigenvalues of m former continuous time units as input, the outputs of the fuzzy controller are the forecast results of the m+1^{-th} time unit. We get the interval of the predicted output $[H_a(\underset{\sim}{B}_i)$, $H_b(\underset{\sim}{B}_i)]$ by properly adjusting the eigenvalues of effective subschema, which describes the range of possible changes in the near future.

5.Experiment and Analysis

5.1.Experimental Environment

In order to verify the performance of proposed model, applying Lincoln laboratory DARPA2000 data, which included tow attack scenario instance: LLDOS1.0 and LLDOS2.0.2. Experimental environment of LLDOS1.0 was built in laboratory. Because the scene of

LLDOS1.0 lasted 180 minutes, data statistics were carried out by one minute. Survival awareness is a comprehensive assessment, which includes security, availability and integrity. In order to clearly describe changes in the system, the DDoS duration and intensity was increased.

5.2.Simulation and Analysis

In the scene of LLDOS1.0, three hosts were attacked. After successful invasion, DDOS tools were installed in these puppet hosts, and to assault the Web services by false IP. Figure.2 showed the time curves of the survival situation, which included the security, the availability and the survivability.

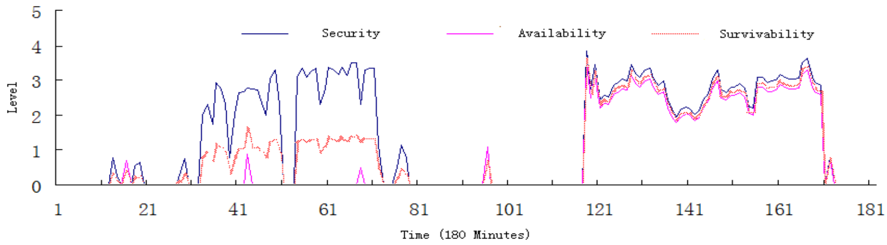

Figure 3. Curves of Survival Situation

Based on mass alarm information of IDS and network performance indicators, reference [11] presented a hierarchical threat assessment model for network security, which focused on the quantization of assessment parameters to accurately assess the security trend. Compared with proposed model, the model lacks the idea of data fusion and the continuity of events and relativity of parameters. References [12] established a three-tier survival situational awareness model based on gray relational analysis. However, the relative membership degree in variable fuzzy theory has wider useful environment than the relative membership degree calculated by triangle bleaching function in gray theory. Moreover, the structure of proposed model is reasonable and clear layer, which adapt to automatically select emergency strategies for each layer.

There are some other methods to predict the trend of survival situation. Each of them has its advantages and disadvantages, and its adaptation range. For example, the algorithm of grey model [5] is simple and easy to implement with faster execution speed, and do not need to set parameters and human intervention in the forecasting process, while forecast results do not reflect the randomness and periodic. The prediction method, based on fuzzy reasoning, has smaller prediction error, and can reflect the randomness and periodic of overall trends, which is closer to the actual condition.

6. Conclusion

In order to provide scientific and objective basis for active formulation and adjustment of survival strategies, situation awareness of system survivability needs to aware and understand the factors that cause the changes of system survivability, and predict the changed trend of system survivability according to historical data. According to conceptual model of situational awareness, this paper proposed a hierarchical model of survival situational awareness. Proposed model emply variable fuzzy sets to study survival situation awareness and fuzzy reasoning to predict the changed trend, which introduce new ideas and methods. The experiments indicate that the model can scientifically and effectively percept the survival trend, which is a better solution with small amount of calculation.

Acknowledgment

The authors would like to acknowledge Shijiazhuang University Of Economics in support with the initial fund of scientific research to pursue our Ph.D degree.

References

[1] LIN Chuang, WANG Yang, LI Quan-Lin. *Stochastic Modeling and Evaluation for Network Security*. Chinese Journal of Computers, 2005, 28 (12): 1943-1955.

[2] WANG Huiqiang, LAI Jibao, HU Mingming, LIANG Ying. *Research on Key Technologies for Implementing Network Security Situation Awareness*. Geomatics and Information Science of Wuhan University, 2008, 33(10): 995-998.

[3] SUN Zhixin, XU Hongxia. Research on a New Network Security Assessment System. *Journal of Nanjing University of Posts and Telecommunications (Natural Science)*, 2006, 26(3): 28-32.

[4] WANG Yifeng, LI Tao, HU Xiaoqin, SONG Chen. A Real-Time Method of Risk Evaluation Based on ArtificialImmune System for Network Security. *Acta Electronica Sinica*, 2005, 33(5): 945-949.

[5] ZHAO Guosheng, WANG Hui-qiang, WANG Jian. A situation awareness model of network security based on grey Verhulst model. *Journal of Harbin Institute of Technology*, 2008, 40(5):798-801.

[6] MAN Dapeng, YANG Wu, YANG Yongtian, ZHOU Yuan. Study on Threat Evaluation Method for Network Security. *Journal of Shenyang Jianzhu University (Natural Science)*, 2008, 24(4): 708-711.

[7] CHEN Shouyu. Theory and model of engineering variable fuzzy set-Mathematical basis for fuzzy hydrology and water resources. *Journal of Dalian University of Technology*, 2005, 45(2): 308-312.

[8] CHEN Shouyu. Philosophical foundation of variable fuzzy sets theory. *Journal of Dalian University of Technology (Social Sciences)*, 2005, 26(1): 53-57.

[9] CHEN Shouyu. Theory of Variable Fuzzy Sets and Variable Model Sets. *Mathematics in Practice and Theory*, 2008, 38(18): 146-152.