A class of Non-statistical Traffic Anomaly Detection in Complex Network Systems

Wenlin Han[1], Wei Xiong[2], Yang Xiao[3], Magdy Ellabidy[4], Athanasios V. Vasilakos[5], Naixue Xiong[6, *]

[1]School of Computer Science and Technology, Huazhong University of Science and Technology, China
[2]Center of Computing & Experimenting, South Central Univ. for Nationalities, Wuhan, China
[3]Deptartment of Computer Science, University of Alabama, Tuscaloosa, AL 35487-0290, USA
[4]Dept. of Computer Science and networking, Wentworth Institute of Technology, USA
[5]Dept of Computer and Telecommunications Engineering, Univ. of Western Macedonia, Greece
[6]School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, China
Email: winni.hann@gmail.com, ccnuxw@sina.com, yangxiao@cs.ua.edu, ellabidym@wit.edu,
vasilako@ath.forthnet.gr, xiongnaixue@gmail.com

Abstract—**Recently Network traffic anomaly detection has become a popular research tendency, as it can detect new attack types in real time. The real-time network traffic anomaly detection is still an unsolved problem of network security. The network traffic appears as a complex dynamic system, precipitated by many network factors. Although various schemes have been proposed to detect anomalies, they are mostly based on traditional statistical physics. In these methods, all factors are integrated to analyze the variation of the network traffic. But in fact, the changing trend of network traffic at some moment is only determined by a few primary factors. In this paper, we present a non-statistical network traffic anomaly detection method based on the synergetic neural networks. For our method, a synergetic dynamic equation based on the order parameters is used to describe the complex behavior of the network traffic system. When the synergetic dynamic equation is evolved, only the order parameter determined by the primary factors can converge to 1. Therefore, the network traffic anomaly can be detected by referring to the primary factors. We evaluate our approach using the intrusion evaluation data set of the network traffic provided by the defense advanced research projects agency (DARPA). Experiment results show that our approach can effectively detect the network anomaly and achieve high detection probability and low false alarms rate.**

Keywords-Anomaly detection; Network traffic; Order parameter; Synergetic neural networks

I. Introduction

Network traffic anomaly detection has become a popular research tendency. Although some significant measurements have been carried out in the network security management, the real-time network traffic anomaly detection is still an unsolved problem of network security.

Generally, the network intrusion detection approaches can be divided into two classes: misuse detection and anomaly detection. The former uses labeled patterns of past anomalies to detect anomalies possibly occurring [1, 2]. The later establishes the normal patterns of the network behavior by training to detect anomalies substantially deviated from these patterns [3].

The early research work on the anomaly detection was mostly signature-based, which needs to update the new signature database frequently and is not suitable for the real-time network anomaly detection. Thus, more investigations on the network traffic anomaly detection have been performed, which can detect new type of anomalies and fulfill the requirement of the real-time anomaly detection.

The proposed approaches of anomaly detection mostly adopt the traditional statistical physics methods to extract the macro features of the network traffic, such as self-similarity [4, 5], entropy [6, 7], probability distribution [8]. Then the network traffic anomaly can be detected by using various pattern recognition techniques, such as neural networks [9], hidden Markov model [10], integrated access control [11], sensor fusion [12], and machine learning [13].

However, the modeling of the network traffic is a complex process which is driven by many factors such as network devices, topology, transfer protocol, as well as the interactive cooperation and competition between the network users. Thus, the network traffic often shows the non-linear, non-stationary and complex nature characteristics [14-18, 20] and is a complex dynamic system. Its macro behavior is yielded by the collaborative activities of these factors. At some moment, the changing trend of the network traffic is only determined by a few primary factors and the contributions of the rest secondary factors are slight [19]. When anomalies occur, the network traffic system will transform from a normal equilibrium to an abnormal equilibrium. The transformation is a catastrophe process, but not stationary [22] and it is promoted by the primary attack factors. The traditional statistical physics methods based on the stationary hypothesis of the network traffic always ignore the real catastrophe process.

In order to solve the above problems, we present a non-stationary network traffic anomaly detection approach based on Synergetic Neural Network (SNN). In our method, a synergetic dynamic equation is used to describe the complex behaviors of the network traffic system. The synergetic order parameters solved by this dynamic equation essentially embody the effect of the primary factors that guide the transformation of network states. When the synergetic order parameters are evolved, only

*In this paper, Naixue Xiong is the corresponding author.

