Network Traffic Anomaly Detection based on Catastrophe Theory

Wei Xiong[1,2], NaixueXiong[3], Laurence T. Yang[4], Athanasios V. Vasilakos[5], Qian Wang[6,] Hanping Hu [1]

[1]From Institute of Pattern Recognition & AI, Huazhong University of Science and Technology, P.R.China.

[2]Center of Computing & Experimenting, the South Central University for Nationalities, P.R.China.

[3]Department of Computer Science, Georgia State University, Atlanta, USA

[4]St. Francis Xavier University, Canada

[5]Dept. of Computer and Telecommunications Engineering, Univ. of Western Macedonia, Greece

[6]Information school, Zhongnan University of Economic & Law, P.R.China

Email: ccnuxw@sina.com, nxiong@cs.gsu.edu, vasilako@ath.forthnet.gr, lyang@stfx.ca,
icedlitchi@163.com, hphu@mail.hust.edu.cn

Abstract-Although various methods have been proposed to detect anomalies, they are mostly based on the traditional statistical physics. The traditional statistical physics methods are based on the stationary hypothesis of the network traffic, which always ignore the real catastrophe process when anomalies occur. In order to reflect the catastrophe process of the abnormal network traffic, we present a non-stationary network traffic anomaly detection approach based on catastrophe theory. The cusp catastrophe model is selected to describe the catastrophe feature of the network traffic and the catastrophe distance is defined as an index to assess the deviation from the normal catastrophe model and the serial of catastrophe distance is the main feature to detect anomaly. We evaluate our approach using the 1999 intrusion evaluation data set of network traffic trace provided by The Defense Advanced Research Projects Agency (DARPA). Experiment results show that our approach can effectively detect network anomalies and achieve high detection probability and low false alarms rate.

***Index Terms*-Anomaly detection, Network traffic, Cusp Catastrophe Model, Catastrophe Distance**

I. INTRODUCTION

It is well-known that the Internet and computer networks are exposed to an increasing amount of security threats. With huge new types of attacks appearing frequently, developing flexible and real-time security oriented approaches is a severe challenge.

At present, the reported approaches of network traffic anomaly detection mostly adopted the statistical physics strategies. In these methods, the macro features of network traffic, such as self-similarity[1,2],entropy[3,4], probability distribution[5,6,7,8] et al, were extracted, followed by using various pattern recognition techniques, such as neural networks[9], hidden Markov model[10], integrated access control[11], and sensor fusion [12], machine learning[13] ,to detect network anomaly.

However, the generation of network traffic is a complex process which driven by many factors, such as network devices, topology, transfer protocol, as well as the interactive cooperation and competition among the network users. Thus, network traffic often shows non-linear, non-stationary and complex nature characteristics [14, 15, 16, 17], which is a complex dynamic system. When attacks occur, the network traffic system will transform from the normal equilibrium to the abnormal equilibrium. This transformed process is catastrophic, but not stationary. The traditional statistical physics methods are based on the stationary hypothesis of the network traffic, which always ignore the real catastrophe process when attacks occur. Thus the accuracy of the real-time detection of these methods is influenced.

In order to reflect the catastrophe process of the network traffic, we present a non-stationary network traffic anomaly detection approach based on catastrophe theory. In this method, a catastrophe potential function is introduced to depict the catastrophe process of the network traffic. The catastrophe distance is proposed to detect the network traffic anomaly. To evaluate the performance of our approach, we run

[1] This work is supported by the National Natural Science Foundation of China (No. 60773192).

In this paper, the corresponding author is Hanping Hu.

this approach on the standard Defense Advanced Research Projects Agency (DARPA) data sets and compare the result with a reported statistical physics method [18] The results show that our approach based on catastrophe theory is effective to detect network anomaly.

The rest of the paper is organized as follows. In Section II we introduce the catastrophe theory and its properties. In Section III we introduce the catastrophe feature of network traffic. In Section IV, we expand our anomaly detection method based on catastrophe. In Section V we show some experiment results and report the performance of our method. At last, we make a conclusion in Section VI

II. . Catastrophe theory

Catastrophe theory is a mathematical model created by French Rene Thom in 1972 and is a branch of bifurcation theory in the study of dynamical systems; it is also a particular special case of more general singularity theory in geometry. Bifurcation theory studies and classifies phenomena characterized by sudden shifts in behavior arising from small changes in circumstances, analyzing how the qualitative nature of equation solutions depends on the parameters that appear in the equation. The model has the ability to produce sudden changes in a variable where a gradual change has been observed and is therefore expected. These sudden changes are titled "catastrophes". The catastrophe theory model possesses five properties that must be satisfied for the model to be considered applicable. These five properties are as follow.

