Intelligent Automation & Soft Computing
DOI:10.32604/iasc.2021.018895
Article

Tech Science Press

Research on Detection Method of Interest Flooding Attack in Named Data Networking

Yabin Xu[1,2,*], Peiyuan Gu[2] and Xiaowei Xu[3]

[1]Beijing Key Laboratory of Internet Culture and Digital Dissemination Research, Beijing, 100101, China
[2]Computer School, Beijing Information Science and Technology University, Beijing, 100101, China
[3]Department of Information Science, University of Arkansas at Little Rock, Little Rock, 72204, USA
[*]Corresponding Author: Yabin Xu. Email: xyb@bistu.edu.cn
Received: 24 March 2021; Accepted: 25 April 2021

Abstract: In order to effectively detect interest flooding attack (IFA) in Named Data Networking (NDN), this paper proposes a detection method of interest flooding attack based on chi-square test and similarity test. Firstly, it determines the detection window size based on the distribution of information name prefixes (that is information entropy) in the current network traffic. The attackers may append arbitrary random suffix to a certain prefix in the network traffic, and then send a large number of interest packets that cannot get the response. Targeted at this problem, the sensitivity of chi-square test is used to detect the change of prefix of interest packets. Interest packets initiated by IFA attackers are usually attached to a real prefix, but with a randomly generated suffix attached. Taking into account of this problem, the similarity of interest packet prefixes is further detected. Finally, the detection results of the two aspects are combined to determine whether interest flooding attack has occurred or not. In addition, according to the symmetric routing characteristic of Pending Interest Table (PIT), we also send the forged interest packet back to the attacker, and then restrict the corresponding port of the attacker, so as to effectively suppress the IFA attack. The experimental results show that the method we proposed can not only detect IFA in NDN at the beginning of the attack, but also is more accurate and effective than other methods.

Keywords: Named data networking; interest flooding attack; chi-square test; self-similarity; information entropy

1 Introduction

Named Data Networking (NDN) is a typical implementation of Information Centric Networking (ICN) [1]. Different from the traditional IP network, NDN focuses on the content itself rather than the location of the content [2]. It routes based on the content name and does not need to rely on IP-like location information. NDN's content caching mechanism also greatly reduces the network load of nodes, making it more suitable for the Internet content sharing mode with ubiquitous data sources [3].

In terms of security, although NDN can effectively shield Denial of Service (DoS) attacks under traditional networks, but nothing can be done about Interest Flooding Attack (IFA). Attackers can deplete Pending Interest Table (PIT) and computing resources by sending a large number of non-existent malicious interest requests, thus giving rise to network service interruption in a short time and having a great attack and destructive ability on NDN [4].

For the above reasons, the research on detection and defense of interest flooding attack that NDN may suffer is of certain forward-looking and important significance for promoting the development of NDN and enhancing the security of NDN.

In order to prevent criminals from launching interest flooding attacks on NDN and improve the security detection and defense capabilities of NDN, this paper proposes a bilateral detection method of interest flooding attack, which combines chi-square test and similarity test. This method can not only realize the early detection of IFA, but also ensure the validity and accuracy of detection.

The contributions of this study are summarized as follows:

1. Apply the chi-square test method to the IFA test of NDN. Use the sensitivity of the chi-square test to perceive the subtle changes in the interest packet traffic in the NDN, thereby improving the sensitivity of IFA detection and discovering possible attacks in time.
2. In order to judge the possible existence of IFA more accurately, the similarity of interest packet traffic is further detected. By implementing this step, the accuracy of IFA detection can be effectively improved, and the misjudgment of attacks caused by normal network fluctuations can be reduced.
3. Apply sliding window method to sample the network traffic data, for ensuring the timeliness and rationality of data sampling. Discriminate the NDN attack behavior based on the detection results of the chi-square test and similarity test.

2 Related Work

In order to realize fast and accurate detection of IFA in NDN, researchers have proposed several effective methods, which are mainly divided into two categories.

The first category is the detection methods based on the abnormal change of PIT state after IFA. Based on the statistics of interest packets entering and leaving PIT and IFA detection, a push-back mechanism is applied to notify other routers that the router receiving the message no longer accepts all interest packets under the malicious prefix, thus preventing the spread of attacks [5]. Literature Wang et al. [6] proposes a detection mechanism (DPE) based on PIT. In DPE, each router has a malicious list called m-list, which records the number of expired interest packets under each prefix in the pending interest table. When the number of expired interest packets exceeds the threshold, all interest packets under the prefix are marked as malicious interest packets.

Literature Afanasyev et al. [7] detects IFA by statistical data on interest packets entering PIT. The method sets a threshold value at the interface. When the statistical data of an interface exceeds the threshold value, IFA is deemed to have occurred, and then interest packets from the interface will not be forwarded. Literature Tang et al. [8] detects IFA according to the combined value of interest package satisfaction rate and PIT occupancy rate. Literature Wang et al. [9] detects IFA in the network by counting the timeout ratio and occupancy ratio of PIT and setting thresholds.

In summary, this category of methods mainly advocates detecting IFA based on the change of PIT status after the attack, which has certain accuracy. However, when the pit state changes, the attack may have already occurred, or even a large-scale attack, so this category of methods has a certain delay.

The second IFA detection method is based on the change of network traffic characteristics. Literature Compagno et al. [10] proposes a framework Poseidon for detecting IFA. Poseidon uses two parameters to detect attacks. The first parameter is the ratio of incoming interest packets to outgoing packets, and the second parameter is the number of interest packets received by each interface. When both parameters exceed the threshold, the interface is considered to have an IFA. Literature Karami et al. [11] proposes an IFA detection method based on multi-target RBF-PSO. In this paper, the author proposes 12 features to detect IFA, and applies a multi-objective optimization algorithm based on RBF neural network to combine with PSO to improve the accuracy of IFA prediction. Literature Xin et al. [12] proposes an IFA detection scheme based on cumulative entropy by monitoring the abnormal distribution of content requests. Literature Goergen et al. [13–15] uses support vector machine to detect and train interest packets, data packets and other related information under different network conditions, thus obtaining classifiers. Finally, the obtained classifiers are used to detect and classify subsequent network nodes in real time, thus judging whether the network nodes are attacked.