1545-0678/12 $26.00 © 2012 IEEE
DOI 10.1109/ICDCSW.2012.7

the order parameter determined by the primary factors can converge to one. Therefore, the network traffic anomaly can be detected by referring to the primary factors. To evaluate the performance of our approach, we tested our system on the standard defense advanced research projects agency (DARPA) data sets and compare the results with a statistical physics method. The results show that our approach is effective in detecting network anomaly.

The rest of the paper is organized as follows. In Section II, we exhibit our anomaly detection based on SNN. Section III shows some experiment results and analyzes the performance of our method. Lastly, we propose a conclusion and give out the future works in Section IV.

II. ANOMALY DETECTION BASED ON SNN

A. *Network traffic anomaly based on SNN*

In this section we discuss our method based on SNN and provide the structure of an algorithm to detect network traffic anomalies.

As we know, network traffic often shows the non-linear, non-stationary and complex nature characteristics which are the hallmark of a complex dynamic system. Its macro behavior is yielded by collaborative activities of many factors. However, at a moment the changing trend of the network traffic is only determined by a few primary factors, the contributions of other factors are slight. The generation of the network traffic is the result of the aggregation of all factors, such as network devices, topology, transfer protocol, the interactive cooperation and competition between network users and attackers. However, all these factors are not equal on the contribution of the network traffic. When anomalies occur, the primary factors is represented by the behavior of abnormal users or attackers and the network traffic system shows the abnormal state determined by the behavior of abnormal users or attackers. Here we call the primary factors as the order parameters, which are the focus of our research of the network anomaly detection.

The motion of the network traffic system depends on the transformations among the equilibrium states determined by the primary factors. In the normal network traffic (when the state of the network traffic is normal, that is to say, no anomalies happened, we call the network traffic normal network traffic.), the network traffic system maintains the stationary variation tendency dominated by the normal primary factors even if the fluctuation of the network traffic generated by other factors may be great. We call this network state normal equilibrium state. After attacks occur, the network traffic system maintains stationary abnormal state dominated by the primary anomaly factors. We call this network state abnormal equilibrium state. When attacks occur, the network state will transform from the normal equilibrium to the abnormal equilibrium. The state change of the network traffic is transient and is a catastrophe process.

The network traffic anomaly detection using SNN is a process of pattern recognition based on the synergetic. In this process, the testing data is called identified patterns and the training data is called prototype patterns. The process to recognize identified patterns is a process of mapping the identified patterns to some existing prototype patterns where the process is dominated by the primary factors called order parameters.

According to the basic concepts of the synergetic [19], a dynamical system researched in the synergetic can be expressed as follows:

$$\dot{q} = -\frac{\partial V}{\partial q^+}\ ,\ \dot{q}^+ = -\frac{\partial V}{\partial q} \qquad (1)$$

The treatment to recognize identified patterns q can be described as a dynamic process: after making the initial patterns labeled as $q(0)$ from an intermediate state $q(t)$ into a prototype pattern v_k, that is, the prototype pattern v_k is near $q(0)$, the process can be described as $q(0) \to q(t) \to v_k$. Where, $q(0)$ is the identified network traffic, v_k is the stored normal network traffic or the various abnormal network traffic, the intermediate state $q(t)$ can be expressed by the order parameter ξ_k determined by the primary factors of the network traffic. This process can be specifically described by a given dynamic Equation (2). Assuming the number of the prototype patterns is M and the dimension of the prototype pattern vector is N, where $M \le N$ is requested.

$$\dot{q} = \sum_k \lambda_k v_k (v_k^+ q) - B \sum_{k \neq k'} (v_{k'}^+ q)^2 (v_k^+ q) v_k - C(q^+ q) q + F(t) \qquad (2)$$

