A router abnormal traffic detection strategy based on active defense

Xin Li, Peng Yi *, Yiming Jiang and Jing Yu

Information Technology Institute, PLA Strategic Support Force Information Engineering University, Zhengzhou, Henan, 450000, China

*Corresponding author’s e-mail: li-xin@alu.hit.edu.cn

Abstract. With the rapid development of network attacks, traditional security protection technology is difficult to deal with unknown threats and persistent attacks. Active defense improves the ability to defend against network attacks by building a dynamic, heterogeneous and redundant endogenous security system. Aiming at the problem of single abnormal arbitrament information of routers in mimic defense, a router abnormal traffic detection strategy based on active defense is proposed. By clustering the traffic information of multiple heterogeneous redundant routing function entities and comparing the distance measurement between them, the routing function entities in abnormal state are determined. The experimental results show that the proposed strategy effectively detects the security threats of routing functional entities and expand the method of mimic router arbitrament.

1. Introduction
In recent years, the rapid development of computer network technology has made the Internet spread to all aspects of society. At the same time, the security problems brought by the Internet are emerging in endlessly, and the attack methods against the Internet are changing with each passing day. As an important device in the Internet, a router transfers information in the Internet through the storage and packet forwarding of data packets between different networks. Therefore, ensuring the security and reliability of routers plays a significant role in maintaining the stability of the Internet system and improving the anti-aggression of the Internet system.

Traditional security protection technology is a static defense [1]. Traditional network security defense is built on the existing network architecture. It provides security protection for the data transmission system through deployment of defense strategies, including firewall technology [2], intrusion detection [3] and other systems. It is improved by passive ways such as constantly digging vulnerabilities and blocking back doors. When the system target is threatened by Trojan horse or backdoor attack, the characteristic library cannot be updated in time, which leads to the lag of the passive defense system. At the same time, passive defense relies on prior knowledge and is difficult to resist unknown security threats. Therefore, this kind of defense effect which can only target the "specific" attack makes the passive defense have a big limitation. Active defense changes the inherent stativity, certainty and similarity of traditional security protection technologies by constructing an endogenous security system with dynamic, heterogeneity and redundancy. It reduces the success rate of exploits and interferes with the controllability of backdoors, thereby significantly increasing the difficulty and cost of attacks. Common active defense technologies include intrusion tolerance technology [4], Moving Target Defense (MTD) [5], etc. Referring to the concept of "Mimic Disguise (MD)" in the biological field, Jiangxing Wu and his team proposed an innovative active defense theory
of Cyberspace Mimic Defense (CMD) [6]. CMD creates an integrated defense system of endogenous security through the selection of dynamic and diversified mechanisms and strategies, which greatly increases the difficulty and cost of attack.

Most of the anomaly detection methods of routers based on mimic defense architecture focus on router configuration ruling and routing table ruling, and the method for discovering anomalies is relatively single. In addition, certain attacks are not aimed at routing table information, making it difficult for the detection mechanism to find abnormalities. In order to solve the above problems, we designed an abnormal router traffic detection strategy based on active defense from the perspective of router traffic information to broaden the dimension of mimic router anomaly detection.

The remaining sections of this paper is organized as follows. The related works are discussed in section 2. The section 3 provides a detailed description of the proposed router anomaly detection strategy based on active defense. Section 4 shows the performance evaluation. Section 5 is concluded with future research. Nakibly, Menahem and Waizel [7] proposed an improved scheme for OSPF routing protocol attack methods and a scheme for automatically detecting the vulnerability of black box routers at the Black Hat conference.

2. Related work

The security threat of routers is reflected in the unavoidable loopholes in the system design and implementation, and the trapdoors brought into unconsciously by using open source code. In [8], a method of remote attack on router based on buffer overflow principle is introduced, and the preventive measures of this kind of attack are analyzed. Attack behaviors such as network scanning attacks and denial of service attacks will change the network traffic. By summarizing the packet load characteristics and session flow state characteristics of router traffic, reference [9] introduced the convolutional neural network and the long-term-time memory circulating neural network to build the router intrusion detection system.