Based on the research status of the above two categories of methods at home and abroad, it can be found that although these methods have certain detection effects on IFA, there is still no effective solution to distinguish between the large fluctuating attack traffic and the real attack traffic. The surge of normal data traffic in the network will also lead to false positives, thus affecting the accuracy of IFA detection. Therefore, it is of great significance to improve the accuracy of IFA detection while ensuring the detection rate and reducing the occurrence of misjudgment for improving the performance of the whole NDN system.

3 General Design of Interest Flooding Attack Detection in NDN

3.1 Analysis of Interest Flooding Attack Principle

In NDN, the Interest packet and Data packet are the basic units used to transmit information [16–18]. Among them, the interest packet is a content request package initiated by the user and carries information such as content name. The data packet carries the information requested by the user to satisfy the interest packet. In NDN architecture, the Pending Interest Table provides two main functions, namely, aggregation of interest packets and forwarding of data packets. Interest packets from each interface will be queried and matched in PIT, and router nodes will store prefix information and detailed name information of unsatisfied interest packets in PIT. When a packet arrives at the PIT, the incoming interface will be obtained from the PIT and outgoing from that interface, while the corresponding entry will be deleted from the PIT.

IFA attackers take advantage of this feature of NDN, a large number of malicious interest packets with real content name prefixes but forged suffixes are initiated to the network, resulting in invalid interest packets being continuously forwarded in each node, thus depleting the PIT resources of the router, causing the PIT on the router to be unable to normally receive interest packets requested by users, thus causing network congestion and even paralysis. On the one hand, attackers can easily attack by forging different false content names without using many network resources; On the other hand, the interest packets sent by attackers are the contents under the same prefix and avoid repeated requests as much as possible, so as to affect the query and aggregation of information names to the greatest extent, thus affecting the normal work of router nodes and causing great destructive power to the network.

3.2 Architecture Design of NDN Interest Flooding Attack Detection System

The overall architecture of interest flooding attack detection in named data network designed in this paper is shown in Fig. 1.

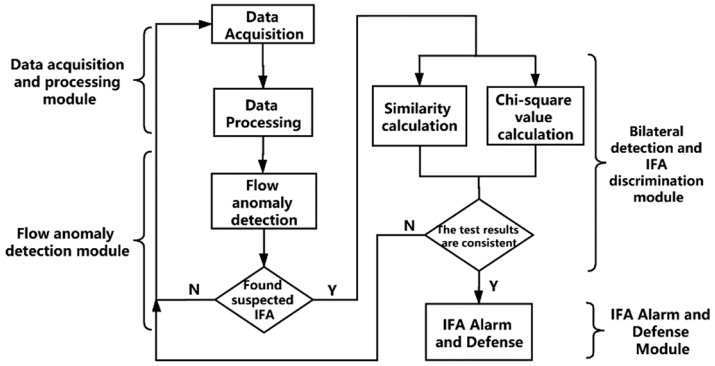

Figure 1: Interest flooding attack detection model

The process of interest flooding attack detection is divided into the following parts:

1. Data acquisition and processing module. The data acquisition and processing module regularly collects traffic data, extracts complete information name prefixes from the collected interest packets and stores them to provide data information for the chi-square traffic determination and similarity test.
2. Flow anomaly detection module. The flow anomaly detection module calculates the chi-square value of the collected data. When the chi-square value exceeds the set threshold value, it indicates that there is a flow anomaly and suspected IFA is generated.
3. Bilateral detection and IFA discrimination module. The bilateral detection and IFA discrimination module is used to carry out the traffic anomaly detection while testing the similarity of traffic, and determined whether there is an interest flooding attack based on the results of the chi-square calculation and similarity calculation.
4. IFA Alarm and Defense Module. The IFA alarm and defense module are used to carry out the IFA alarm according to the above judgment results, and corresponding defense measures are taken to inhibit IFA.

3.3 Setting of Detection Window

In the process of IFA detection, it is particularly important to determine the size of the detection window, and the setting of the window value will have a great impact on the detection results. If the window size is large, may lead to a large scale attacks when the attacks are detected, and the attack behavior cannot be detected in advance. If the window size is small, will get inaccurate traffic distribution. Thus, it is essential to select an appropriate window size to obtain the current information.

According to Shannon's theory [19], information entropy can be used to measure the random change of an information sequence. The smaller the entropy value, the more stable the information sequence. The greater the entropy value, the greater the randomness of the information sequence will be. In NDN, users initiate content requests and obtain required information through interest packets and data packets. Since the information name is the unique identification of information transmission in the whole forwarding process, the change of information name in network traffic can reflect the change of network traffic.

Based on the above theory, we can determine the size of the window according to the distribution of information name prefixes in the current network traffic. Using the method of information entropy, we use a larger window when the prefix distribution is complex (the information entropy value is large) and a smaller window when the prefix distribution is simple (the information entropy value is small).

According to Shannon's theory, the calculation method of information entropy is shown in Eq. (1).

$$E(\mathrm{x}) = -\sum_{i=1}^{r} [p(xi)\,\mathrm{log}p(\mathrm{xi})] \tag{1}$$

where r refers to r independent events in the information sequence, which in this paper represents the number of interest packets received in the network within each window; $p(x_i)$ refers to the probability that each event occurs, which in this paper represents the proportion of interest packets with the same information name prefix. When an attack occurs, quantity of illegal interest packets in the network will destroy the distribution probability of information name prefixes under normal conditions, thus causing changes in information entropy.