Where, q is the testing network traffic pattern vector with the initial input pattern value $q(0)$. λ_k is the attention parameter. Only when it is positive, testing patterns can be identified. $F(t)$ is the fluctuation factor of the network traffic and can be ignored. B and C are specified coefficients and must be greater than 0. v_k is the prototype pattern vector, $v_k = (v_{k,1}, v_{k,2}, \cdots, v_{k,N})^T$. v_k^+ is the adjoint vector of v_k and they meet:

$$(v_k^+, v_{k'}) = v_k^+ v_{k'} = \delta_{k,k'} = \begin{cases} 1, k = k' \\ 0, k \neq k' \end{cases} \qquad (3)$$

v_k must be prepared by normalizing and zero-mean:

$$\sum_{l=1}^{N} v_{k,l} = 0\ ,\ \|v_k\|_2 = (\sum_{l=1}^{N} v_{k,l}^2)^{1/2} = 1 \qquad (4)$$

Vector q can be decomposed into a prototype vector v_k and the remaining vector w:

$$q = \sum_{k=1}^{M} \xi_k v_k + w\ ,\ v_k^+ w = 0 \qquad (5)$$

The adjoint vector of q is defined as follows:

$$q^+ = \sum_{k=1}^{M} \xi_k v_k^+ + w^+\ ,\ w^+ v_k = 0 \qquad (6)$$

Obviously there is a relationship:

$$(v_k^+, q) = (q^+, v_k) \qquad (7)$$

Typing (5) into (7), according to the orthogonal relationship, the order parameter ξ_k is defined as following:

$$\xi_k = (v_k^+, q) = v_k^+ q \qquad (8)$$

Style described in (2) is a powerful dynamics equation and if we neglect the fluctuation factor $F(t)$ of the network traffic, the potential function V can be described as following:

$$V = -\frac{1}{2}\sum_{k=1}^{M}\lambda_k (v_k^+ q)^2 + \frac{1}{4}B\sum_{k\neq k'}(v_k^+ q)^2 (v_{k'}^+ q)^2 + \frac{1}{4}C(\sum_{k=1}^{M}(v_k^+ q)^2)^2 \; (9)$$

According to the Equations (1), (2) and (8), correspondingly the dynamic equations and the potential function described by the order parameter are as follows:

$$\dot{\xi}_k = \lambda_k \xi_k - B\sum_{k'\neq k}\xi_{k'}^2 \xi_k - C(\sum_{k'=1}^{M}\xi_{k'}^2)\xi_k \qquad (10)$$

$$V = -\frac{1}{2}\sum_{k=1}^{M}\lambda_k \xi_k^2 + \frac{1}{4}B\sum_{k'\neq k}\xi_{k'}^2 \xi_k^2 + \frac{1}{4}C(\sum_{k'=1}^{M}\xi_{k'}^2)^2 \qquad (11)$$

We know that when the potential energy of a system reaches the lowest, the system dominated by the order parameters comes into the most stable state. In this case, the order parameters almost can not change. That is to say, the steady state of the network system is decided by the following formula:

$$\dot{\xi}_k = 0 \, , 0 \leq k \leq M \qquad (12)$$

That is:

$$\dot{\xi}_k = \lambda_k \xi_k - B\sum_{k'\neq k}\xi_{k'}^2 \xi_k - C(\sum_{k'=1}^{M}\xi_{k'}^2)\xi_k = 0 \qquad (13)$$

If it is defined:

$$D = (B + C)\sum_{k'=1}^{M}\xi_{k'}^2 \qquad (14)$$

Then the Equations (11) and (13) can be rewritten as

$$\dot{\xi}_k = \xi_k(\lambda_k - D + B\xi_k^2) \qquad (15)$$

$$\xi_k(\lambda_k - D + B\xi_k^2) = 0 \qquad (16)$$

According to Equation (15), the model is constructed with three layers (Figure 1). The top layer is the input layer in which unit j receives component $q_j(0)$ of need-recognized pattern vector's original value $q(0)$. All order parameter components form the middle layer, where order parameter ξ_k is gotten by all summing angle index j through each input value $q_j(0)$ multiplying its joint unit $v_{k,j}^+$. The active order parameter ξ_k recognize special prototype pattern chosen by angle index k. According to the dynamics equation, the network will be evolved to the end state that only one of the order parameters survives and q_j is gotten through reciprocity and competition of D. The down layer is the output layer in which output pattern can be expressed $q_j(t) = \sum_k \xi_k(t) v_{k,j}$, where q_j is active of output unit j and ξ_k is the end state of the middle layer. There is $\xi_k = 1$(8) if $k = k_0$, otherwise $\xi_k = 0$. $v_{k,j}$ is the component j of prototype vector.

