Computers, Materials & Continua CMC, vol.60, no.3, pp.1171-1187, 2019

YATA: Yet Another Proposal for Traffic Analysis and Anomaly Detection

Yu Wang[1, 2, *], Yan Cao[2], Liancheng Zhang[2], Hongtao Zhang[3], Roxana Ohriniuc[4], Guodong Wang[5] and Ruosi Cheng[6]

Abstract: Network traffic anomaly detection has gained considerable attention over the years in many areas of great importance. Traditional methods used for detecting anomalies produce quantitative results derived from multi-source information. This makes it difficult for administrators to comprehend and deal with the underlying situations. This study proposes another method to yet determine traffic anomaly (YATA), based on the cloud model. YATA adopts forward and backward cloud transformation algorithms to fuse the quantitative value of acquisitions into the qualitative concept of anomaly degree. This method achieves rapid and direct perspective of network traffic. Experimental results with standard dataset indicate that using the proposed method to detect attacking traffic could meet preferable and expected requirements.

Keywords: Anomaly detection, cloud model, forward cloud transformation, backward cloud transformation, quantitative data to qualitative concept.

1 Introduction

Network traffic analysis is one of the most interesting topics in the research of basic theory of computer network. With the rapid development of Internet services and improvement of network performance, network threats are becoming more and more significant. Many kinds of anomaly events are mixed with normal traffic, especially viruses/Trojans, Botnet, XSS/CSRF, DoS and other attacks emerge in an endlessly stream. In face of the traditional and newborn threats, traffic analysis and anomaly detection technology is facing severe challenges, notably on how to effectively identify and perceive potential unknown attacks [Cheang, Wang, Cai et al. (2018); Gokcesu and Kozat (2017); Zhao, Luo, Gan et al. (2018); Zu, Luo, Liu et al. (2018)].

Anomaly detection has two main components: the model, and the algorithm. Existing

[1] Henan University of Engineering, Xinzheng, 451191, China.

[2] China National Digital Switching System Engineering & Technological R&D Center, Zhengzhou, 450002, China.

[3] Zhengzhou University, Zhengzhou, 450001, China.

[4] Florida Atlantic University, Boca Raton, FL 33341, USA.

[5] Massachusetts College of Liberal Arts, North Adams, MA 01247, USA

[6] Luoyang Electronic Equipment Test Center of China, Luoyang, 471003, China.

[*] Corresponding Author: Yu Wang. Email: stonchor@163.com.

detection methods mainly include signal processing and machine learning based technologies. The former includes statistical analysis [Hu, Xiao, Fu et al. (2006)], frequency spectrum analysis [Chen, Wang, Zhao et al. (2011); Ningrinla, Amar and Kumar (2018)], wavelet analysis [Sun and Tian (2014)] and principal component analysis [Xie, Li, Wang et al. (2018)]. The latter's representative methods cover data mining [Sukhanov, Kovalev and Stýskala (2016)], neural network [Naseer, Saleem, Khalid et al. (2018)] and immunology theory [Jiang, Ling, Chan et al. (2012)]. With the increase number of network applications, the overall complexity of network traffic characteristics has sharply raised. The key problem is with the subjective judgment differences between collected data and the analyzer perceptions.

Traditionally, the major problem in network traffic analysis [Li, Sun, Hu et al. (2018); Zhou, Wang, Ren et al. (2018)] is the subjectivity of the analyzer when collecting and analyzing data, many and multidisciplinary experts are needed to analyze traffic anomalies and characteristics, and the final network analysis result is created by a team of experts not just individuals. The experts would like to utilize the threat ratings or levels [Treurniet (2011); Yao, Shu, Cheng et al. (2017)] to characterize the degree of potential damage that anomalous network traffic can bring. It is rated according to a fix number ordinal qualitative scale, like 5, or 10. The maximum number is considered to represent the most frequent and severe threat.

Existing models have been considered as suffering from a lot of drawbacks. First, the threat ratings or levels are subjective and linguistic in nature, which could not be determined precisely using a scale from 1 to 5, or 1 to 10. Second, in view of different backgrounds and understanding levels to the identified abnormal traffic, experts usually have different perceptions and interpretations, hence, the same linguistic grade level always has different meanings from different experts.

The Cloud Model [Li, Liu and Gan (2009); Wang, Xu and Li (2014)] proposed by Li Deyi has been proved to achieve bidirectional cognitive transformation between the qualitative concept and quantitative data. Because of its advantages, the Cloud Model has been widely used in unsupervised communities detection [Gao, Jiang, Zhang et al. (2013)], failure mode and effect analysis (FMEA) [Liu, Li, Song et al. (2017)], monocular visual odometry (MVO) [Yang, Jiang, Wang et al. (2017)] and other fields [Liu, Xue, Li et al. (2017); Peng and Wang (2018); Yang, Yan, Peng et al. (2014)]. In this paper, we propose an anomaly traffic analysis method based on the cloud model. The cloud model theory is introduced to depict multiple assessment information given by experts who utilize qualitative concepts to describe the key characteristics of network traffic, so as to establish the gauge cloud. Based on that, we generate abnormal membership cloud and abnormal matrix for the traffic to be determined, in order to analyze and judge the underlying situation of network traffic.

