Research on the Detection of Distributed Denial of Service Attacks Based on the Characteristics of IP Flow

Dongqi Wang, Guiran Chang, Xiaoshuo Feng, and Rui Guo

Northeastern University, Computing Centre 310-1,
Shenyang, China
zorrorily@163.com,
chang@neu.edu.cn,
{shuoner,happyachilles}@163.com

Abstract. IP Flow is classified into the Micro-flow and the Macro-flow, which provides a way of selecting proper features used to detect DDoS. Five abstracted features' capabilities of recognizing DDoS are analyzed through experiments. With these features as inputs, a neural network classifier is used to detect DDoS. Experiments' results show that these IP Flow based features can be very helpful to DDoS detection if they are put together.

Key words: DDoS; IP Flow; Detection; Neural Network

1 Introduction

Distributed Denial of Service (DDoS) attack presents a very serious threat to the internet. There are mainly two kinds of researches on DDoS detection: One is how to select the features to be tested. Another is looking for effective techniques to find out the abnormities shown by features during the attack, which is to be detected. Feature selecting methods can be divided into wrapper approach and filter approach [1]. Wrapper approach exploits machine learning algorithm to evaluate the goodness of features, and performance of learning algorithm is used as evaluation criterion. On the other hand, filter approach uses underlying characteristics of features as evaluation criterions. One example of wrapper approach is [2] in which Gavrilis Dimitris et al. used neural network classifier's performance as a genetic algorithm's evaluation criterion to identify a hypo-optimal feature set. The examples of filter approach are [3] and [4]. Xu Tu et al. [3] employed an underlying characteristic of network flow OWCD to detect DDoS, Cheng Guang et al. [4] get features through sampling measurement of statistics in high speed network. In order to find out the features' abnormities during an attack, researchers had employed many kinds of techniques, such as neural network, hidden Markov model, SVM, data mining [5-8] and so on.

In this paper, a filter approach which selects five statistical features from IP flow is proposed, and a neural network classifier is designed to find out the

abnormities shown by these features during an attack. The information such as protocols used by attackers, packet's size of an attack etc, as byproducts, can be gained during the generation of the features, which are very useful for filtering DDoS attack.

2 IP Flow Based Feature Selection

IP flow is composed of IP packets arriving one after another. As the basic data carrying unit of Internet, IP packet holds the upper layer's information and can be easily caught and handled. In the following part of this section IP flow will be divided into the Micro-Flow and the Macro-Flow and we are going to research how to select effective IP flow based detecting features.

2.1 The Concepts of Micro-Flow and Macro Flow

The Micro-Flow A Micro-Flow is a packet set who is composed of packets belonging to the same time interval of Internet, and all these packets have the same specific characteristics [9,10]. These same specific characteristics are called keys [10]. A group of commonly used keys are ¡Protocol, SrcIP, SrcPort, DestIP, DestPort¿. Protocol is the protocol used by the upper layer, SrcIP and SrcPort are the source IP address and the source port number separately. DestIP and SrcIP are the destination IP address and the destination port number separately.

The definition of Micro-Flow is helpful in two ways. First, each key group corresponds to one connection from SrcIP to DestIP, so keys can be used to describe DDoS connection. Second, a key group contains much information which can be used by routers and firewalls to operate each packet.

The Macro-Flow All the packets belonging to one time interval compose a set which is called the Macro-Flow. Macro-Flow is pooled by Micro-Flows.

The definition of Macro-Flow is helpful in two ways too. First, Detecting features can be formed on the base of Macro-Flow. Second, the information contained in the Macro-Flow is the complementarities to keys.

In experiments, we intercept network traffic by time interval i equals to 10s randomly. On one hand, in order to form the Micro-Flow based features, we classify packets by different keys. On the other hand, we abstract the Macro-Flow based features from the whole i directly.

2.2 IP Flow Based Features

Micro-Flow Based Features:

1. Average Number of Packets in Per Flow(ANPPF). Continuously and randomly generated "legitimate" IP are usually used in attack, so the generating speed of Micro-Flow is quickened, and the packet amount in per flow decrease. There are commonly 1-3 packets in per flow [9].

$$ANPPF = (\sum_{j=1}^{FlowNum} PacketsNum_j)/FlowNum \quad (1)$$

PacketsNumj is the quantity of packets in the jth flow of a time interval. FlowNum is the quantity of packets of the whole interval. Figure 1 shows the experimental comparison of ANPPF between normal traffic and DDoS traffic (110i-180i).The ANPPF of DDoS traffic which is near 1(attacking traffic is the mix of DDoS traffic generated by tfn2k and normal traffic of internet. ANPPF of tfn2k generating traffic is 1) differs from normal ANPPF (ruleless distribution) significantly.