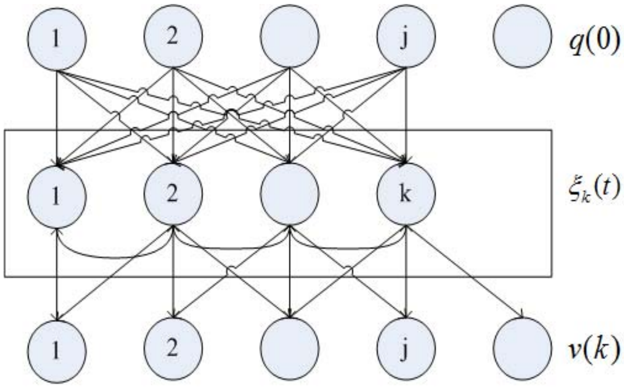


Figure 1. The framework of SNN

B. The anomaly detection method

In this method, the network traffic is extracted as the network traffic time series in packets per second and then the time series is dealt with by SNN to detect anomalies. In each type of the network traffic, we select some network traffic time series sharing the same size N that can reflect the normal and the abnormal, these makeup a prototype pattern set that includes M components.

The process of the anomaly detection includes two stages: the training stage to learn the prototype patterns of each type of the network traffic and the test stage to detect the network traffic anomaly. The detection steps are given as follows.

1) The training stage

a) Choose the training pattern vectors $\{X_1, X_2, ..., X_n\}$ from the training data set.

b) By using the training pattern vectors $\{X_1, X_2, ..., X_n\}$, compute the prototype pattern vector v_k with normalized and zero-mean condition.

c) Compute the corresponding adjoint vector v_k^+ of the prototype pattern vector v_k.

2) The test stage

a) Test on the test pattern vector $q(0)$ consists of the test network traffic data with normalized and zero-mean condition.

b) Achieve the corresponding order parameter ξ_k of each prototype pattern.

c) Keep evolving according to (17), which is an order parameter dynamic equation, until the synergetic neural network converges to a specific prototype pattern. Now, we've finished the anomaly detection processes and get the detection result of the pattern vector $q(0)$. Thus the anomaly detection processes have been completed.

$$\xi_k(n+1) - \xi_k(n) = \gamma(\lambda_k - D + B\xi_k^2(n))\xi_k(n) \qquad (17)$$

$$D = (B + C)\sum_{k'}\xi_{k'}^2(n) \qquad (18)$$

Where, γ is the iterative step.

III. Experimental Analysis

A. Experimental data

The network traffic data used in our experiments was obtained from the DARPA intrusion detection evaluation database. It includes five weeks of network traffic data. In this data, the traffic data of week 1 and week 3 contain no attacks. The data of week 2 lack the labeled information of the exact time when the attacks occur. Thus we only used network traffic data of weeks 4 and 5. Four types of network attacks, including Denial of Service (DoS), PROBE (Surveillance /Probing), U2R (User to Super user (root) and R2L (Remote to Local user) are collected. The attacks on Friday, week 5 are with the longest short-length (287) among all days and contain 4 types of attacks, which are extracted as the training data. In details, for each type of the attacks and the normal data on Friday, week 5, a sub-serial with a length of 143 (287/2) was extracted as a training vector. Then an amount of the sub-serials with the same length separated from the all-day's data (except the training data) were provided as the testing data. The mean generation rate of the attack for the testing data is 8%.