2 Cloud model

The Cloud model proposed is based on probability theory and fuzzy mathematics, which can not only reflect the uncertainty of the concept of natural language, but also can depict the event relationship between the randomness and fuzziness. This model is implemented to make transformation between the qualitative concept and quantitative instantiation.

The definition of the cloud is as follows.

2.1 Definitions

Suppose U is a quantitative domain described by numerical values, and C is a qualitative concept with U. Let x be a random instantiation of C, as well as $x \in U$. If $\mu(x)$ represents the certainty degree for C, i.e., $\mu(x) \in [0,1]$, and satisfy:

$$\mu : U \to [0,1], \forall x \in U,\ x \to \mu(x) \tag{1}$$

Define the distribution of x on U to be a cloud, denoted as C(X), in which each x is called a cloud drop.

A cloud drop is one of the instantiation of qualitative concept C in the form of numerical value, and there is no orderly relationship between them. However, the overall status of cloud drops will reflect the characteristics of the qualitative concept. The certainty degree of cloud drop indicates the extent of characterizing this qualitative concept, the greater occurrence probability of cloud drop, the higher certainty degree.

The cloud model describes the overall qualitative concept by three numerical characteristics including:

• Ex (Expectation). It is the expectation of the cloud drops, which is considered to be the most representative and typical sample of concept C.

• En (Entropy). It is used to represent the uncertainty measurement of concept C. On the one hand, En indicates the dispersing extent of cloud drops which measures the degree of randomness; on the other hand, En reflects the range of the universe that concept C can accept, which measures the degree of fuzziness.

• He (Hyper Entropy). It is the entropy of En, namely the uncertain degree of En, reflecting the degree of condensation of cloud drops, which is expressed as the dispersion and thickness of cloud.

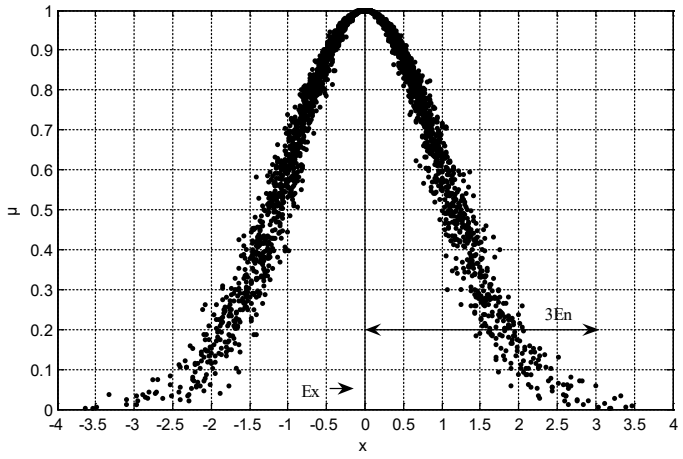

Figure 1: The normal cloud model

As shown in Fig. 1, the normal cloud is a cloud model that has been frequently used by researchers, in which x is one of the cloud drops which randomly realizes the concept C, and μ is the certainty degree of x on concept C. The thickness of the normal cloud is uneven demonstrating the randomness and fuzziness. More than this, each cloud model probably owns different degree of distribution and degree of discreteness, reflecting the randomness and fuzziness features of specific cloud model. This indicates the following: the greater the value of En, the wider the distribution range; the greater the value of He, the larger the degree of discrete.

2.2 Cloud generator algorithms

Algorithm 1: Forward Cloud Transformation (FCT). Transform the qualitative concepts into quantitative representations, that is, to use Ex, En and He to generate cloud drops which satisfy the current network situation.

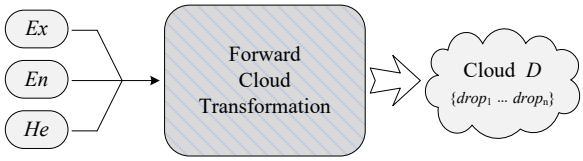

Figure 2: Forward cloud generator

Input: *Ex*, *En*, *He*

Output: Cloud drops

Steps:

1. Take *En* as the expectation, and take *He* as the standard deviation, to produce normal random value |*En*'|.

2. Take *Ex* as the expectation, and take |*En*'| as the standard deviation, to produce normal random value x, which is denoted as a cloud drop within this domain.

3. Based on *Ex* and |*En*'|, μ could be calculated as follows:

$$\mu = e^{-\frac{(x-Ex)^2}{2(En')^2}} \tag{2}$$

Here, μ is defined as the degree of certainty that x belongs to the qualitative concept of C.

4. Iterations from Step 1 to Step 3 are preformed to generate n cloud drops, and every drop could be represented as $drop_i=(x_i, \mu_i)$.

Algorithm 2: Backward Cloud Transformation (BCT). Conduct uncertainty transformation from the quantitative numeric value into qualitative concepts.