In order to obtain the distribution of network traffic under normal network conditions, we need to simply record the information names passing through PIT. In this paper, a counter is added to record all information names, prefixes and times that pass through PIT. The counter is located at the front of the whole PIT and it is completely transparent to the data flow. It does not make any modification to interest packets or data packets, only records the data. Moreover, the counter only records the prefix and quantity records of incoming information, does not record specific information, and has little demand for storage space. The process of construction is shown in Fig. 2.

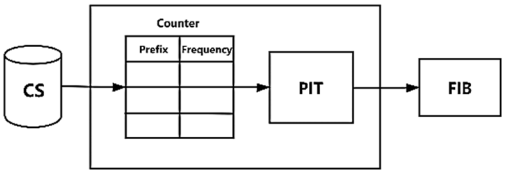

Figure 2: The structure of PIT with counter

When the detection module is started, the detection window is determined. The occurrence time and corresponding probabilities of interest packet prefixes in the counter at this time are recorded. Because of the randomness of network traffic, the total number of prefixes recorded in the counter is different every time the detection window is calculated, so only the first n information prefixes with the highest occurrence times are recorded, and the probability is recorded as $p(xi)$ $(i = 1,\ 2,\ 3\ldots\ n)$. The information entropy value of the first n information prefixes at this time is calculated by using the information entropy formula.

Since the value interval of $p(x_i)$ is (0, 1), $log(p(x_i)) \leq 0$ and increases monotonically with $p(x_i)$, then $E(x)$ also increases monotonically with $p(x_i)$. Therefore, the number of interest packets contained in larger windows and smaller windows can be set in advance, and the threshold value α of information entropy can be determined. When $E(x)$ is greater than α, it means that the more chaotic and uncertain the request number distribution of the first n information prefixes is. In this case, a larger window is adopted to obtain the data flow distribution under the current situation, which improves the accuracy of detection. When $E(x)$ less than the threshold α, it indicates that the distribution purity of the number of requests for the first n information prefixes is higher and the uncertainty is smaller. At this time, a smaller window is used for subsequent detection to discover the attack behavior as soon as possible.

Too many or too few interest packets selected in the window may lead to a large deviation between the actual flow distribution and the theoretical flow distribution, which will further affect the detection effect of the detection model proposed in this paper. Therefore, in the specific determination of window size, we select the appropriate detection window size by comparing IFA detection effects under different windows through experiments.

In order to improve the accuracy of subsequent window selection, after each detection obtains the required window size, all data in the counter are emptied and recorded again.

4 IFA Detection and Defense

4.1 Chi-Square Detection of Interest Packet Prefixes

When an IFA is launched, the attacker will attach any random suffix to a prefix in the network traffic, and then send a large number of interest packets that cannot be responded to. In this attack situation, those large number of interest packets will converge under a certain prefix. Due to the convergence of traffic, the number of PIT entries of network nodes close to content providers or central locations increases rapidly, resulting in PIT cache overflow, central nodes denying service and unable to respond to legitimate interest requests. To solve this problem, this paper uses the chi-square test method to detect IFA that may occur.

Chi-square test is a widely used hypothesis test method, which belongs to the category of nonparametric test. It aims to detect the deviation degree between the actual observation value and the theoretical calculation value of statistical samples [20,21]. The larger the chi-square value, the greater the deviation will be. The smaller the chi-square value, the more consistent the actual situation with the theoretical situation will be.

This paper uses this principle to detect interest flooding attacks. Under normal network traffic, assuming that the probability of prefix occurrence for each packet of interest is equal, the chi-square value will float within a fixed range. When IFA occurs, the network traffic will show a very high proportion of interest packets under a certain prefix, and the chi-square value calculated from this will rise sharply. According to the chi-square value, we can discern the existence of IFA.

The specific detection scheme is as follows: extract all interest packet information from the detection window, extract complete information name prefixes from these interest packets, and put them into W set:

$$W = \{(\mathrm{x1}, \mathrm{y1}),(\mathrm{x2}, \mathrm{y2}),\ldots,(\mathrm{xi}, \mathrm{yi})\} \tag{2}$$

where, x_i is the prefix of the information name, y_i is the number of times the prefix of the information occurs, and the chi-square value calculation formula is as follows in Eq. (3):

$$\chi^2 = \sum_{\mathrm{i}}^{K} \frac{(\mathrm{yi} - NP\mathrm{i})^2}{NP\mathrm{i}} \tag{3}$$

where, K represents the amount of data in the W set, i.e., the number of non-repeating information prefixes; N is the size of the sampled data, and p_i is the frequency of prefix occurrence of each packet of interest under normal circumstances, which is the theoretical value.

4.2 Similarity Test of Interest Packet Prefixes

Interest packets launched by IFA attacks are usually attached to a real prefix, but with randomly generated suffixes, the excessive number of such interest packets destroy the distribution probability of information names under normal conditions, thus causing the similarity of interest packet information name prefixes to change. Therefore, we can detect IFA through the similarity changes of interest packet information name prefixes in different time periods.

The prefix similarity test method of interest packet designed in this paper is as follows:

By counting the prefix distribution of interest packets in adjacent time periods (windows) T_1 and T_2 respectively, and using the prefixes of the top N bits of traffic in this period as the original data of similarity calculation, the data of T_1 and T_2 time periods are processed as follows:

List all interest packet prefixes and arrival times arriving at PIT in T_1 period, calculate the proportion of each prefix, and record it as $x_1, x_2 \cdots x_n$.

According to the prefix of each interest packet in T_1 time period, the proportion of interest packet prefixes corresponding to T_2 time period is listed and recorded as $y_1, y_2 \cdots y_n$. If there are other interest packet prefixes other than T_1 period records in T_2 period, the proportion of this prefix is defined as 0.

The correlation coefficient is calculated for the obtained two groups of percentage data to obtain the similarity of network traffic in T_1 and T_2 time periods.