B. Results

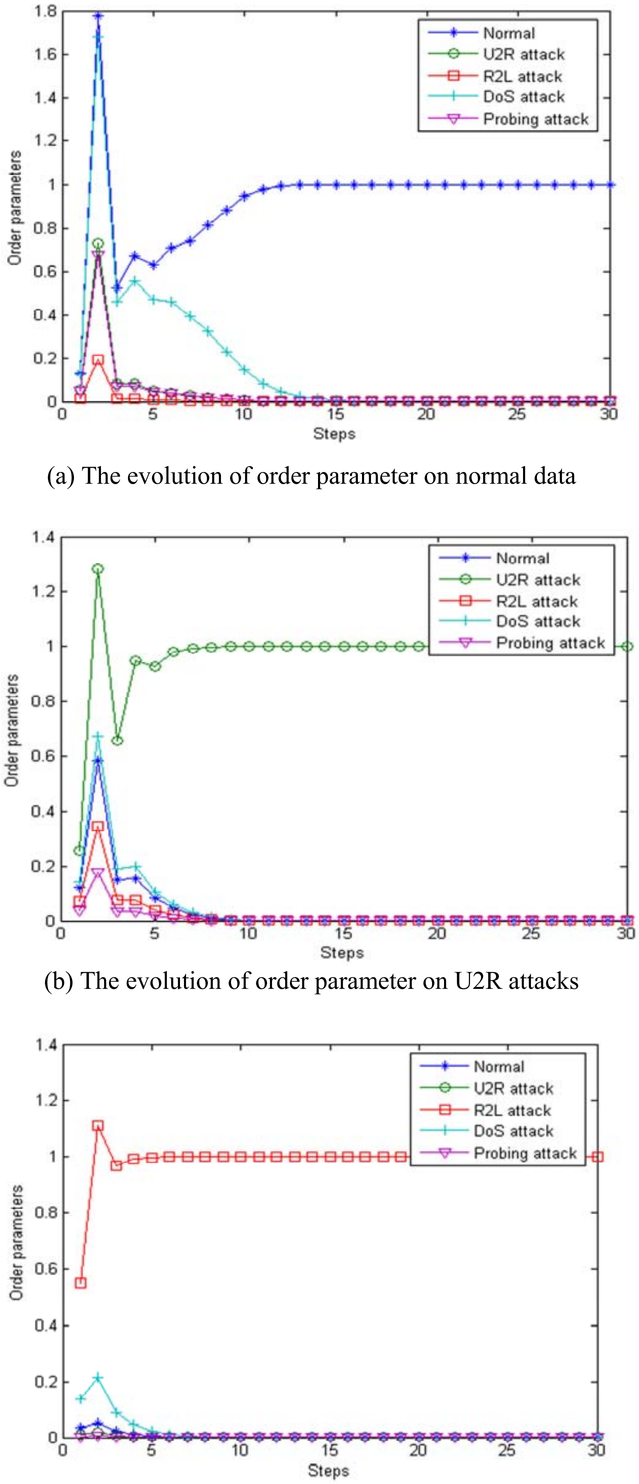
In our experiments, each point of the network traffic was detected as one of the five patterns (including four attack patterns and one normal pattern) based on the classifying method using our SNN model.

The given constants B and C were chosen as $B = C = 1$, which are greater than 0 to guarantee the convergence of SNN. The attention parameter λ_k is also initialized as 1 to make the same attentions with five patterns.

Figures 2(a-e) show the evolving curves of the order parameters of the network traffic data of the normal, U2R attack, R2L attack, Probing attack and Dos attack, respectively. It can be seen that only one order parameter converges to 1 and the others converge to 0. Moreover, the order parameter with the converged value 1 just corresponds to the real prototype pattern of the testing data. Figure 3 shows the detection result of a sub-serial network traffic data on Friday, week 5. Most attacks in this sub-serial are detected as one pattern of the attacks with pattern values of 2 to 5, and the normal traffic data are corresponding to the normal pattern with the pattern value of 1.

We summarize the detection results of the weeks 4 and 5 in Table 1. The *DR* and *FAR* of a continuous sub-serial whose length are longer than 3000 have been computed for each day. The best detection result obtains on Friday in week 5, where the *DR* is 97% and the *FAR* is 8.9%. The mean values of *DR* and *FAR* of all the days are 83% and 8.3%, respectively.