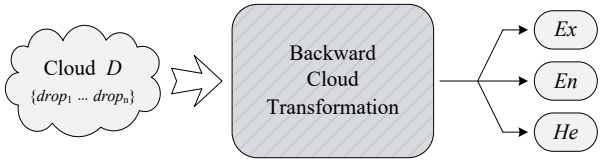

Figure 3: Backward cloud generator

Input: Set of cloud drops with total amount of N, namely D=$\{drop_1 \ldots drop_n\}$.

Output: *Ex*, *En* and *He*, which corresponds to the qualitative concepts represented by the cloud.

Steps:

1. Calculate the mean value of the input value based on X:

$$\overline{X} = \frac{1}{N} \cdot \sum_{i=1}^{N} x_i \tag{3}$$

Also, the first order absolute central moment is:

$$\frac{1}{N} \cdot \sum_{i=1}^{N} \left| x_i - \overline{X} \right| \tag{4}$$

Besides, the sample variance is:

$$S^2 = \frac{1}{N-1} \cdot \sum_{i=1}^{N} (x_i - \overline{X})^2 \tag{5}$$

2. Therefore *Ex*, *En* and *He* could be obtained:

$$Ex = \overline{X} \tag{6}$$

$$En = \sqrt{\pi / 2} \cdot \frac{1}{N} \cdot \sum_{i=1}^{N} \left| x_i - Ex \right| \tag{7}$$

$$He = \sqrt{\left| S^2 - En^2 \right|} \tag{8}$$

By utilizing BCT and FCT, the cloud model could implement reciprocal conversion between qualitative concept and quantitative value based on the interaction between probability theory and fuzzy mathematics.

3 Traffic analysis and anomaly detection model based on cloud model

3.1 Analytical characteristics of traffic

Considering the potential differences among the characteristics of network traffic, the anomalies and normal traffic could be identified accordingly. Under usual circumstances, the anomaly traffic characteristics caused by non-human factors are more significant and easy to detect. Oppositely, the abnormal traffic generated by malicious attackers is obvious in terms of traffic flow rate, average packet length, and the distribution of different protocol packets.

The key characteristics, which mainly reflect the security degree of network traffic, could generally be divided into three levels including strong, weak and neutral according to the given network environment. In the light of above analysis, the characteristics of traffic volume, the average packet length of different applications, the average packet arrival interval, the number of connection requests, the number of port pairs in a single flow, are regarded as especially important. According to the division of anomaly levels, 7 characteristics are selected to depict the security state of target traffic qualitatively.

Table 1: Characteristics of traffic

#	Characteristics	Descriptions
1	Flow rate of traffic	The velocity of the packet access
2	Length of packet	The average length of packets
3	Duration of single connection	The degree of stability of the application
4	Failure connections ratio per unit time	The proportion of failure connections
5	Same service ratio per connection	The frequency of the changes of the services
6	Erroneous segments ratio per unit time	The proportion of wrong segmentation
7	Ratio change of application protocol	The frequency of the changes of the applications

In this manner, the anomaly state of network traffic could be described and depicted combined with one or multiple characteristics, so as to indicate the traffic qualitatively. For instance, a qualitative description of the characteristics of traffic include: per unit time (assuming 10 minutes, mainly for the TCP) accounted for the rate of change of 5%, the average connection duration is 15 minutes, the failed connection accounted is 2%, the error segment accounted for 0.03%, the same service (assuming WWW is the main access) accounted for 55%, the flow rate is 0.8M Pkt/Min, and the average packet length is 800 Byte.

3.2 Membership degree of traffic characteristics

After characteristic of traffic have been qualitatively described, each characteristic will have three anomaly levels: Low, Medium, and High. These levels are evaluated based on expert evaluations of the ranges of values of each characteristic. That is, there are three gauge clouds corresponding to each characteristic, and the gauge clouds are taken as the criterion of anomaly determination. In face of given target traffic, the characteristic will be compared with the gauge cloud, and the degree of anomaly membership (High/Middle/Low) of different characteristics could be obtained.

Anomaly membership cloud for each characteristic is established to describe the distribution of membership values of qualitative concepts, which should also be a normal cloud. Each degree of membership is a random value that follows the normal distribution in order to reflect the uncertainty of the qualitative concept and quantitative values. Suppose the number of experts is n, the expert set is $P=\{ p_1, p_2, \dots , p_n \}$, expert p_i evaluates the anomaly membership of one characteristic to be three levels, as h_i, m_i, l_i respectively, and satisfying the condition that $0 \le h_i, m_i, l_i \le 1$. Based on this, three membership evaluation results could be processed using the backward cloud generator, to obtain digital features of Ex, En, and He, which reflect the anomaly level (High/Medium/Low) of the qualitative concept.

Table 2: Scope of characteristics value

#	Characteristics	Scope
1	Flow rate of traffic	Slow(<5), Medium(5~10), Fast(>10)(M Pkt/Min)
2	Average length of packet	Short(<700), Medium(700~1250), Long(>1250)(Byte)
3	Average duration of single connection	Short(<40ms), Medium(40ms~10Min), Long(>10Min)
4	Failure connections ratio per unit time	Low(<20%), Medium(20%~30%), High(>30%)
5	Same service ratio per connection	Low(<40%), Medium(40%~80%), High(>80%)
6	Erroneous segments ratio per unit time	Low(<2%), Medium(2%~5%), High(>5%)
7	Ratio change rate of application protocol	Low(<20%), Medium(20%~30%), High(>30%)

It is a great number of cloud drops that could effectively reflect the relationship between the qualitative concept and quantitative numerical value, therefore the digital characteristics obtained from many experts are merged and taken to produce the cloud diagram based on forward cloud generator. The cloud diagram corresponds to the three anomaly levels of qualitative concept, namely High, Middle and Low. Considering the fair randomness of different opinions from experts, the fused characteristics are good at depicting the fuzziness degree of qualitative concept and the randomness degree of cloud drops.