For the router software system, the unit for message processing is defined as the routing function entity, and the router software system contains various routing protocols and routing function entities of management software [10]. Vulnerabilities and backdoors in the routing function entity can be scanned and exploited by attackers for privilege escalation, system control, and information acquisition. In [10], a dynamic heterogeneous redundant (DHR) router mimic defense system model was designed, as shown in Figure 1. According to the function of the software system, the model introduces multiple heterogeneous and redundant routing functional entities to process the same input, and vote their output messages in multi-mode, so as to identify which functional routing functional entity outputs abnormal messages, and then conduct security defense of the routing system.

![Figure 1. Router mimic defense architecture model based on DHR.](image)
In a heterogeneous multi-modal architecture, external threats trigger a single routing functional entity to be abnormal, which in turn makes it inconsistent with other routing functional entities in terms of traffic characteristics. Aiming at the problem of the single ruling information of the current mimic routers, we propose a router abnormal traffic detection strategy based on active defense. By analyzing the multi-dimensional statistical characteristics of heterogeneous routing functional entities from the traffic perspective, security threats in the network are discovered and the method of mimic router arbitrament is improved.

3. A router anomaly detection strategy based on active defense

Figure 2 shows the block diagram of router exception detection based on active defense. Externally, routing functional entities exhibit consistent characteristics, but their internal structure is different. Vulnerability and backdoor attacks have different effects on routing entities, and they exhibit abnormal traffic statistics. In this paper, a router anomaly detection strategy based on active defense is proposed. Through collecting and preprocessing traffic statistics, the unsupervised clustering algorithm is combined with the multi-dimensional arbitrament system based on mimic defense to achieve the purpose of traffic anomaly detection. The multi-dimensional arbitrament architecture of mimic defense ensures the accuracy and reliability of detection results, and broadens the anomaly detection methods of mimic routing functional entities.

3.1. Preprocessing of router traffic data

The CICFlowMeter tool is used to extract the features of traffic data from routing functional entities. Due to the different attributes of the features, the traffic data has different dimensions and orders of magnitude. To eliminate the influence of different orders of eigenvalues on the calculation results and ensure the reliability of the calculation results, it is necessary to standardize the data. Meanwhile, the value corresponding to the features is normalized to the interval of [0, 1], and the normalized expression is as follows:

$$X^* = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

(1)

In order to reduce the redundant features of the data, we use principal component analysis (PCA) to decrease the dimensionality of the data. PCA is a spatial mapping method, which uses linear transformation to transform the original data into linearly independent representations of each dimension, which is used to extract the main feature components of the data.

3.2. Cluster analysis process

The K-means algorithm is used to conduct cluster analysis on the traffic dataset after data preprocessing and feature selection. The basic idea is to calculate the distance between the data points...
in the dataset and the cluster center. By constantly iterating and updating the mean value of each cluster center, the cluster points are adjusted until the objective function converges. The steps of K-means algorithm are as follows.

The cluster sample set is defined as \( X = \{x_1, x_2, \cdots, x_n\} \). \( x \) is a sample and \( n \) is the total number of samples. The mean value of the sample is:

\[
z = \frac{1}{n_k} \sum_{x \in z_k} x
\]  

(2)

where, \( Z_k = \{z_1, z_2, \cdots, z_k\} \) represents the category of the cluster, \( k \) is the number of clusters and \( n_k \) represents the number of samples in the \( k_{th} \) category. The objective function is defined as follows:

\[
J = \sum_{i=1}^{n_k} \sum_{j=1}^{k} d_{ij}(x_i, c_j)
\]  

(3)

where, \( \{c_1, c_2, \cdots, c_k\} \) is \( k \) cluster centers, \( d_{ij}(x_i, c_j) \) is the Euclidean distance between the \( i_{th} \) data and the \( j_{th} \) cluster center. The objective function \( J \) represents the sum of squared distances from all points in each class of the sample dataset to the cluster center. According to \( d_{ij}(x_i, c_j) \), the Euclidean distance between each point and the cluster center is calculated, and the sample data is grouped into the cluster containing the cluster center with the nearest sample point. After constant updating of the clustering center, the iteration is completed when the minimum mean square deviation is reached and no change is made, and then the algorithm converges.