Firstly, the covariance ($Cov(X, Y)$) is calculated for the two sets of data. The larger the value of covariance, the greater the degree of similarity (similarity) between X and Y variables, and vice versa. Covariance is different from the variance. It is a statistic used to measure the relationship between two random variables. Covariance represents whether the two variables deviate from the mean at the same time and whether the deviation direction is the same or not.

$$Cov(X, Y) = \frac{\sum_{i=1}^{n} (Xi - \bar{X})(Yi - \bar{Y})}{n - 1} \tag{4}$$

Then, according to the flow distribution in each time period, the variance $D(X)$ and $D(Y)$ of the two groups of random variables are calculated:

$$S^2 = \frac{\sum_{i=1}^{n} (Xi - \bar{X})^2}{n - 1} \tag{5}$$

The similarity value P is the quantity value reflecting the similarity degree of flow distribution in T_1 and T_2 time periods.

$$P\text{xy} = \frac{Cov(X, Y)}{\sqrt{D(X)D(Y)}} \tag{6}$$

The closer the value of P is to 1, the more linear the flow distribution in the two time periods, which indicates that the flow distribution in T_1 and T_2 time period is more similar.

Statistics of these data are carried out at regular intervals, and by comparing the similarity of the data in adjacent time windows, the changes in the flow distribution of adjacent time windows are found. When the similarity coefficient P (range from 0 to 1) exceeds the preset set threshold, it indicates that an abnormality may occur.

4.3 IFA Defense Design

Since PIT uses symmetric routing, that is, packets take the same path as their corresponding interest packets, but in the opposite direction. Therefore, we can use this feature to trace the forged interest packet back to the attacker and restrict the corresponding port of the attacker, thus effectively suppressing IFA.

The specific design scheme is as follows: when a node detects IFA, it quickly locates the first information name prefix in the counter (i.e., the prefix containing the largest number of interest packets), then the node sends out false data packets and interest packets under the prefix, and forwards them back to the initiator by looking for PIT in the intermediate node. When a fake packet arrives at the edge router, the edge router finds the host to which the attacker is connected according to the interface when the interest packet is passed in. Then, the edge router restricts the transmission of interest packets from the interface by discarding interest packets from the interface. The specific implementation process is shown in Fig. 3.

As a result, the corresponding PIT entries in the nodes through which malicious interest packets pass will be satisfied by false data packets, and these PIT entries will be deleted, freeing up space to meet the interest

requests of normal users. At the same time, because the sending of malicious interest packets is restricted from the root, malicious interest packets will not spread in the network, thus inhibiting IFA.

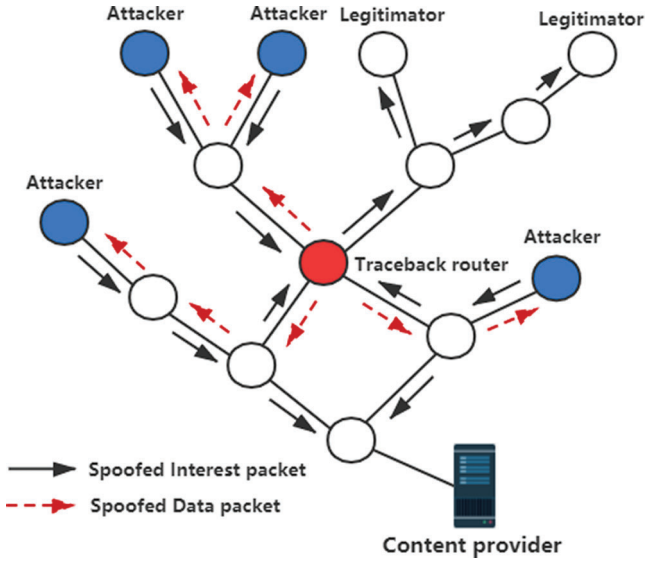

Figure 3: IFA defense process

5 Simulation Experiment and Analysis

5.1 Simulation Experiment Environment and Parameter Configuration

We use ndnSIM for simulation experiments. The ndnSIM is an open-source network simulation platform, and all NDN routing and forwarding experiments can be implemented on ndnSIM [22]. See Tab. 1 for the configuration of the simulation experiment environment. Set the lifetime of interest packets in PIT to 1 second and the size of returned packets to 1024 bytes.

Table 1: Experimental environment configuration table

Main Modules	Specific Configuration
CPU	Intel (R) Core (TM) i5-4590 (4 cores, main frequency: 3.30 GHz)
Memory	8GB
Operating System	Ubuntu 12.04
System Bits	64-bit
NdnSIM Version	2.3

The topology of the experimental network is shown in Fig. 4.

In Fig. 4, R1, R2, R3, and R4 are edge routers used to access end-users. Where R1 and R2 connect to the normally requested user and R3 and R4 connect to impersonate an attacker. R5 is the intermediate router, which realizes the convergence and forwarding function. R6 is directly connected to the data source Producer. The experiment will adjust the attack strength by controlling the number of legitimate requesting users and the attacker's sending rate.