2. Percentage of Correlative Flow (PCF). During attack, though the victim still has capability to reply to attacking packets' "requests", the replying packets can not get to the zombies, because the attacking IP addresses are faked. If flow x is from SrcIPx=A to DestIPx=B, and flow y is from SrcIPy=B to DestIPy=A, then we call flow x and y is a pair of Correlative Flow.

$$PCF = CFNum/FlowNum \quad (2)$$

CFNum is two times of the pairs of Correlative Flow. PCF represents the "there is going-out but no coming-back" characteristic of DDoS. As is shown in figure 2, when DDoS happens (110i-180i), PCF is near 0, while the PCF of normal traffic is 0.4-0.6. The difference between them is distinguishable.

3. One Direction Generating Speed (ODGS). Flow generating speed quickens when attack happens or busy time comes. In order to distinguish these two kinds of situations, ODGS is proposed.

$$ODGS = (FlowNum - CFNum)/interval \quad (3)$$

ODGS reflects the sudden increase of traffic when DDoS happens, and it also reflects the "there is going-out but no coming-back" characteristic of DDoS. Figure 4 gives the experimental comparison of ODGS between normal traffic (110i-180i) and abnormal traffic. ODGS' order of magnitude in normal traffic (102i) is much smaller than that in the abnormal traffic (104i).

4. Ports Generating Speed (PGS).

$$PGS = PortsNum/interval \quad (4)$$

PortsNum is the number of distinct port in one time interval. Some researchers select the size of port[2] as a detecting feature, while we find that many newly emerged services and applications (such as famous p2p application BT) use port

number bigger than 1024, so approach of [2] is not suitable anymore. Through deeper investigation, we realize that attackers continuously and randomly generate port too, so PGS is proposed. As is shown in figure 4, the PGS of normal traffic is not bigger than 200, while PGS of attacking traffic (110-180i)is over thousands.

Macro-Flow Based Feature:

Percentage of Abnormal Packets(PAP.) In order to increase the efficiency of attacking, attacking packets' content parts are usually unfilled or only filled with very few useless bytes (such as famous attacking tools tfn2k, trinoo). This kind of procedure results in the increase of abnormal small packets (for example, some TCP packets are only a little bigger than 40bytes, and UDP packets are only a little bigger than 28bytes). PAP presents this characteristic of DDoS attack by counting the percentage of abnormal packets in the one i(a Macro-Flow). Figure 5 is the comparison of PAP of normal traffic and abnormal traffic. As we can see, there is a significant change of PAP from near 0 to more than 0.9 when DDoS happens (110i-180i).

3 Neural Network Classifier

To detect the abnormities shown by features we choose to use a BP neural network with two layers. There are 5 neurons (because there are 5 features) in layer 1, and these neurons all use hyperbolic tangent sigmoid function as their transfer function. Layer 2 has only one neuron using logarithmic sigmoid function as the transfer function. The output of layer 2 is the output of the whole neural network, and its value is between 0 and 1. Value 0 presents the normal traffic, and value 1 presents the abnormal traffic. Mean squared error function is used as our error performance function, and Levenberg Marquardt (LM) algorithm is chosen to adjust weights and thresholds.

These five features mentioned earlier are used as the inputs of neural networks to do convergence test of different adjusting algorithms. In the experiment, weights and thresholds are randomly initialized, and inputs (PAP, ANPPF, PCF, ODGS, PGS) belonging to 50 intervals are applied. Table 1 is the comparison of four commonly used adjusting algorithms' convergence performance, and the values in table 1 are the average values of 50 tests to each algorithm. As is shown, LM algorithm suits us better from convergence rate to error performance on average level.