A. Catastrophes

A catastrophe is a sudden change in a parameter when gradual changes were previously experienced and were expected to continue. While using perfect delay convention, there is a jump from disappeared minimum to absolute minimum or partial minimum, and the numerical value of potency varies discontinuously. The numerical value of potency varies continuously, but its derivate is discontinuous.

B. Divergence

A marginal change in path causing major change in behavior is the definition of divergence. The limited variation of control will lead to the variety of state variable in equilibrium. In general case, minute disturbance of control variable maybe result in a great change of last value of state variable. The instability of disturbance of control variable is called divergence.

C. Bimodality

Bimodal behavior means that for certain values of the control variables there are at least two stable values for the state variable. There could be two or more different states in system. That is to say, the potency of system in certain range of control variable may have two minimums.

D. Inaccessibility

The middle fold in the catastrophe surface and the vertical slice catastrophe both represent an inaccessible surface. Values that occur on these surfaces are not stable and quickly move to a more stable point on the upper or lower surfaces. There is an unstable equilibrium location, where it is may either be continuous or discontinuous. It is not differentiable in mathematics.

E. Hysteresis

The hysteresis phenomenon in catastrophe theory is that the catastrophes to the congested regime do not necessarily occur at the same values of the variables as the catastrophes from the congested regime. In a catastrophe model using the perfect delay convention, this phenomenon is a necessary property.

III. The catastrophe feature of network traffic

The catastrophe characteristics of a system are dependent on whether it has the catastrophe properties. The motions of the network traffic system depend on the transformations among equilibriums. For instance, in the normal network traffic (when the state of the network traffic is normal, that is to say, there are no anomaly happened, we call the network traffic as the normal network traffic.), the network traffic system maintains the stationary changing tendency where the network state is called as the normal equilibrium. After anomalies have occurred, the network traffic system maintains stationary abnormal state where the network state is called as the abnormal equilibrium. When anomalies occur, the network state will transform from the normal equilibrium to the abnormal equilibrium. The change of network traffic state is a catastrophe process. A volume of network traffic is versus two network traffic state values (a normal volume value and an abnormal volume value), which is accordance with the bimodality properties. In cases, the network traffic state is whether in a normal state or in an abnormal state. The catastrophe properties of network traffic system can also be shown in Fig.1. From the plot of the observed data, we could also find that there are few pints or regions.

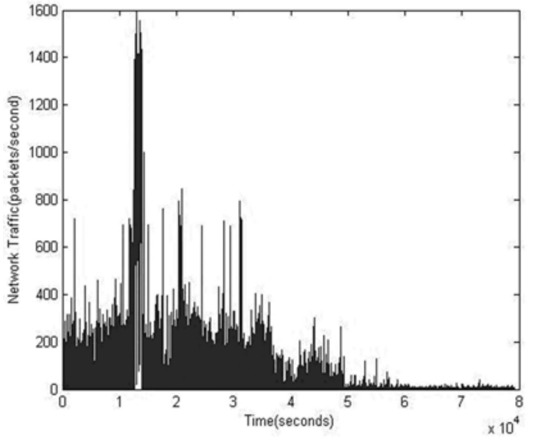

Fig.1. The sudden jump of the network traffic

Considering from network users psychology and economy, a time benefit makes driver not linger in these regions or web sites in order to reduce delay and operation time as possible (unless there are hot news or sole owner information). Thus, there exists inaccessible behavior. Besides, when the network traffic system reaches its capacity, it means a state of critical equilibrium. But the equilibrium state is an ideal one and it exists only in an instant equilibrium state that would be destroyed once it is disturbed by outside factors. The transformation of equilibriums is not a gradual process but a sudden jump. This process may be regarded to be unstable. The hysteresis phenomenon occurs only when using perfect delay convention.

As stated above, network traffic exhibits catastrophic behavior. It is feasible to explain traffic operation by using catastrophic theory. When a system has catastrophe properties as mentioned above, catastrophe model can be established to describe it.

Catastrophe theory analyses degenerate critical points of the potential function — points where not just the first derivative, but one or more higher derivatives of the potential function are also zero. These are called the germs of the catastrophe geometries. The degeneracy of these critical points can be unfolded by expanding the potential function as a Taylor series in small perturbations of the parameters.