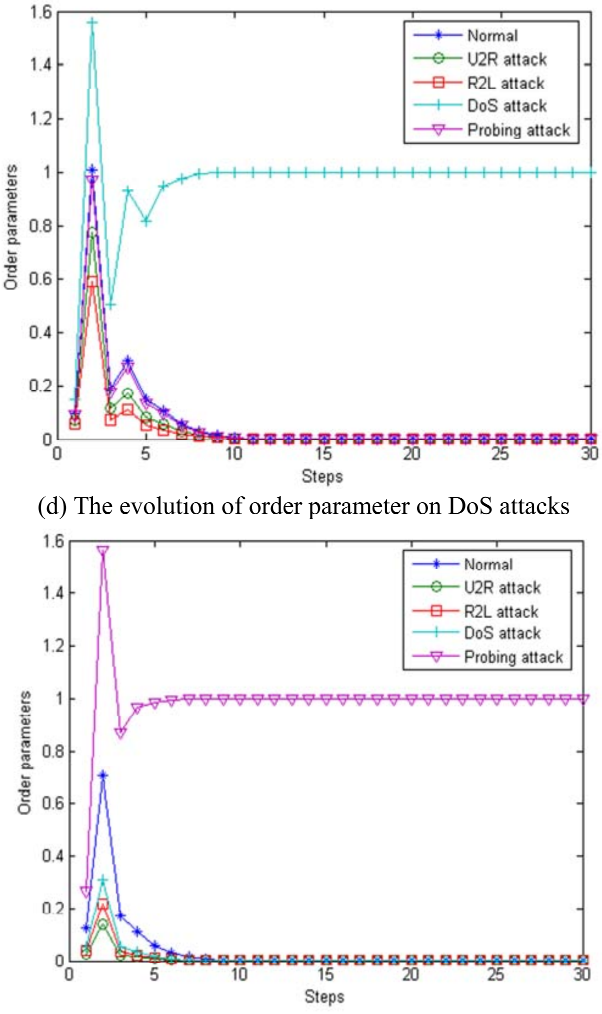
For further validating our method, we make a contrast with a traditional statistical physics method using the auto-correlation function (ACF)[20]. We compare its results with the ACF on the same detection data. The *DR* and *FAR* of the two methods of each day in the week 4 and 5 are shown in Figure 4 and Figure 5. The mean *DR* of SNN method is increased 22% than that of ACF. Simultaneously, the mean



(a) The evolution of order parameter on normal data

(b) The evolution of order parameter on U2R attacks

(c) The evolution of order parameter on R2L attacks



(d) The evolution of order parameter on DoS attacks

(e) The evolution of order parameter on Probing attacks

Figure 2. The evolution results of order parameters.
(a) The evolution of order parameter on normal data,
(b) The evolution of order parameter on U2R attacks,
(c) The evolution of order parameter on R2L attacks,
(d) The evolution of order parameter on DoS attacks,
(e) The evolution of order parameter on Probing attacks.

Table 1.The results of the anomaly detection based on SNN

Day	*DR*	*FAR*
Monday, week 4	78.00%	9.30%
Tuesday, week 4	76.00%	8.21%
Wednesday, week 4	73.21%	6.70%
Thursday, week 4	83.40%	10.35%
Friday, week 4	82.35%	6.25%
Monday, week 5	93.33%	13.64%
Tuesday, week 5	79.53%	9.76%
Wednesday, week 5	81.72%	10.04%
Thursday, week 5	84.21%	7.69%
Friday, week 5	97.30%	8.90%

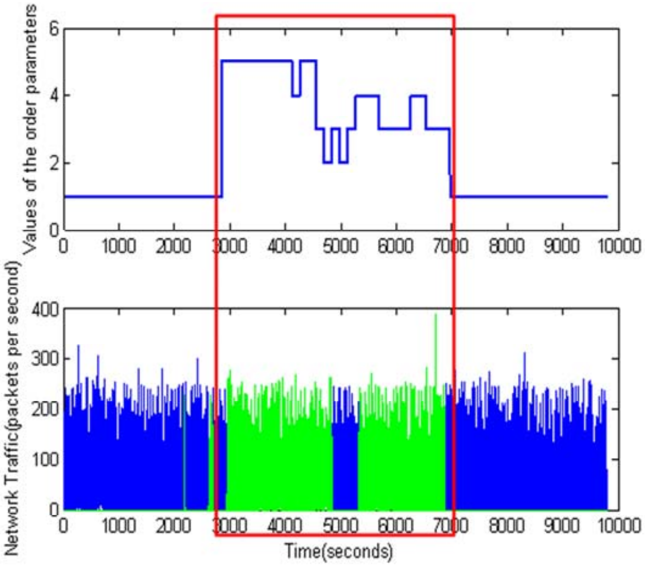


Figure 3. The detection results of the network traffic anomaly based on SNN. (Top): the detection values by involving the order parameters –Friday, the fifth week (with attacks) (Bottom): the network traffic with attacks (the blue parts represent the network traffic with no attacks, the green parts represent the network traffic with attacks) –Friday, fifth week. The rectangle identifies the time period there were attacks happened.