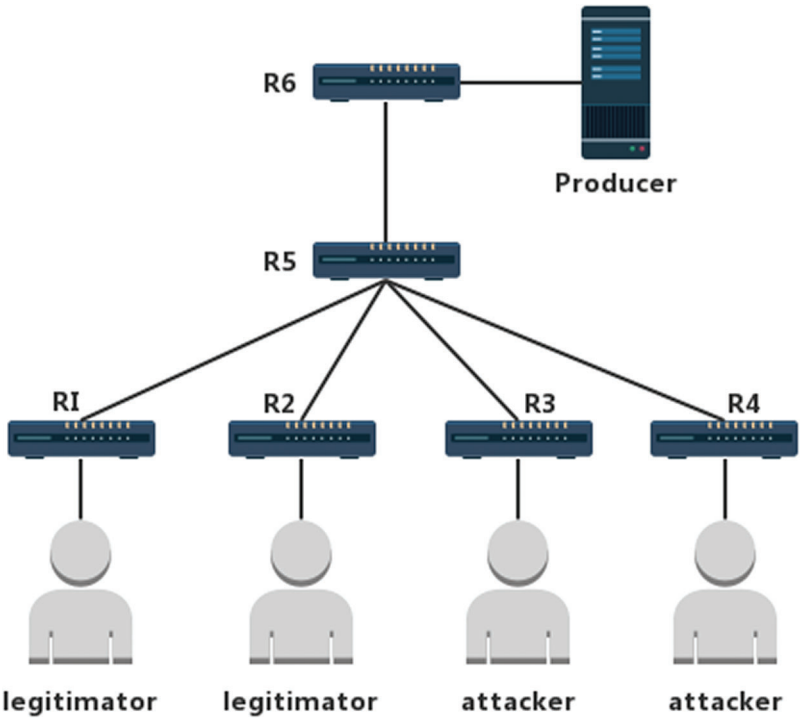

Figure 4: Experimental network topology diagram

5.2 Experimental Results and Analysis

5.2.1 Test Window Value Determination Experiment

The dynamic change of the window size is very important to improve the effectiveness of IFA detection. Under the condition of large fluctuation of network traffic, the distribution of interest packets under different information name prefixes in the network is more complex. If a small window is adopted, the traffic distribution under normal conditions cannot be correctly obtained, and misjudgment is likely to occur. However, when the fluctuation of network traffic is small, it indicates that the traffic distribution at this time is relatively simple, so the real traffic distribution can be obtained by using a small window. In the subsequent similarity test, the small fluctuation can be detected by the similarity test module. Therefore, we conducted the following experiments to determine the appropriate window size. In the experiment, the number of interest packets in the window is set to [50,100], [100,200], [200,400] respectively, and three groups of experiments are carried out to obtain which group of window sizes can effectively improve IFA detection and determine the appropriate window size.

Firstly, we compared the detection time. From Tab. 2, it can be concluded that under the three attack flows, the attack detection time of the larger window experimental group is more than that of the smaller window experimental group. However, in the experimental group with window size [200,400], the detection time has increased greatly, indicating that due to the large number of windows set and the large number of interest packets contained in each window, the detection time of the system has been obviously affected.

Secondly, we have carried out a comparative experiment to detect the misjudgment rate. In the experiment, the running time of the system is set to 40 s, and the number of network fluctuation phenomenon is set to 20 times. The fluctuation traffic is 125% of the normal request traffic, and the duration is 0.5 s. A total of 10 experiments were carried out, and the experimental results are shown in Fig. 5.

Table 2: Test time comparison

Proportion of attack traffic	Window Size [50,100]	Window Size [100,200]	Window Size [200,400]
25%	2.47	2.80	3.32
50%	1.42	1.69	2.05
75%	0.81	0.97	1.26

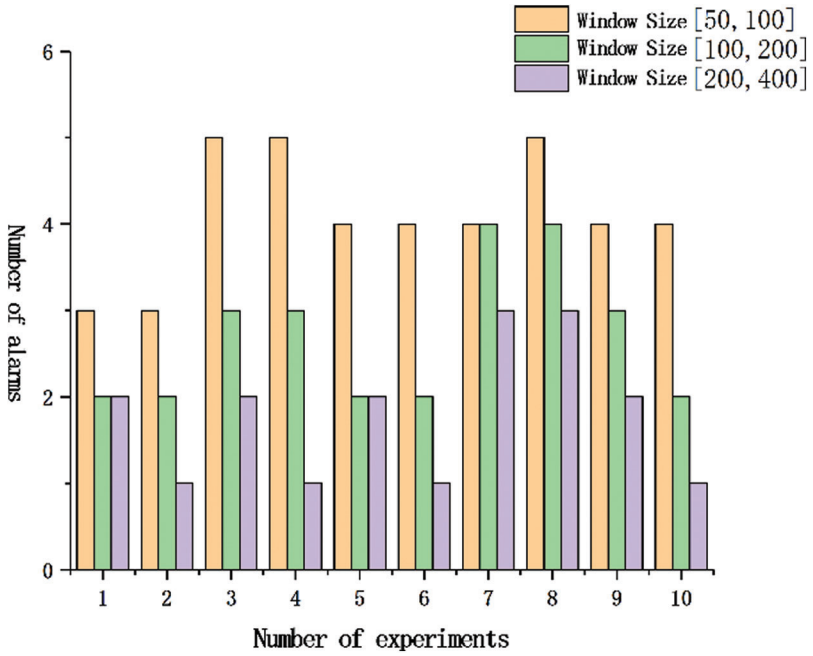

Figure 5: Comparison of alarm times of different window sizes

With the increase of window size, the detection range expands, and the number of interest packets contained in each window increases, thus the network traffic distribution at that time can be more accurately obtained, the misjudgment rate will also decrease. As can be seen from Fig. 5, there is no absolute negative correlation between the number of alarms and the window size, the overall trend in the figure is that the number of alarms decreases with the increase of the window size. That is, with the increase of window size, the misjudgment rate decreases and the detection performance gradually improves.

Considering the detection time and detection accuracy, we set the window size [100,200] as the window size used in the subsequent experiments, i.e., the value of the detection window is set to be that the smaller window contains 100 interest packets, and the larger window contains 200 interest packets.

5.2.2 Chi-Square Value Detection Experiment of Interest Packet Prefix

Here, it is necessary to determine the chi-square value of interest packet prefix under normal network traffic and IFA in NDN through experiments, to prove the sensitivity of the chi-square test method to traffic changes, and to determine the threshold of the chi-square test. The experiment increases or decreases the strength of IFA by controlling the rate values of normal network traffic and IFA traffic, thus obtaining the chi-square value changes under different attack intensities. The experimental results are shown in Fig. 6.