When the degenerate points are not merely accidental, but are structurally stable, the degenerate points exist as organizing centers for particular geometric structures of lower degeneracy, with critical features in the parameter space around them. If the potential function depends on two or fewer active variables, and four or fewer active parameters, then there are only seven generic structures for these bifurcation geometries, with corresponding standard forms into which the Taylor series around the catastrophe germs can be transformed by diffeomorphism. Three common catastrophe types are, respectively, cusp catastrophe, swallowtail catastrophe and butterfly catastrophe. Moreover, the cusp catastrophe model is the most common model and our paper uses the cusp catastrophe model to describe the network traffic anomalies. Their potential functions are shown in table 1.

Table 1 The three common catastrophe types. Where *F(x)* is the potential function, *x* is the state variable and *a*, *b*, *c* and *d* are control variables.

Catastrophe type	The potential function
Cusp catastrophe	$F(x)=x^4+ax^2+bx$
Swallowtail catastrophe	$F(x)=x^5+ax^3+bx^2+cx$
Butterfly catastrophe	$F(x)=x^6+ax^4+bx^3+cx^2+d.$

IV. NETWORK TRAFFIC ANOMALY DETECTION BASED ON CATASTROPHE THEORY

In this paper, the cusp catastrophe model is selected to detect anomalies of network traffic. The potential function $F(x)$ of cusp catastrophe model is as follows:

$$F(x)=x^4+ax^2+bx$$

Where x is state variable; a , b is the control variables.

The critical point set of potential function $F(x)$ compose a balance surface. By seeking a derivative of $F(x)$ and making $F'(x)=0$, the balance surface equation can be gotten:

$$4x^3+2ax+b=0$$

Making $F''(x)=0$, the singularity set of the cusp catastrophe is gotten:

$$6x^2+a=0$$

Though the equations $F'(x)=0$ and $F''(x)=0$, the difference set $G(a,b)$ of cusp catastrophe model is as follows:

$$8a^3+27b^2=0$$

The difference set of cusp catastrophe model, which can reflect the relationship between each control variables and the state variable presented by the state variable, is the most import because the difference set is in the control space that can be observed and in which all the sudden jumps will happen. Fig.2 shows the basic form of the cusp catastrophe model. The top surface is the equilibrium surface of cusp catastrophe model which is divided into upper sheet A and lower sheet C. When the state of system transfers from the stable equilibrium state B to another stable equilibrium state D, there is a sudden jump between the stable state B and D and the catastrophe phenomena appear. The bottom one is the control space represented by the control variables.

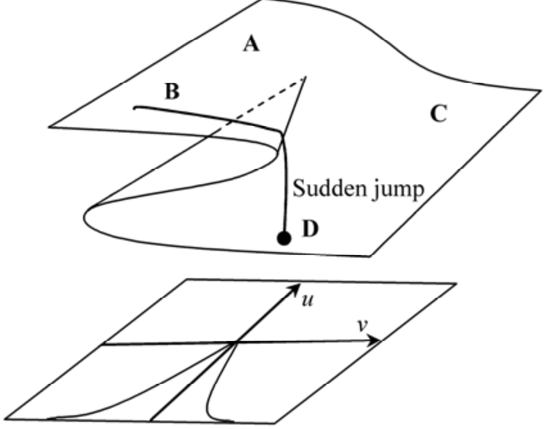

Fig.2. A cusp catastrophe model

The steps of the network traffic anomaly detection based on catastrophe theory are as follows.

(1)Considering the time series $X_1,...,X_n$ of the training data, for each time t , construct the vector set

$\{ \ Y_t^w = (X_{t-w+1},...,X_t) \mid t = 1,...,n-w+1 \ \}$ with the time window Win_p .

(2)Obtain the series of the state variable $\{ x_t \}$ and the control variables $\{ u_t \}$ and $\{ v_t \}$ based on normalized features extracted from each vector Y_t^w .

(3)Compute the parameters a, b of the cusp catastrophe model using the series $\{ x_t \}$, $\{ u_t \}$ and $\{ v_t \}$.

In the testing stage, the main steps are as following.

(1)Construct the vector Y_i^w (with the same time window Win_p in the training stage) of the testing data at the observed time i , which is labeled as observed point P_i .

(2)Extract the selected normalized features to present the state variable x_i and control variables u_i , v_i .

(3)Compute the catastrophe distance between the observed point P_i and the bifurcation set $G(u,v)$, labeled as d_p . When the catastrophe distance d_p is beyond a given threshold η , the observing point P_i can be argued that there is an anomaly existed.