FAR of SNN method is decreased 1% than that of ACF. In other words, SNN method improved the *DR* greatly and maintained the FAR in a low level either.

IV. Conclusions

Recently Network traffic anomaly detection has become a popular research tendency, as it can detect new attack types s in real time. The real-time network traffic anomaly detection is still an unsolved problem of network security. The network traffic appears as a complex dynamic system, precipitated by many network factors. Although various schemes have been proposed to detect anomalies, they are mostly based on traditional statistical physics. In these methods, all factors are integrated to analyze the variation of the network traffic. But in fact, the changing trend of network traffic at some moment is only determined by a few primary factors. In this paper, a new non-statistical network traffic anomaly detection method based on SNN to detect the network traffic anomaly is presented. By introducing a synergetic dynamic equation based on the order parameters, the complex system of the network traffic has been described exactly. When the synergetic dynamic equation is evolved, only the order parameter determined by the primary factors can converge to 1. Therefore, the network traffic anomaly can be detected by referring to the primary factors. We evaluate our approach using the intrusion evaluation data set of the network traffic provided by the defence advanced research projects agency (DARPA). Furthermore, the experiment results on the DARPA intrusion detection evaluation data set indicate that our approach can effectively detect the network anomaly and achieve high detection probability and low false alarms rate. For the future, we will try to carry out this method in engineering applications.

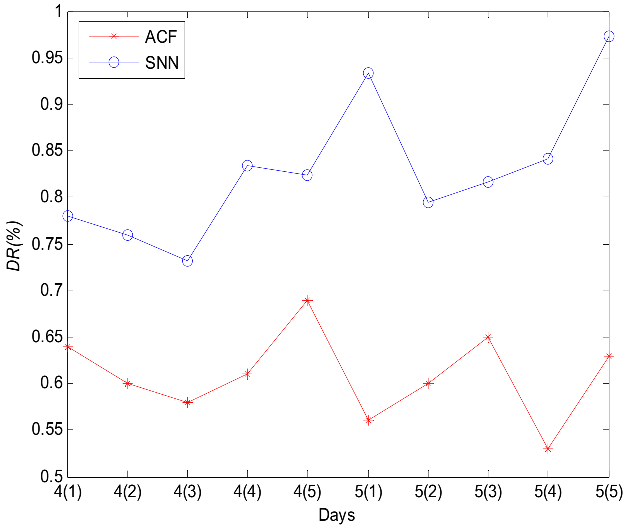


Figure 4. *DR* of two methods of each day in weeks 4 and 5.

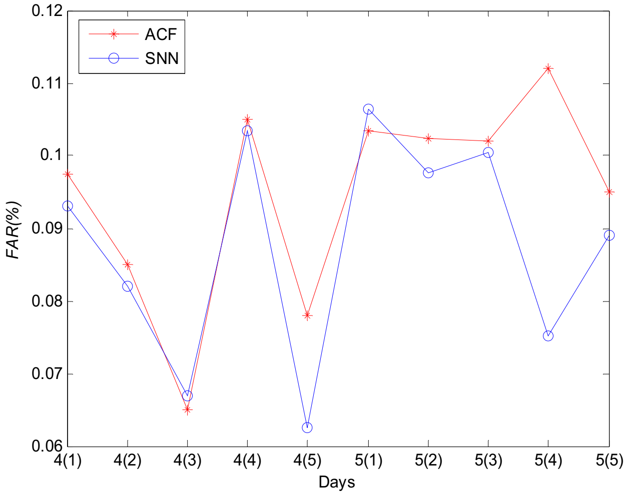


Figure 5. *FAR* of two methods of each day in weeks 4 and 5.

ACKNOWLEDGMENT

This research has been supported by Natural Science Foundation of China (No: 61162009 and No: 60803105), and partly supported by the China National Science Foundation (NSF): Adative and Dynamic Recovery Strategies for Real-time Databases in Embedded and Mobile Computing Environment (ID: 60863016).

REFERENCES

[1] M. Roesch, “Snort–lightweight intrusion detection for networks.” in Proceeding of the 13th USENIX conference on System administration, USENIX Association Berkeley, CA, USA, 1999.

[2] V. Paxson, “Bro: A system for detecting network intruders in real-time,” Comput. Networks, vol. 31, no. 23, pp. 2435-2463, 1999.