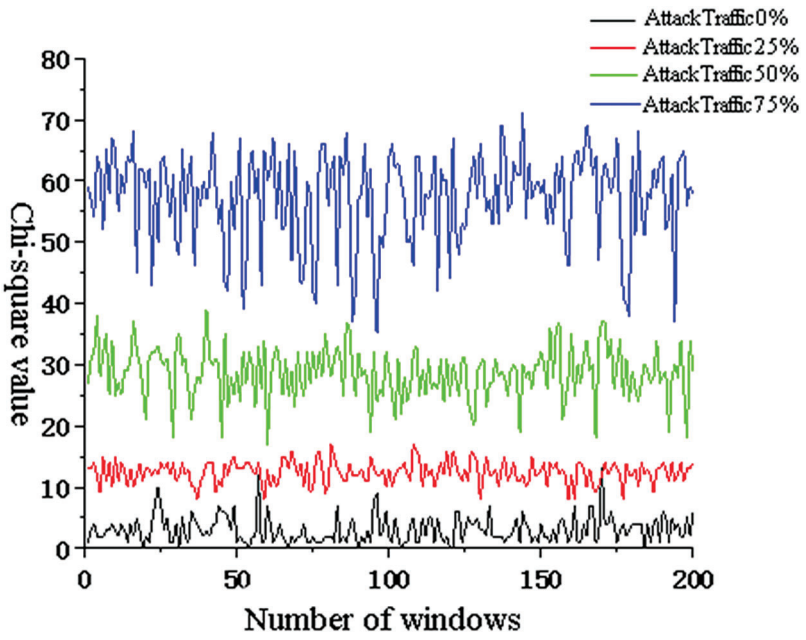

Figure 6: Change of chi-square value under different attack traffic

Through experiments, it can be found that the chi-square value appears an upward trend with the increase of attack traffic, and the chi-square value at 75% attack traffic can reach 10 times of the normal chi-square value, thus proving that the method has high sensitivity to IFA detection. Under normal network traffic, the chi-square value basically remains between 0 and 10. At 25% of the attack traffic, the chi-square value basically remains between 7 and 17. Under 50% of the attack traffic, the chi-square value basically remains between 18 and 38. At 75% of the attack traffic, the chi-square value basically remains between 40 and 70. From the results, we can judge that if the attack traffic is more than 25%, a suspected IFA is considered to have occurred. At the same time, the chi-square calculation threshold for IFA can be determined to be 10.

5.2.3 Similarity Test Experiment of Interest Packet Prefix

IFA detection method based on similarity can find the change of information name prefix distribution in two adjacent time windows. In the experiment, the similarity change under different attack intensities is obtained by controlling the rate value of normal network traffic and IFA traffic, thus proving the feasibility of similarity test method for IFA detection and determining the threshold of similarity test. The experiment starts to release attack traffic when the interest packet is sent to the 1000-th and ends the attack when it is sent to the 3000-th. The experimental results are shown in Fig. 7.

As can be seen from Fig. 7, with the increase of attack traffic, the value of similarity drops sharply, indicating that the addition of attack traffic leads to a change in the proportion of information name prefixes. After the attack lasted for a while, it can be found that the value of similarity shows an upward trend. Analysis shows that this is because attackers send a large number of interest packets with random suffixes for a certain prefix. With the continuous transmission of attack traffic, the proportion of information name prefixes in the whole network tends to stabilize again. When the attack is over, the value of similarity drops instantly and then gradually rises again, which is also because the disappearance of attack traffic causes the prefix ratio in network traffic to change.

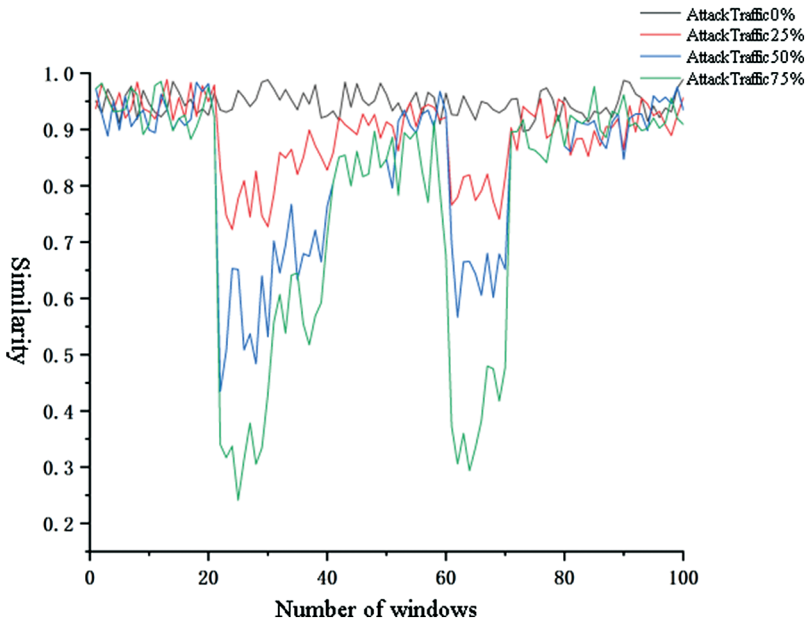

Figure 7: Similarity changes under different attack traffic

As can be found from Fig. 7, under normal network traffic, the lowest range of similarity is between 0.9 and 0.95. Under 25% attack traffic, the lowest similarity range is between 0.7 and 0.8. At 50% attack traffic, the lowest range of similarity is between 0.45 and 0.65. In 75% of attack traffic, the minimum similarity range is between 0.25 and 0.45. From this, it can be considered that 25% of the attack traffic occur as an attack. Correspondingly, the threshold value of similarity test can be set to 0.8.

5.2.4 IFA Detection Accuracy Comparison Experiment

In the experiment, a short network traffic fluctuation is designed to simulate the normal network fluctuation, thus detecting the accuracy of the method in this paper for IFA detection. In the experiment, a total of 20 network fluctuations were set in 40 s. Each network fluctuation traffic was 125% of the normal request traffic, and the fluctuation traffic lasted for 0.5 s. A total of 10 experiments were carried out, and the experimental results are shown in Fig. 8.