V. Experimental result

The DARPA Intrusion Detection Evaluation has been carried out in 1998 and 1999 by MIT Lincoln Laboratory and sponsored by Defense Advanced Research Projects Agency (DARPA). The purpose of its evaluation is to contribute significantly to the intrusion detection research field by providing direction for research efforts and an objective calibration of the current technical state-of-the-art. The taxonomy of attacks used for DARPA evaluation data set is characterized as follows. (1)DoS: Denial of Service, (2) PROBE: Surveillance/Probing, (3) U2R: User to Super user (root), (4) R2L: Remote to Local user.

In this work we focus on the 1999 evaluation data set. With the tcpdump tool, we extract the aggregated network traffic in bits/bytes/packets per second from the data files to evaluate our approach of network traffic detection. Three weeks of training data were provided for the 1999 DARPA Intrusion Detection off-line evaluation. The first and third weeks of the training data do not contain any attacks. This data was provided to facilitate the training of anomaly detection systems. The second week of the training data contains a select subset of attacks from the 1998 evaluation in addition to several new attacks. In addition, the network traffic data of the forth and fifth weeks where attacks are mixed in the normal background data were also provided as the test data.

We process the network traffic data using the anomaly detection algorithm based on the catastrophe theory. The data of each day of the first and third weeks were used to construct the control variables for the cusp catastrophe model. We then observed the traffic of each day of the forth and fifth weeks and we investigated how this deviates from the reference traffic of the similar day of the first and third weeks. To detect the attacks we use the K-nearest algorithm [19].According the K-nearest algorithm, we make the serial of catastrophe distance as the input data. Keeping the near two numbers of the series as a vector with no repetition, we implement the anomaly detection algorithm while the betterment is that the vector which is be considered as normal will replace the oldest training vector in order to make the whole vector aggregate not be larger. The anomaly detection of catastrophe distance of the second day of fifth week is shown in Fig 3.

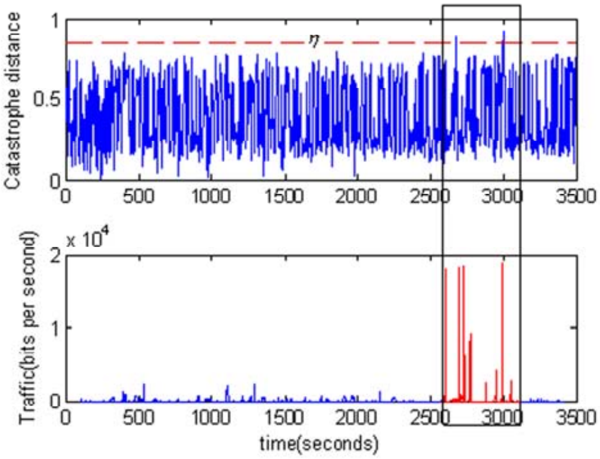

Fig.3. the anomaly detection of catastrophe distance

To verify the performance of our anomaly detection approach, we quote the detection rate of abnormal events (*DR*) and the rate of the fault point (*FAR*) to quantitatively evaluate the detected results. Their definitions are as following

$$DR = \frac{NA_{\text{detected}}}{NA_{all}} \ , \ FAR = \frac{NAF_{\text{detected}}}{NN_{all}} \ .$$

Where NA_{detected} is the number of the attack points having been detected correctly, NA_{all} is the number of all attack points, NAF_{detected} is the number of the normal points having been detected falsely as attacks, NN_{all} is the number of all points to be detected.

We summarize the detection results of the week 4 and 5 in Fig 4 and Fig 5. The *DR* and *FAR* have been computed for each day. The best detection result obtains in the Friday in week 5, where the *DR* is 96% and the *FAR* is 11.42%. The mean *DR* and *FAR* of all the days are 86.62% and 9.06%, respectively.

In our earlier work, we studied a traditional statistical physics method using the auto-correlation function (ACF)[18] to detect the network traffic anomaly. To evaluate the new method in this paper more thoroughly, we compare its results with the traditional statistical physics method 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 Fig 4 and Fig 5. The mean *DR* of the method based on catastrophe theory is increased 25.72% than that of ACF. Simultaneously, the mean *FAR* of the method based on catastrophe theory is decreased 0.4% than that of ACF. In other words, the method based on catastrophe theory improved the *DR* greatly and maintained the *FAR* in a low level either.