Considering the slow convergence speed of the K-means algorithm, Canopy algorithm is adopted to improve the convergence speed of clustering. Canopy algorithm is a simple, fast and accurate object clustering method, which represents all objects as a point in the multi-dimensional feature space. A fast-approximate distance measurement and a comparison between two distance thresholds were used to achieve fast coarse clustering. The Canopy algorithm has an advantage in dividing data into different overlapping subsets through rough distance calculations. Meanwhile, computing sample data vectors in the same overlapping subset reduces the number of samples requiring distance calculations.

Since different \( k \) has a great impact on the clustering results of the K-means algorithm, we adopted the silhouette coefficient method to select the optimal number of clusters. The silhouette coefficient is a theory that combines the cohesion and separation of the cluster to evaluate the effect of clustering. For the \( i_{th} \) vector in the cluster, its silhouette coefficient is:

\[
S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}
\]  

(4)

where, \( a(i) \) is the degree of dissimilarity within the cluster, which represents the average of the dissimilarity of the first vector to the other points in the cluster. \( b(i) \) is the degree of dissimilarity between the cluster, which represents the minimum of the average dissimilarity of the \( i_{th} \) vector to the other clusters. The average silhouette coefficient of all data points is defined as follows:

\[
S = \frac{\sum_{i=1}^{m} s(i)}{m}
\]  

(5)

The value range of \( S \) is [-1,1]. The closer it is to 1, the better the clustering effect. When \( S \) reaches the maximum, \( K \) is the optimal number of clusters.
3.3. Abnormal detection and arbitrament of routing functional entities

Figure 3 shows the structure of abnormal detection and arbitrament for routing functional entities.

As the clustering center of K-means algorithm is determined by the mean value of samples, the clustering result is susceptible to the influence of outliers. When the routing function entity is attacked, abnormal traffic data causes the change and shift of the cluster center. Considering this characteristic, we determine the occurrence of abnormal traffic by comparing the distance measures and similarity measures between the cluster centers of each two routing functional entities. Distance measurement is used to measure the distance of individuals in space. The farther the distance, the greater the difference between routing function entities. Euclidean distance is a classical distance measurement used to measure the absolute distance between points in a multidimensional space. It is defined as follows:

$$ EuclideanDist(p, q) = \left( \sum_{i=1}^{n} (p_i - q_i)^2 \right)^{1/2} $$ (6)

When all routing function entities are running normally, the clustering center of traffic data of each two routing function entities is relatively close. When a routing function entity is attacked and abnormal traffic is generated, the distance between the abnormal routing function entity and the clustering center of the normal routing function entity increases obviously. When a routing function entity is attacked and generates abnormal traffic, the distance of cluster centers between the abnormal and other routing functional entities increases significantly. By comparing the distance measurement of the cluster centers of all routing functional entities, the distance judgment mechanism is used to determine the final response result. When the distance between a certain routing functional entity and its neighbors are obviously greater than the distance between other routing functional entities, it is judged as abnormal by the ruling mechanism. Furthermore, the result of the arbitrament is added to the historical record, and the credibility of the abnormal routing function entity is reduced. The system sends out alarm information and replaces abnormal routing entities according to the scheduling rules in the mimicry defense architecture to ensure the normal operation of the system.