4 Detecting Experiments

On one hand, victims usually do not want to publicize the details of the attack they had suffered. On the other hand, currently, few data can describe the whole

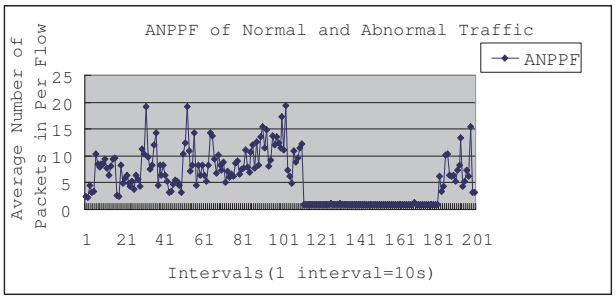

Fig. 1. ANPPF of Normal and Abnormal Traffic

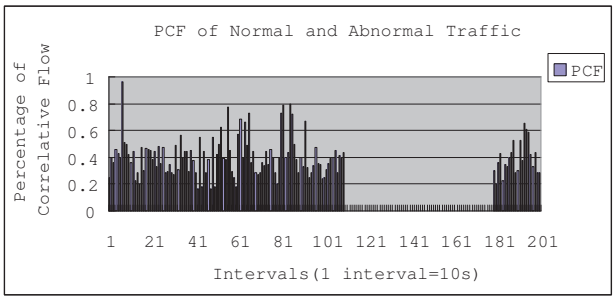

Fig. 2. PCF of Normal and Abnormal Traffic

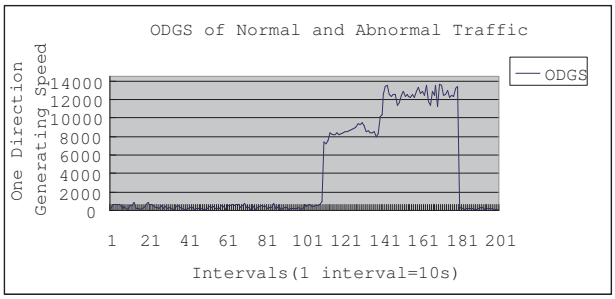

Fig. 3. ODGS of Normal and Abnormal Traffic

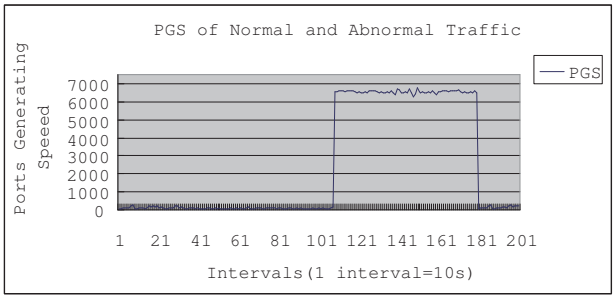

Fig. 4. PGS of Normal and Abnormal Traffic

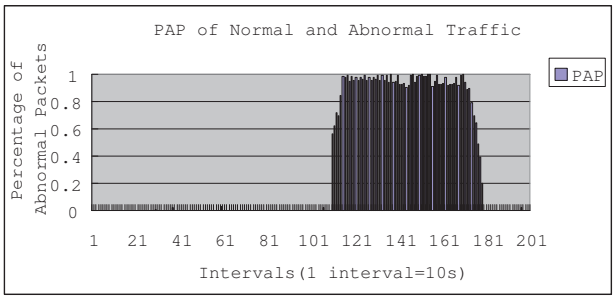

Fig. 5. PAP of Normal and Abnormal Traffic

Table 1. The Comparison of BP Algorithms

	BFGS	OSS	RPOP	LM
Average Convergence Time	0.9141s	1.079s	0.658s	0.6893s
Average Iteration Times	74.25	100	100	74.87
Average Error Performance	10^{-10}	10^{-7}	10^{-7}	10^{-18}

profile of DDoS attack. So we use both the UCLA's data set [11] and the data generated by our own simulation.

In our own simulation, 200 time intervals randomly intercepted from an outgoing router (kilo mega) of the central node are used as the normal samples of neural network. Another 200 time intervals are randomly intercepted when we use 50 hosts running tfn to attack one server behind this router, and these intervals containing both attack traffic and normal traffic are used as the attack samples.

UCLA's data is stored as pure text, and each row of the text is a packet composed of SrcIP, DestIP, SrcPort, DestPort, packet length and ACK (TCP packet) et. The attack launched in our own simulation is constant rate attack, so we choose the constant rate UDP attack data of UCLA as the attack samples.

The ROC curves in figure 6 and figure 7 show the sensitivity and accuracy of the neural network. A ROC curve is a plot with the false positive rate on the X axis and the true positive rate on the Y axis. The area below the curve reflects the sensitivity of the neural network. As we can see, the curve is close to both the Y axis and the point (0, 1) which means that we obtained low false positives and the classification capability is good.