[3] P. Gogoi, B. Borah, and D. Bhattacharyya, “Anomaly Detection Analysis of Intrusion Data using Supervised & Unsupervised Approach,” Journal of Convergence Information Technology, vol. 5, no. 1, pp. 95-110, 2010.

[4] C. Sastry, S. Rawat, A. Pujari et al., “Network traffic analysis using singular value decomposition and multiscale transforms,” Information Sciences, vol. 177, no. 23, pp. 5275-5291, 2007.

[5] M. Li, “Change trend of averaged Hurst parameter of traffic under DDOS flood attacks,” Computers & Security, vol. 25, no. 3, pp. 213-220, 2006.

[6] A. Ziviani, A. Gomes, M. Monsores et al., “Network anomaly detection using nonextensive entropy,” IEEE Communications Letters, vol. 11, no. 12, pp. 1034-1036, 2007.

[7] Y. Gu, A. McCallum, and D. Towsley, “Detecting anomalies in network traffic using maximum entropy estimation.” in Proceeding of the 5th ACM SIGCOMM conference on Internet Measurement (IMC05), USENIX Association Berkeley, CA, USA, 2005.

[8] S. Janakiraman, and V. Vasudevan, “ACO based distributed intrusion detection system,” JDCTA: International Journal of Digital Content Technology and its Applications, vol. 3, no. 1, pp. 66-72, 2009.

[9] S. Lee, and D. Heinbuch, “Training a neural-network based intrusion detector to recognizenovel attacks,” IEEE Transactions on Systems, Man and Cybernetics, Part A, vol. 31, no. 4, pp. 294-299, 2001.

[10] Y. Qiao, X. Xin, Y. Bin et al., “Anomaly intrusion detection method based on HMM,” Electronics Letters, vol. 38, no. 13, pp. 663-664, 2002.

[11] T. Ryutov, C. Neuman, K. Dongho et al., “Integrated access control and intrusion detection for web servers,” IEEE transactions on parallel and distributed systems, vol. 14, no. 9, pp. 841-850, 2003.

[12] C. Nelson, and D. Fitzgerald, "Sensor fusion for intelligent alarm analysis." in Proceeding of the 30th Annual International Carnahan Conference on Security Technology, Lexington, USA, pp. 143-150, Oct 1996.

[13] T. Shon, and J. Moon, “A hybrid machine learning approach to network anomaly detection,” Information Sciences, vol. 177, no. 18, pp. 3799-3821, 2007.

[14] V. Frost, and B. Melamed, “Traffic modeling for telecommunications networks,” IEEE Communications Magazine, vol. 32, no. 3, pp. 70-81, 1994.

[15] K. Chandra, C. You, G. Olowoyeye et al., “Non-Linear Time-Series Models of Ethernet Traffic,” Technology Report, 1999.

[16] L. Amaral, and J. Ottino, “Complex networks,” The European Physical Journal B-Condensed Matter and Complex Systems, vol. 38, no. 2, pp. 147-162, 2004.

[17] A. Adas, “Traffic models in broadband networks,” IEEE Communications Magazine, vol. 35, no. 7, pp. 82-89, 1997.

[18] B. Chen, Y. Yang, B. Lee et al., “Fuzzy adaptive predictive flow control of ATM network traffic,” IEEE transactions on Fuzzy Systems, vol. 11, no. 4, pp. 568-581, 2003.

[19] H. Haken, “Synergetic computers and cognition: a top-down approach to neural nets,” Book by Springer-verlag Berlin Heidelberg, 2004.

[20] N. Xiong, X. Jia, L. T. Yang, A. V. Vasilakos, Y. Li, Y. Pan, "A Distributed Efficient Flow Control Scheme for Multirate Multicast Networks," IEEE Trans. Parallel Distrib. Syst. 21(9): 1254-126, 2010.

[21] W. Xiong, H. ping, and Y. Yue, "Anomaly detection of network traffic based on autocorrelation principle," Journal of Communication and Computer, vol. 4, no. 008, pp. 15-19, 2007.

[22] N. Xiong, A. V. Vasilakos, L. T. Yang, L. Song, Y. Pan, R. Kannan, Y. Li, "Comparative analysis of quality of service and memory usage for adaptive failure detectors in healthcare systems," IEEE Journal on Selected Areas in Communications 27(4): 495-509, 2009.