4. Simulation results and analysis

The CICIDS2017 dataset is used for the experiment, which is a novel network intrusion detection dataset released by the Canadian Institute for Cybersecurity (CIC). It extracted more than 80 network flow features from the generated network traffic using CICFlowMeter. We annotate traffic records according to different attack periods and types, and standardize and normalize the dataset. Due to the excessive amount of data contained in the analyzed CSV file, problems such as excessively long time consuming and slow convergence rate of the model will occur when the host is used for model training.
Therefore, we simplified and reintegrated these CSV data files while preserving the original attack features. The data of routing function entity 1 and 2 are normal traffic data, while routing function entity 3 adds 5% "PortScan", 10% "DoS" and 20% "DDoS" to the normal traffic data respectively to simulate attack traffic and test the performance of routing anomaly detection strategy based on active defense. The specific category composition of the datasets is shown in Table 1.

Table 1. The specific category composition of the datasets.

| Dataset  | Attack types | The number of normal samples | The number of abnormal samples | Number of Samples |
|----------|--------------|------------------------------|-------------------------------|-------------------|
| Test set 1 | PortScan     | 2240                         | 112                           | 2352              |
| Test set 2 | DoS          | 2138                         | 214                           | 2352              |
| Test set 3 | DDoS         | 5160                         | 1032                          | 6192              |

Table 2 shows the optimal number of clusters (ONC), average silhouette coefficient (ASC) and cluster center vectors of routing functional entities for the three types of anomalies.

Table 2. Experimental results of different test sets.

| Dataset  | ONC | ASC  | Cluster center vector of routing function entity 1 | Cluster center vector of routing function entity 2 | Cluster center vector of routing function entity 3 |
|----------|-----|------|---------------------------------------------------|---------------------------------------------------|---------------------------------------------------|
| Test set 1 | 2  | 0.9041 | 0.0442,0.0371,0.0428,0.0402,0.0416,0.0174        | 0.0448,0.0376,0.0416,0.0401,0.0411,0.0174        | 0.0429,0.0372,0.0396,0.0391,0.0402,0.0174        |
| Test set 2 | 5  | 0.9298 | 0.0067,0.0611,0.0362,0.0035,0.0235,0.0238        | 0.0067,0.0609,0.0037,0.0293,0.0375,0.0027        | 0.0071,0.0579,0.0049,0.0394,0.0017,0.0024        |
| Test set 3 | 2  | 0.9108 | 0.0549,0.0358,0.0537,0.0362,0.0479,0.0294        | 0.0552,0.0367,0.0533,0.0358,0.0435,0.0296        | 0.0481,0.03053,0.0509,0.0313,0.0393,0.03125      |

Because different types of anomalies are affected by different characteristics, the optimal clustering number is not consistent. The silhouette coefficient reflects the influence of the number of clusters on the clustering effect. It can be seen from Table 2 that the average silhouette coefficients of the experimental datasets are all above 0.9, indicating that the introduction of Canopy algorithm and the silhouette coefficient increase the cohesion degree within the clusters and the separation degree between clusters. The improved clustering performance is helpful to get more accurate clustering results. Observing the cluster center vectors of different routing functional entities, it is found that there are obvious differences between the feature vectors of normal and abnormal traffic. It shows that the abnormal flow changes the characteristic value, which results in the deviation of the flow characteristic vectors. Furthermore, due to the characteristic that the K-means algorithm is sensitive to outliers, the obtained cluster center vectors effectively reflects the changes in abnormal traffic.

Table 3. Abnormal detection results of routing functional entities.

| Dataset  | Routing function entity number | Distance measurement | The result of arbitration |
|----------|--------------------------------|----------------------|--------------------------|
| Test set 1 | Routing function entity 1 and 2 | 1.256E-03            | Routing function entity 3 is abnormal |
| Test set 1 | Routing function entity 1 and 3 | 7.813E-03            |                          |
| Test set 1 | Routing function entity 2 and 3 | 1.158E-02            |                          |
| Test set 1 | Routing function entity 1 and 2 | 1.776 E-03           |                          |
| Test set 2 | Routing function entity 1 and 3 | 2.326 E-02           | Routing function entity 3 is abnormal |
| Test set 2 | Routing function entity 2 and 3 | 1.972 E-02           |                          |
| Test set 2 | Routing function entity 1 and 2 | 3.526 E-03           |                          |
| Test set 3 | Routing function entity 1 and 1  | 5.663 E-02           | Routing function entity 3 is abnormal |
| Test set 3 | Routing function entity 2 and 3 | 6.975 E-02           |                          |