5 Conclusion

In this paper we present five effective detecting features base on the characteristics of IP flow: PAP, ANPPF, PCF, ODGS and PGS. These five features can

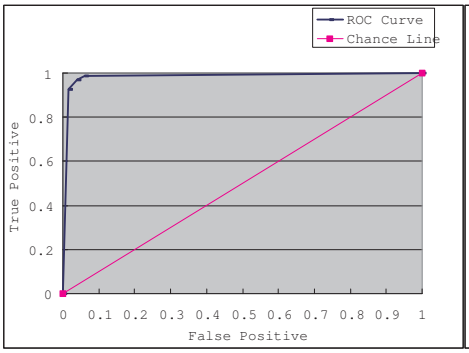

Fig. 6. Our Own Data ROC Curve

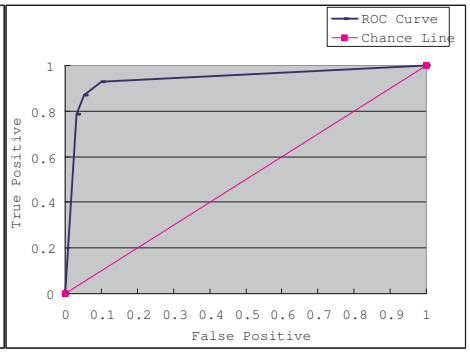

Fig. 7. UCLA ROC Curve

exploit the abnormalities during DDoS attack. Byproducts of features generation are helpful for filtering. We prove the capabilities of these five features through experimental comparison between their normal values and values in attack. A neural network using LM algorithm to adjust the rights and thresholds is used to detect abnormalities shown by features. Experiments using our own simulating data and UCLA data are carried out separately, and the experimental result are satisfying. As an effective approach, ours is easy to understand and easily to be carried out. It can be used as a part of common security tool in the core network, and it can also be attached a filter to in the edge network to relieve the threat of DDoS. Our future work is to improve the classifier's performance in real-time. Filtering Research based on this approach will also be done.

References

1. Jong Sou Park,Khaja Mohammad Shazzad,Dong Seong Kim: Toward modeling lightweight intrusion detection through correlation-based hybrid feature selection. In: Infromation Security and Cryptology. First SKLOIS Conference, CISC 2005, pp. 279–289. Beijing,China(2005)
2. Gavrilis Dimitris, Tsoulos Ioannis, Dermatas Evangelos: Feature Selection for Robust Detection of Distributed Denial-of-Service Attacks Using Genetic Algorithms. Methods and Applications of Artificial Intelligence. 3025, 276–281 (2004)
3. Xu Tu, He Da-ke: Time series Analysis for One-Way Connection Density of Network Flow. Journal of SiChuan University (Engineer Science Edition). 39(3),136–140 (2007)
4. Cheng Guang, Gong Jian, Ding Wei: A Real-Time Anomaly Detection Model Based on Sampling Measurement in a High-Speed Network. Journal of Software. 14(3),594-599 (2003)
5. Yang Xiang, Wanlei Zhou: Intelligent DDoS Packet Filtering in High Speed Network. In: 3rd International Symposium on Parallel and Distributed Processing and Applications, ISPA 2005, pp. 395-408. Nanjing, China (2005)

6. Xie Yi, Yu Shun-zheng: A novel model for detecting application layer DDoS attacks. In: First international on Computer and Computational Sciences, pp. 2,56–63. IEEE press,Hanzhou, China (2006)
7. Jungtaek Seo, Cheolho Lee, Taeshik Shonet al: A New DDoS Detection Model Using Multiple SVMs and TRA. In: Embedded and Ubiquitous Computing -EUC 2005 Workshops. EUC 2005 Workshops: UISW, NCUS, SecUbiq,USN, and TAUES, pp. 976–985. Nagasaki,Japan (2005)
8. Gao Neng,Feng Deng-guo,Xiang Ji:A Data-Mining Based DoS Detection Technique.Chinese Journal of Computers.29(6),944-951(2006)
9. Christos SiaterlisVasilis Maglaris: Detecting incoming and outgoing DDoS attacks at the edge using a single set of network characteristics. In: Proceedings 10th IEEE Symposium on Computers and Communications, pp. 469–475. Murcia, Spain (2005)
10. SUI SONG, LI LING: Flow-based Statistical Aggregation Schemes for Network Anomaly Detection. In: 2006 IEEE International Conference on Networking, Sensing and Control, pp. 786–791. IEEE Press, Ft.Lauderdate,FL,USA (2006)
11. Ucla.Sanitized UCLA CSD traffic traces, `http://lever.cs.ucla.edu/ddos/traces/`