From Fig. 8, it can be found that compared with the detection results using the chi-square test method and similarity test method, the double-sided detection method proposed in this paper has the least number of alarms obtained in each experiment, obviously lesser than the former two detection methods. From this, it can be proved that the double-sided detection method proposed in this paper has a lower misjudgment rate than the single the chi-square test method or similarity test method.

5.2.5 Comparison of Attack Detection Time

In order to verify the superiority of IFA detection method proposed in this paper, this method is compared with the detection method based on information entropy adopted in literature [12]. The comparison results are shown in Tab. 3.

As can be seen from Tab. 3, the detection time of the two detection methods for IFA decreases with the increase of attack intensity, but under different attack intensities, the IFA detection method proposed in this paper can find attack behavior in a shorter time.

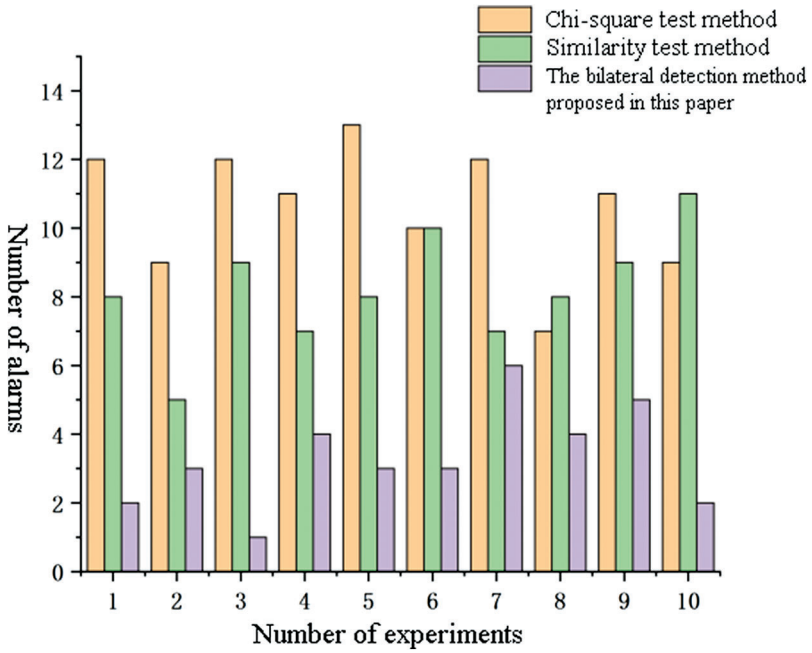

Figure 8: Comparison of alarm times of different detection methods

Table 3: Comparison of detection time

Proportion of attack traffic	Detection time (s) in literature [12]	The detection time (s) of the algorithm in this paper
25%	3.15	2.72
50%	2.07	1.65
75%	1.18	0.92

5.2.6 Packet Loss Rate Detection Experiment

Here, the packet loss rate refers to the proportion of all interest packets entering the PIT that have been replaced by the replacement policy or deleted after reaching the timeout. Under normal circumstances, the packet loss rate is basically maintained at nearly 0%. When IFA occurs, the packet loss rate will rise sharply due to the occupation of plenty of malicious interest packets on the PIT, showing a higher level. Therefore, the change of packet loss rate can effectively reflect the real-time impact of IFA on network traffic. In the experiment, three groups of experiments were carried out according to different attack intensities. The attack traffic was set to be released when the 500-th interest packet was sent, and the packet loss rate of each window was calculated. The experimental results are shown in Fig. 9.

As can be seen from Fig. 9, under a normal network circumstances, the packet loss rate is basically maintained at a level of nearly from 0% to 5%. During the IFA, due to the occupation of a huge amount of malicious interest packets on the pending interest table, legal interest packets cannot be queried and processed normally, so the packet loss rate shows a high level. Under the circumstances of 75% attack traffic, the packet loss rate reaches as high as 75%. In addition, in Fig. 8, we can also find that under the three kinds of attack traffic, the packet loss rate all increases briefly and then decreases rapidly, and then tends to be stable. This is because the interest backtracking method proposed in this paper enables PIT

entries generated by malicious interest packets to be satisfied by false data packets, thus releasing PIT space for normal users' interest requests. Therefore, the packet loss rate can gradually return to the normal level.

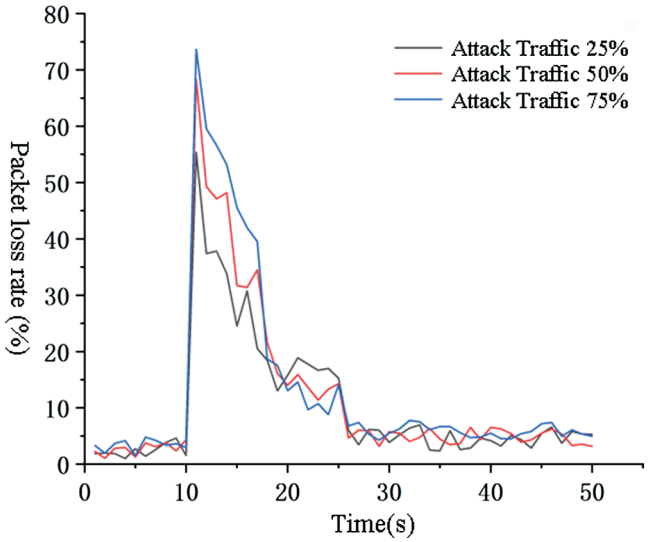

Figure 9: Change of packet loss rate of different attack traffic

6 Conclusion

According to the characteristic that NDN interest flooding attackers will continuously send plenty of interest packets that cannot be responded to, this paper proposes a bilateral detection model based on the chi-square test and similarity test. The method comprises a data acquisition and processing module, a flow anomaly detection module, a bilateral detection and IFA discrimination module, and an IFA alarm and defense module. The implementation process of the method proposed in this paper is as follows: Firstly, the structure of the PIT is expanded, a counter is added to record information, and the sliding window method is used to dynamically adjust the size of the detection window. The sensitivity of the chi-square test is used to preliminarily detect the attack behavior. Then, the similarity test method is used for further detection, thus realizing bilateral detection. Finally, based on ndnSIM simulation platform, the simulation experiment of this method is carried out, and the feasibility and effectiveness of this method are verified by the chi-square anomaly detection, similarity anomaly detection, accuracy detection, time detection and packet loss rate detection of IFA, respectively.