3.3 Anomaly detection and determination

The fundamental process of anomaly detection and determination algorithm based on the cloud model is shown in Fig. 4.

Stage 1. Compare a specific characteristic of the traffic with corresponding gauge cloud in order to obtain expectations of this specific characteristic including three levels as High, Middle and Low, taken as the membership degrees (md) to depict the anomaly situation.

Stage 2. Calculate all the membership degrees of characteristics within target traffic based on Stage 1 respectively. Suppose the number of traffic characteristics is n, and every characteristic corresponds to three values of membership degrees, thus the Matrix (n×3) could be composed and represented as formula (9).

$$Matrix = \begin{bmatrix} md_H^1 & md_M^1 & md_L^1 \\ . & . & . \\ . & . & . \\ . & . & . \\ md_H^n & md_M^n & md_L^n \end{bmatrix} \tag{9}$$

$$WM = \begin{bmatrix} \alpha\kappa^1 md_H^1 & \beta\kappa^1 md_M^1 & \gamma\kappa^1 md_L^1 \\ . & . & . \\ . & . & . \\ . & . & . \\ \alpha\kappa^n md_H^n & \beta\kappa^n md_M^n & \gamma\kappa^n md_L^n \end{bmatrix} \tag{10}$$

Stage 3. According to the differences among anomaly levels of High, Medium, Low, different weights are given to mdH, mdM, mdL, and the proportion is α : β : γ. Besides, in view of the difference in mapping and depicting the anomalies, each characteristic is given a different weight, Wmd1: Wmd2:…:Wmdn=κ1:κ2:…:κn, and the WM (Weighted-Matrix, n×3) could be denoted as formula (10).

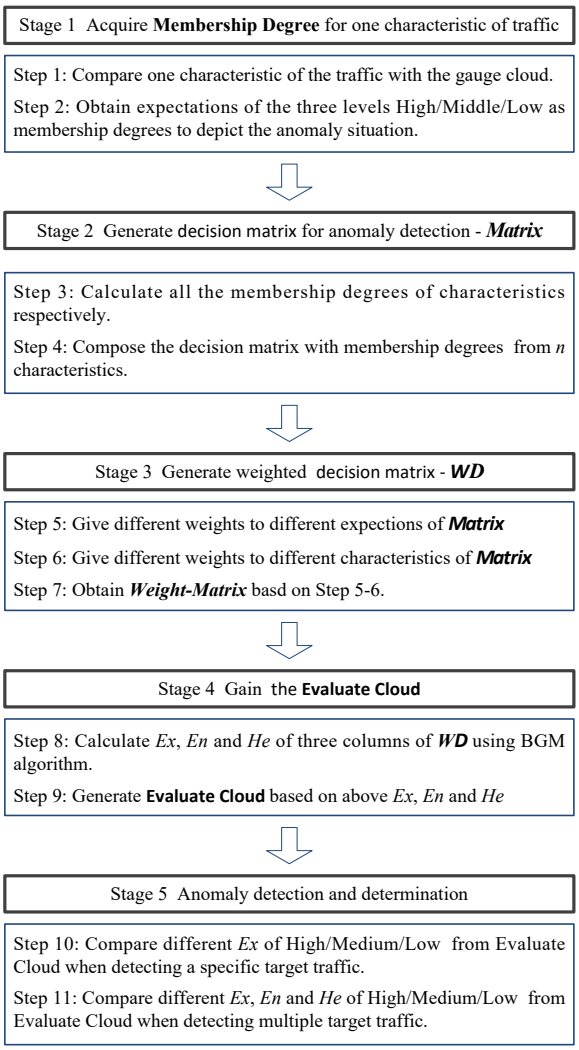

Figure 4: Flowchart of anomaly detection and determination algorithm

Stage 4. BCT is utilized to pursue digital features of Ex, En, and He corresponding to the anomaly degree memberships of High/Medium/Low respectively based on WM. Later, the Evaluation Cloud will be created from above Ex, En, and He with FCT algorithm, which covers three sub-clouds referring to the anomaly degrees of High/Medium/Low levels, ensuring that full information originated from traffic characteristics will be fused under these circumstances.

Stage 5. Considering expectation value (Ex) is the most representative indicator of "anomaly", the determination process abides by the following criterion: (1) when detecting a specific traffic, the expectation values of High, Medium and Low will be compared to select the maximum referring to the maximal membership degree of the qualitative concept. When multiple samples of traffic need to be compared, the expectation of High cloud from samples could be directly selected and compared, the larger value means the higher degree of anomaly. When two expectations are the same incidentally, the smaller entropy value (En) of qualitative concept will be decided. In addition, if both the value of expectation pairs and the entropy pairs are the same, the smaller hyper entropy (He) has to be chosen to find the better qualitative concept.