VI. CONCLUSION

In this paper, we present a new network traffic anomaly detection approach based on catastrophe theory. We evaluate our approach using the 1999 DARPA data set. Experiment results show that our approach is effective to detect network anomalies and achieve high detection probability and low false alarms rate.

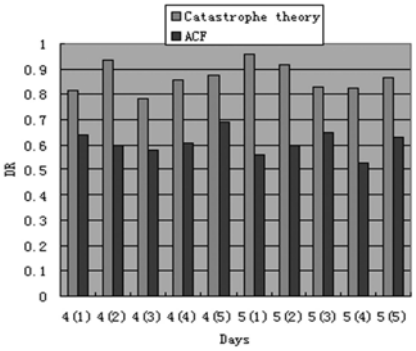

Fig.4. The *DR* of the two methods of each day in the week 4 and 5

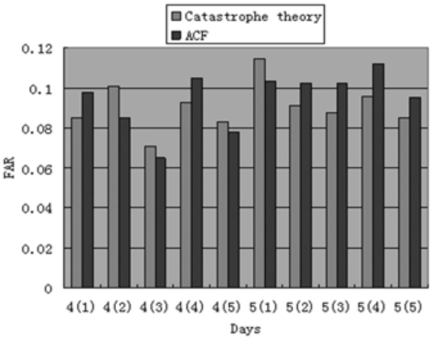

Fig.5. The *FAR* of the two methods of each day in the week 4 and 5

REFERENCES

[1]C. Sastry, S. Rawat, A. Pujari and V. Gulati, *Network traffic analysis using singular value decomposition and multiscale transforms*, Information Sciences **177** (2007), no. 23, 5275-5291.

[2] M. Li, *Change trend of averaged hurst parameter of traffic under ddos flood attacks*, Computers & Security **25** (2006), no. 3, 213-220.

[3]A. Ziviani, A. Gomes, M. Monsores and P. Rodrigues, *Network anomaly detection using nonextensive entropy*, IEEE Communications Letters **11** (2007), no. 12, 1034-1036.

[4]Y. Gu, A. McCallum, and D. Towsley, *Detecting anomalies in network traffic using maximum entropy estimation*, in Proc. ACM/SIGCOMM Internet Measurement Conference – IMC 2005, Oct. 2005.

[5]I. Paschalidis and G. Smaragdakis, *Spatio-temporal network anomaly detection by assessing deviations of empirical measures*, IEEE/ACM Trans. Networking.

[6]M. Burgess, *Probabilistic anomaly detection in distributed computer networks*, Science of Computer Programming **60** (2006), no. 1, 1-26.

[7]L. Zhang, Z. Zhu, K. Jeffay, J. Marron and F. Smith, *Multi-resolution anomaly detection for the internet*, 2008, p.^pp. 1-6.

[8]A. Lakhina, M. Crovella and C. Diot, *Mining anomalies using traffic feature distributions*, ACM New York, NY, USA, 2005, p.^pp. 217-228

[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 **31** (2001), no. 4, 294-299.

[10]Y. Qiao, X. Xin, Y. Bin and S. Ge, *Anomaly intrusion detection method based on hmm*, Electronics Letters **38** (2002), no. 13, 663-664.

[11]T. Ryutov, C. Neuman, K. Dongho and Z. Li, *Integrated access control and intrusion detection for web servers*, IEEE transactions on parallel and distributed systems **14** (2003), no. 9, 841-850.

[12]C. Nelson and D. Fitzgerald, *Sensor fusion for intelligent alarm analysis*, 1996, p.^pp. 143-150.

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

[14]V. Frost and B. Melamed, *Traffic modeling for telecommunications networks*, IEEE Communications Magazine **32** (1994), no. 3, 70-81.

[15]K. Chandra, C. You, G. Olowoyeye and C. Thompson, *Non-linear time-series models of ethernet traffic*, Submitted to INFOCOM '99 (1998).

[16]L. Amaral and J. Ottino, *Complex networks*, The European Physical Journal B-Condensed Matter and Complex Systems **38** (2004), no. 2, 147-162.

[17]A. Adas, *Traffic models in broadband networks*, IEEE Communications Magazine **35** (1997), no. 7, 82-89.

[18]Xiong W., H. ping, and Y. Yue. Anomaly detection of network traffic based on autocorrelation principle, Journal of Communication and Computer.2007,4(008): 15-19.

[19]N.B. Waite, A real-time system-adapted anomaly detector, Information Sciences, Volume 115, Number 1, April 1999 , pp. 221-259(39).