Funding Statement: This research is supported by The National Natural Science Foundation of China under Grant (No. 61672101), Beijing Key Laboratory of Internet Culture and Digital Dissemination Research (No. ICDDXN004) and Key Lab of Information Network Security of Ministry of Public Security (No. C18601).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] L. X. Zhang, D. Estrin, J. Burke, V. Jacobson, J. D. Thornton *et al.,* "Named Data Networking(NDN) Project, NDN-0001," 2010. [Online]. Available at: https://www.researchgate.net/publication/267821166_Named_data_networking_NDN_project.

[2] K. Lei, *Information center network and named data network*. Beijing, China: in Peking University Press, 60–68, 2015.

[3] L. X. Zhang, A. Afanasyev, J. Burke, C. Papadopoulos and C. S. Univ, "Named data networking," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 3, pp. 66–73, 2014.

[4] S. Choi, K. Kim, S. Kim and B. H. Roh, "Threat of DoS by interest flooding attack in content-centric networking," in *Proc. of the Int. Conf. on Information Networking*, Bangkok, Thailand, pp. 315–319, 2013.

[5] H. Dai, Y. Wang, J. Fan and B. Liu, "Mitigate DDoS attacks in NDN by interest traceback," in *Computer Communications Workshops*, Toronto, Ontario, Canada: IEEE, pp. 381–386, 2014.

[6] K. Wang, H. Zhou, Y. Qin, Jia Chen and H. Zhang, "Decoupling malicious interests from pending interest table to mitigate interest flooding attacks," in *Proc. IEEE Global Workshop MENS*, pp. 963–968, 2013.

[7] A. Afanasyev, P. Mahadevan, I. Moiseenko, E. Uzun and L. Zhang, "Interest flooding attack and counter-measures in named data networking," in *Proc. of the IEEE IFIP Networking Conf.*, New York, USA, pp. 1–9, 2013.

[8] J. Q. Tang, H. C. Zhou, Liu Ying and H. K. Zhang, "Mitigating interest flooding attack based on prefix identification in content-centric networking," *Journal of Electronics and Information*, vol. 36, no. 7, pp. 1735–1742, 2014.

[9] K. Wang, H. Zhou, Y. Qin and H. Zhang, "Cooperative-filter: Countering interest flooding events in named data networking," *Soft Computing*, vol. 18, no. 9, pp. 1803–1813, 2014.

[10] A. Compagno, M. Conti, P. Gasti and G. Tsudik, "Poseidon: mitigating interest flooding DDoS attacks in named data networking," in *Proc. of the IEEE Conf. on Local Computer Networks*, Sydney, Australia, pp. 630–638, 2013.

[11] A. Karami and M. Guerrero-Zapata, "A hybrid multi-objective RBF-PSO method for mitigating DoS attacks in Named Data Networking," *Neurocomputing*, vol. 151, no. 16, pp. 1262–1282, 2015.

[12] Y. Xin, Y. Li, W. Wei, W. Li and C. Xin, "A novel interest flooding attacks detection and counter-measure scheme in NDN," in *Proc. of the IEEE Global Communications Conf.*, Washington, USA, pp. 1–7, 2016.

[13] D. Goergen, T. Cholez, J. Françoi and T. Engel, "Security monitoring for content-centric networking," *Lecture Notes in Computer Science*, vol. 7731, pp. 274–286, 2013.

[14] J. Chen, Z. Zhou, Z. Pan and C. Yang, "Instance retrieval using region of interest based on features," *Journal of New Media*, vol. 1, no. 2, pp. 87–99, 2019.

[15] P. Cen, K. X. Zhang and D. S. Zheng, "Sentiment analysis using deep learning," *Journal on Artificial Intelligence*, vol. 2, no. 1, pp. 17–27, 2020.

[16] P. Gasti, G. Tsudik, E. Uzun and L. Zhang, "DoS and DDoS in named-data networking," in *ICCCN 2013*, Nassau, Bahamas, pp. 1–7, 2013.

[17] M. Luo, K. Wang, Z. Cai, A. Liu and Y. Li, "Using imbalanced triangle synthetic data for machine learning anomaly detection," *Computers, Materials & Continua*, vol. 58, no. 1, pp. 15–26, 2019.

[18] M. Luo, K. Wang, Z. Cai, A. Liu and Y. Li, "Using imbalanced triangle synthetic data for machine learning anomaly detection," *Computers, Materials & Continua*, vol. 58, no. 1, pp. 15–26, 2019.

[19] O. Aimbola, V. Suresh and H. Alex, "Entropy clustering approach for improving forecasting in DDoS attacks," in *12th IEEE Int. Conf. on Networking, Sensing and Control*, Taiwan, China, pp. 315–320, 2015.

[20] Y. T. Tang, "Research on a technology of mitigating DDoS attacks based on SDN," Guiyang: M.S. dissertation, Guizhou University, 2018.

[21] X. Shi, Y. Li, H. Xie, T. Yang, L. Zhang *et al.,* "An Openflow based load balancing strategy in SDN," *Computers, Materials & Continua*, vol. 62, no. 1, pp. 385–398, 2020.

[22] A. Afanasyev, I. Moiseenko and L. Zhang, *NdnSIM: NDN Simulator for NS-3*, NDN: Technical Report NDN-0005, 2012. [Online]. Available at: http://named-data.net/techreports.html.