4 Simulation

Experiments to verify the YATA method are carried out by using a standard dataset through two phases: the preparation part is to generate the gauge clouds for follow-up work, and the Implementation part is to demonstrate the proposed method in two logically progressive sections.

4.1 Preparation

In order to verify the effectiveness of proposed method, the simulation process takes advantage of the kddcup_data_10_percent.zip in the benchmark data sets of KDD Cup 1999. Because there are totally 494021 items of traffic records in the kddcup.data_10_percent.txt after decompression, to simplify the simulation, we had selected 6648 items to run experiments, involving normal traffic and three sorts of attacks, including denial of service attack (DoS), scan and sniffing (Probe), unauthorized access from a remote machine (R2L), as the anomaly detection dataset for simulation. Among the 6648 records, almost 2/3 of them are chosen for the generation of gauge cloud, denoted as DB-Gauge, and the remainder 1/3 are kept for the YATA verification, named as DB-VEF. Besides, both of DB-Gauge data set and DB-VEF data set include normal, DoS, Probe and R2L traffic.

It is significant to generate the gauge clouds for all the characteristics of traffic, which is prepared for the next stage, namely, anomaly detection and determination. On the basis of classifications towards attack and normal traffic within DB-Gauge, all the numeric selections of 7 traffic characteristics would be estimated by analogue experts, 770 for each. In this way, the membership degrees of High/Medium/Low levels of every characteristic could have been obtained.

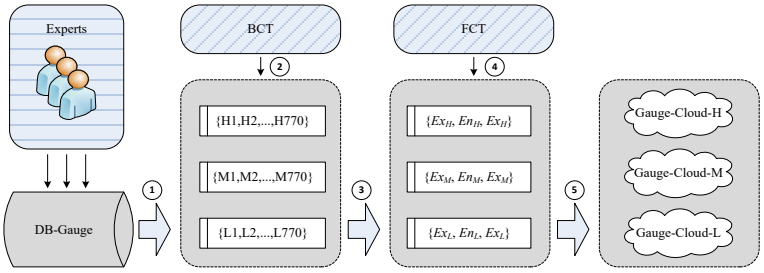

Figure 5: Process of gauge cloud generation for one characteristic

Taking the characteristic flow rate of traffic for instance, three sets including {H1,H2,…,H770}, {M1,M2,…,M770} and {L1,L2,…,L770} will be acquired by the assessments of analogue experts, which produces numerical values including {ExH, EnH, HeH}, {ExM, EnM, HeM}, {ExL, EnL, HeL} based on BCT algorithm. Subsequently, FCT is utilized to generate the gauge clouds corresponding to different anomaly levels as High, Medium, and Low for the characteristic of flow rate of traffic. Therefore, the gauge clouds of other six characteristics would be built the same way.

4.2 Implementation

Section 1: Single Traffic Analysis of DoS.

In this section, certain traffic within the DoS category from DB-VEF is selected to demonstrate the YATA method to check it could determine the degree of anomaly for this traffic.

First, gain the Matrix. The sample of DoS traffic extracted from DB-VEF is denoted as F1 and has the following characteristics: average rate is 9.6M Pkt/Min, average duration of single connection lasts for 70.4 ms, failure connections ratio per unit time is up to 25.1%, and average length of packet is 51 Bytes, etc. Next, each of the 7 characteristics is compared with the corresponding gauge cloud respectively, to obtain the decision matrix following the process which is illustrated in Fig.4. Because there are 7 characteristics and 3 levels (H, M, L), the 7×3 Matrix-F1 is calculated as follows.

$$Matrix\text{-}F1 = \begin{bmatrix} 0.727 & 0.551 & 0.503 \\ 0.898 & 0.580 & 0.406 \\ 0.814 & 0.624 & 0.479 \\ 0.698 & 0.739 & 0.320 \\ 0.770 & 0.656 & 0.387 \\ 0.807 & 0.611 & 0.528 \\ 0.792 & 0.593 & 0.445 \end{bmatrix} \tag{11}$$

Secondly, gain the WD-F1 based on Matrix-F1. According to the importance and influence of every characteristic as well as the different levels of anomaly, we have the weight specifications as shown in Tab. 3 and Tab. 4.

Table 3: Weight Specifications for Levels

Levels	Weight
High	1
Medium	0.6
Low	0.3

Table 4: Weight Specifications for Characteristics

Characteristics	Weight
Flow rate of traffic Average length of packet Average duration of single connection	1
Failure connections ratio per unit time Same service ratio per connection Ratio change rate of application protocol	0.8
Erroneous segments ratio per unit time	0.7

Therefore, the WD-F1 shown below would be obtained through formula (10):

$$WD\text{-}F1 = \begin{bmatrix} 0.727 & 0.331 & 0.151 \\ 0.898 & 0.348 & 0.122 \\ 0.814 & 0.374 & 0.144 \\ 0.698 & 0.443 & 0.096 \\ 0.616 & 0.315 & 0.093 \\ 0.646 & 0.293 & 0.127 \\ 0.554 & 0.249 & 0.094 \end{bmatrix} \tag{12}$$

Thirdly, BCT algorithm would be used to calculate Expectation, Entropy and Hyper Entropy according to the different levels of High/Medium/Low based on WD-F1.