According to the cluster center vectors of routing function entities, Euclidean distance between two routing function entities is used as distance measurement, and the results are shown in Table 3. When the routing function entity is abnormal, the distance measurement between it and the adjacent routing functional entity becomes larger. In the test set 1 where 5% of the "PortScan" attack is added, the
distance between the abnormal routing functional entity 3 and the them is 7.813E-03 and 1.158E-02 respectively, which is 6.22 and 9.21 times of the distance between the two normal routing function entities. It indicates that the proposed routing anomaly detection strategy based on active defense can effectively detect the abnormal routing function entities, even if only a small part of the abnormal traffic is generated. In the test set 3 with 20% “DDoS” attacks, it increased to 16.06 and 19.78 times. As a result, the abnormal routing function entity is clearly detected, and timely send an alarm message to the system. According to the scheduling rules in the mimic architecture, the routing function entities are replaced and cleaned to ensure the stable operation of the system.

5. Conclusions and future work
With the development of network technology, network attack is updated rapidly. Traditional network defense is difficult to resist the unknown threats and persistent attacks faced by key services of cyberspace information system. As an active defense method, mimicry defense technology blocks the accessibility of attack by increasing the system dynamics, heterogeneity and redundancy. In order to solve the problem that the abnormal arbitrament information of router in mimic defense is relatively single, we propose a router abnormal traffic detection strategy based on active defense from the perspective of traffic. The Canopy algorithm is used to improve the convergence speed of K-means algorithm by clustering the traffic information of multiple heterogeneous redundant routing entities, and the silhouette coefficient method is used to select the optimal cluster number. Abnormal routing entities are determined by comparing the distance measurement between routing entities. The experimental results show that the proposed strategy can effectively detect the abnormal routing functional entities, and the detection results become more obvious with the increase of abnormal data. For future work, we consider deploying the proposed strategy in a real environment to verify the effectiveness and identify attacks by comparing the types of abnormal features.

Acknowledgments
This research is partially supported by the National Key Technology Research and Development Program of China, grant number 2017YFB0803204 and 2018YFB0804002.

References
[1] Wu J. (2017) Introduction to The Defense of Mimicry in Cyberspace: Vol. 2. Science Press, Beijing.
[2] Li W. (2010) Computer Information Security Technology. National University of Defense Technology Press, Changsha.
[3] Meng W. (2018) Intrusion detection in the era of IoT: Building trust via traffic filtering and sampling. Computer, 51(7): 36-43.
[4] Pal P, Webber F, Schantz R E, et al. (2000) Intrusion tolerant systems. In: IEEE Information Survivability Workshop. pp. 24-26.
[5] S. Jajodia et al. (2011) Moving target defense: creating asymmetric uncertainty for cyber threats. Springer Science & Business Media.
[6] Wu J. (2016) Research on cyber mimic defense. Journal of information security, 1(4).
[7] Nakibly G, Menahem E, Waizel A, et al. (2013) Owning the Routing Table. Part II. Black Hat.
[8] Dan F, ZOU M. (2007) Research on Buffer Overflow attack of Cisco IOS System.
[9] Liu Y, Liu S, Wang Y. (2017) Route intrusion detection based on long short term memory recurrent neural network. DEStech Transactions on Computer Science and Engineering.
[10] Ma H, Yi P, Jiang Y, et al. (2017) Router Mimicry Defense Architecture based on dynamic heterogeneous redundancy mechanism. Journal of Information Security, (1): 29-42.