- $Ex_H = 0.708$, $En_H = 0.113$, $He_H = 0.035$
- $Ex_M = 0.336$, $En_M = 0.056$, $He_M = 0.026$
- $Ex_L = 0.118$, $En_L = 0.026$, $He_L = 0.009$

Finally, FCT algorithm would be adopted to generate evaluation clouds combined with the above parameters.

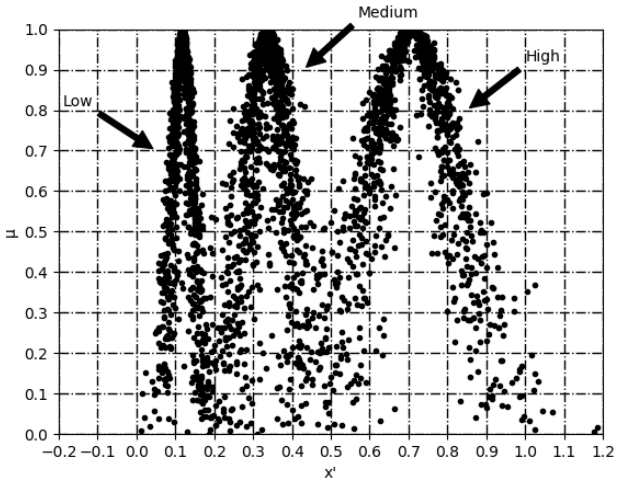

Figure 6: Evaluation clouds of F1 including three anomaly levels

As depicted in Fig. 6, three peaks correspond to the evaluation clouds of High/Medium/Low levels. In each area every dot represents a cloud drop, namely $drop_i = (x_i, \mu_i)$ where μ_i indicates the degree of certainty that x_i belongs to the qualitative concept of anomaly traffic.

It can be seen from the diagram that the majority of the cloud evaluations have a level of High which represents the highest degree of membership. Therefore, the evaluation cloud of level High should be focused on to reflect the current situation of traffic, compared with Medium and Low levels.

The sample traffic F1 whose average rate is up to 9.6M Pkt/Min, significantly indicating a DoS attack. Moreover, considering that the failure connections ratio is 25.1% and the single connection average duration lasts for 70.4 ms both fall into the high gauge level estimated by experts. Therefore, the evaluation cloud of High level preferably reflects the anomaly degree of the F1.

Section 2: Single Normal Traffic Analysis.

At this time, certain normal traffic from DB-VEF would be selected to demonstrate the ability of YATA method to distinguish whether the traffic is suspected anomalous or not. The normal traffic marked as F2 covers the following characteristics: average rate is 3.4K Pkt/Min, average duration of single connection lasts for 175.2 s, failure connections ratio per unit time is less than 0.3%, and average length of packet is 684 Bytes, etc. The decision matrix of F2, together with WD-F2 could be obtained as follows.

$$Matrix\text{-}F2 = \begin{bmatrix} 0.102 & 0.490 & 0.876 \\ 0.075 & 0.323 & 0.971 \\ 0.093 & 0.447 & 0.858 \\ 0.112 & 0.201 & 0.842 \\ 0.084 & 0.352 & 0.974 \\ 0.069 & 0.399 & 0.884 \\ 0.101 & 0.326 & 0.945 \end{bmatrix} \quad (13)$$

$$WD\text{-}F2 = \begin{bmatrix} 0.102 & 0.294 & 0.263 \\ 0.075 & 0.194 & 0.291 \\ 0.093 & 0.268 & 0.257 \\ 0.112 & 0.121 & 0.253 \\ 0.067 & 0.169 & 0.234 \\ 0.055 & 0.191 & 0.212 \\ 0.071 & 0.137 & 0.199 \end{bmatrix} \quad (14)$$

WD-F2 and evaluation clouds would be calculated based on BCT and FCT algorithms consequently. Therefore, the Ex, En and He values are:

- $Ex_H = 0.082$, $En_H = 0.022$, $He_H = 0.008$
- $Ex_M = 0.196$, $En_M = 0.061$, $He_M = 0.152$
- $Ex_L = 0.244$, $En_L = 0.031$, $He_L = 0.136$

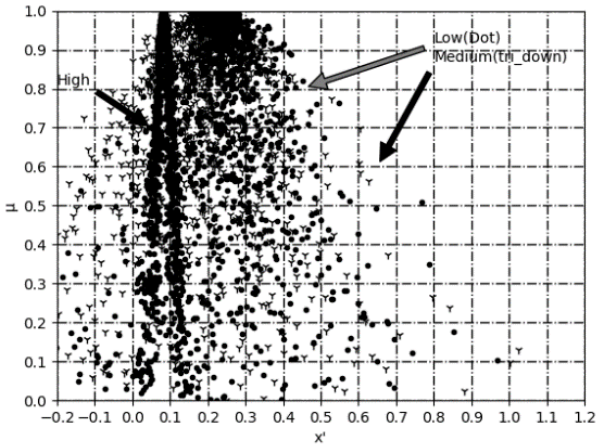

Figure 7: Evaluation clouds of F2 including three anomaly levels

In Fig. 7, there exist three areas respectively indicating the evaluation clouds of High/Medium/Low levels for traffic F2. It is quite different from Fig. 6, because here the Low evaluation cloud and Medium evaluation cloud are partly overlapping. At the same time, they are both cluster to the right of the High evaluation cloud, which means that the F2 sample traffic is very probably normal.

Section 3: Multiple Traffic Analysis.

In order to implement comparisons between different types of traffic, a sample of Probe traffic is extracted from DB-VEF denoted as F3, which has the following characteristics: average rate is 0.8K Pkt/Min, average duration of single connection lasts for 89.7 ms, failure connections ratio per unit time is up to 4.1%, and average length of packet is 135 Bytes, etc. The decision matrix of F3, together with WD-F3 could be obtained as follows:

$$Matrix\text{-}F3 = \begin{bmatrix} 0.590 & 0.443 & 0.247 \\ 0.633 & 0.417 & 0.292 \\ 0.627 & 0.390 & 0.306 \\ 0.776 & 0.378 & 0.118 \\ 0.709 & 0.495 & 0.250 \\ 0.652 & 0.502 & 0.293 \\ 0.590 & 0.461 & 0.179 \end{bmatrix} \tag{15}$$

$$WD\text{-}F3 = \begin{bmatrix} 0.590 & 0.266 & 0.074 \\ 0.633 & 0.250 & 0.088 \\ 0.627 & 0.234 & 0.092 \\ 0.776 & 0.227 & 0.035 \\ 0.567 & 0.238 & 0.060 \\ 0.522 & 0.241 & 0.070 \\ 0.413 & 0.194 & 0.038 \end{bmatrix} \tag{16}$$

Then, WD-F3 and evaluation clouds would be calculated based on BCT and FCT algorithms consequently.

- $Ex_H = 0.708$, $En_H = 0.113$, $He_H = 0.035$
- $Ex_M = 0.336$, $En_M = 0.056$, $He_M = 0.026$
- $Ex_L = 0.118$, $En_L = 0.026$, $He_L = 0.009$

Similarly in Fig. 8, there are three areas describing the evaluation clouds of High/Medium/Low levels for sample F3. Still the evaluation cloud of High level is moved prominently to the right which means it is the most persuasive indicator of underlying situation. Because evaluation cloud of High level locates on the right side of the middle, showing that it is in a slightly invaded situation, and generally in conformity with the situation of Probe attack.

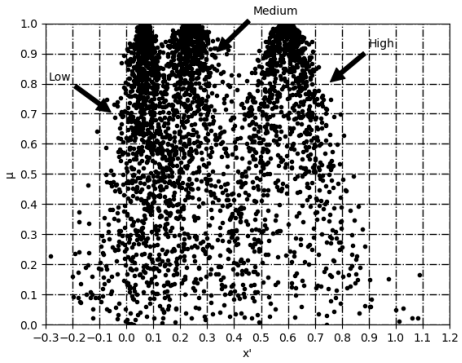

Figure 8: Evaluation clouds of F3 including three anomaly levels

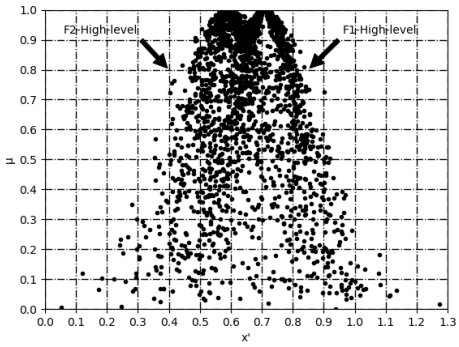

Figure 9: Comparison of F1 and F3

In aspect of the High level cloud from Fig. 9, it is obvious that F1 is grouped to the right of F3, which means F1 behaves more anomalous than F3. Since F1 comes from DoS traffic part of DB-VEF, while F3 comes from the Probe, the result illustrated in Fig. 9 shows good consistency with the varied sorts of traffic, and verifies the effectiveness of YATA, which could reveal the extent of anomaly from a qualitative point of view in a convenient and intuitive way.

5 Conclusion

To protect the network from being exploited by malicious traffic, we propose an anomaly traffic detection and determination method named YATA. In virtue of the Cloud Model, this method is capable of transforming quantitative data to the qualitative concept rapidly and directly, which improves the expressiveness of traffic situation for the security administrators to take further measures. We deploy and demonstrate the feasibility of this method based on KDD Cup 1999.

Acknowledgement: We would like to present our thanks to any anonymous reviewers for their helpful suggestions. This work is supported in part by National Natural Science Foundation of China (No. 61802115), Henan province science and technology projects (Nos. 182102310925, 192102310445), Key Scientific Research projects of Henan Province Education Department (Nos. 18A520004, 19A520008).

References

Cheang, C.; Wang, Y.; Cai, Z.; Xu, G. (2018): Multi-VMs intrusion detection for cloud security using Dempster-Shafer theory. *Computers, Materials & Continua*, vol. 57, no. 2, pp. 297-306.

Chen, L.; Wang, X.; Zhao, X.; Li, W. (2011): Research of botnet anomaly detection algorithm based on private protocol. *IEEE International Conference on Broadband Network and Multimedia Technology*, pp. 55-59.

Gao, H.; Jiang, J.; Zhang, L.; Yuchao, L.; Li, D. (2013): Cloud model: detect

unsupervised communities in social tagging network. *Proceedings of the International Conference on Information Science and Cloud Computing*, pp. 317-323.

Gokcesu, K.; Kozat, S. (2017): Online anomaly detection with minimax optimal density estimation in nonstationary environments. *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 133-143.

Hu, W.; Xiao, X.; Fu, Z.; Xie, D.; Tan, T. et al. (2006): A system for learning statistical motion patterns. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 28, no. 9, pp. 1450-1464.

Jiang, F.; Ling, S.; Chan, K.; Chaczko, Z.; Leung, F. et al. (2012): An immunology-inspired multi-engine anomaly detection system with hybrid particle swarm optimisations. *IEEE International Conference on Fuzzy Systems*, pp. 1-8.

Li, D.; Liu, C.; Gan, W. (2009): A new cognitive model: cloud model. *International Journal of Intelligent Systems*, vol. 24, no. 3, pp. 357-375.

Li, R.; Sun, Y.; Hu, J.; Ma, T.; Luo, X. (2018): Street-level landmark evaluation based on nearest routers. *Security & Communication Networks*, vol. 2018.

Liu, H.; Li, Z.; Song, W.; Su, Q. (2017): Failure mode and effect analysis using cloud model theory and PROMETHEE method. *IEEE Transactions on Reliability*, vol. 66, no. 4, pp. 1058-1072.

Liu, H.; Luan, X.; Li, Z.; Wu, J. (2017): Linguistic petri nets based on cloud model theory for knowledge representation and reasoning. *IEEE Transactions on Knowledge & Data Engineering*, vol. 30, no. 4, pp. 717-728.

Naseer, S.; Saleem, Y.; Khalid, S.; Bashir, K. M.; Han, J. et al. (2018): Enhanced network anomaly detection based on deep neural networks. *IEEE Access*, vol. 6, pp. 48231-48246.

Ningrinla, M.; Amar, T.; Kumar, P. (2018): Detecting byzantine attack in cognitive radio networks by exploiting frequency and ordering properties. *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 816-824.

Peng, H.; Wang, J. (2018): A multicriteria group decision-making method based on the normal cloud model with Zadeh's z-numbers. *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 6, pp. 3246-3260.

Sukhanov, A.; Kovalev, S.; Stýskala, V. (2016): Fuzzy interpretation for temporal-difference learning in anomaly detection problems. *Bulletin of the Polish Academy of Sciences Technical Sciences*, vol. 64, no. 3, pp. 625-632.

Sun, T.; Tian, H. (2014): Anomaly detection by diffusion wavelet-based analysis on traffic matrix. *International Symposium on Parallel Architectures*, pp. 241-248.

Treurniet, J. (2011): A network activity classification schema and its application to scan detection. *IEEE/ACM Transactions on Networking*, vol. 19, no. 5, pp. 1396-1404.

Wang, G.; Xu, C.; Li, D. (2014): Generic normal cloud model. *Information Sciences*, vol. 280, pp. 1-15.

Xie, K.; Li, X.; Wang, X.; Cao, J.; Xie, G. et al. (2018): On-line anomaly detection with high accuracy. *IEEE/ACM Transactions on Networking*, vol. 26, no. 3, pp. 1222-1235.

Yang, S.; Jiang, R.; Wang, H.; Ge, S. (2017): Road constrained monocular visual localization using Gaussian-Gaussian cloud model. *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 12, pp. 3449-3456.

Yang, X.; Yan, L.; Peng, H.; Gao, X. (2014): Encoding words into cloud models from interval-valued data via fuzzy statistics and membership function fitting. *Knowledge-Based Systems*, vol. 5, pp. 114-124.

Yao, D.; Shu, X.; Cheng, L.; Stolfo, S.; Bertino, E. et al. (2017): Anomaly detection as a service: challenges, advances, and opportunities. *Synthesis Lectures on Information Security Privacy & Trust*, Morgan & Claypool Publishers, USA.

Zhao, F.; Luo, X.; Gan, Y.; Zu, S.; Cheng, Q. et al. (2018): IP geolocation based on identification routers and local delay distribution similarity. *Concurrency and Computation: Practice and Experience*, pp. 1-15.

Zhou, M.; Wang, Q.; Ren, K.; Koutsonikolas, D.; Su, L. et al. (2018): Dolphin: real-time hidden acoustic signal capture with smartphones. *IEEE Transactions on Mobile Computing*, vol. 18, no. 3, pp. 560-573.

Zu, S.; Luo, X.; Liu, S.; Liu, Y.; Liu, F. (2018): City-level IP geolocation algorithm based on pop network topology. *IEEE Access*, vol. 6, pp. 64867-64875.