DDoS Detection and Prevention Based on Artificial Intelligence Techniques

Boyang Zhang
International school
Beijing University of Posts and Telecommunications
Beijing, China
e-mail: zbykid@bupt.edu.cn

Tao Zhang
Automobile Research Institute
Geely
Ningbo, China
e-mail: oucseczt@126.com

Zhijian Yu
School of Economics and Management
Beijing University of Posts and Telecommunications
Beijing, China
e-mail: yuzhijian@bupt.edu.cn

***Abstract*—DDoS attacks have been the major threats for the Internet and can bring great loss to companies and governments. With the development of emerging technologies, such as cloud computing, Internet of things, artificial intelligence techniques, attackers can launch a huge volume of DDoS attacks with a lower cost, and it is much harder to detect and prevent DDoS attacks. Because DDoS traffic is similar to normal traffic. Some artificial intelligence techniques like machine learning algorithms have been used to classify DDoS attack traffic and detect DDoS attacks, such as Naive Bayes and Random forest tree. In the paper, we survey on the latest progress on the DDoS attack detection using artificial intelligence techniques and give recommendations on artificial intelligence techniques to be used in DDoS attack detection and prevention.**

Keywords-DDoS; detection; prevention; artificial intelligence; machine learning

I. INTRODUCTION

Distributed denial of service (DDoS) attack is an attack using multiple distributed resources against targets [1], [2], which will deprive authorized client from services. Attack targets include system resources, network bandwidth and other resources. DDoS attacks have been the most common and fatal attacks to the Internet. However, DDoS attack is hard to be detected, because attack traffic is similar to normal traffic in most case [1], [3]. DDoS attack is a major threat to availability, because it tries to prevent legitimate traffic between clients and servers. DDoS attacks can be huge volumes of traffic in short time, low volumes of traffic in long time, huge volumes of traffic in long time [4], of which the latter is hard to detect and prevent. With the development of cloud computing, Internet of things(IoT), artificial intelligence techniques, DDoS attacks have been changing and it becomes harder to detect and prevent DDoS attacks. Even IoT devices can be used to launch DDoS attacks, such as light bulbs. In Oct 2016, a Domain name system(DNS) service provider, Dyn, encountered a DDoS attack from an IoT botnet, which disrupted service of many famous American website[5].

II. DDOS DETECTION AND PREVENTION

A. DDoS Classifications and Features

DDoS attack can multiply the power of attack and have a large impact on the victims. IP spoofing and flooding attack are two particular DDoS attacks. In IP spoofing, attackers impersonate as a trusted source [6]. While in flooding attack, attackers send too many packets to disrupt the availability of services.

There are three kinds of flooding dos attacks [7], [8]:

TCP flood attack. Attackers will send too many TCP connection requests without acknowledging the SYN-ACK response server to the target victim server. The server will be down because these half connections consume too many system resources. TCP flood DDoS attack is one of the most commonly used attacks.

ICMP flood attack (smurf attack). ICMP flood attack is to send ICMP packets with a spoofed IP source address. The owner of the spoofed IP address will be the potential victim, because it will be destination of many ICMP responses and be flooded.

UDP flood attack. UDP flood attack is to send too much UDP packets to different port of a target in random way.

DNS amplification attack. DNS amplification attack is an attack that attackers spoofs the source address of the victim. The attacker sends a small request to the DNS server and DNS server will reply with a large responses.

B. DDoS Detection and Prevention

Most common mechanisms to detect and prevent DDoS include attack prevention, attack detection, and attack reaction [1]. It is hard to detect DDoS attacks, because it is hard to differentiate the attack traffic and normal traffic. When detecting DDoS attacks, the first step is to detect the abnormality from traffic. And machine learning classification methods can be used to differentiate the good

978-1-5090-6352-9/17/$31.00 ©2017 IEEE

and bad packets. Packets that are classified as attack traffic will be dropped. Some features to detect DDoS attack are number of packets, average of packet size, time interval variance, packet size variance, number of bytes, packet rate and bit rate [9].

C. Artificial Intelligence Techniques

Typical artificial intelligence techniques include machine learning, speech Recognition, and natural language processing [10]. Machine learning algorithms have been applied to DDoS detection and defense, anomaly detection in particular. Most frequently used techniques are Naive Bayes, neural network, and support vector machine [1].

1) Bayes classification.

Bayes classifiers are commonly used machine learning classifying methods based on the application of Bayes theorem. Naive Bayes models include simple Bayes and independence Bayes[11].

2) Artificial neuron network.

Artificial neuron networks are composed of artificial neurons that can communicate with other neurons [12]. Artificial neuron networks are to solve problems as the brain works and have been used in different fields.

3) Support vector machine.

Support vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis [13]. Support vector machines can efficiently perform linear and non-linear classification.

D. Trend of DDoS Attacks

According to Verisign's report [14], we can summarize the DDoS trends as follows.

1) Largest volumetric attack and highest intensity flood.

The number of attacks is decreasing, however, volume, peak attack size and speed are all becoming larger. 36% of attack size peaked over 5 Gbps.

2) Multi-vector DDOS attacks are the norm.

30% of DDOS attacks in Q1 2017 uses more than three attack types and 6% utilizes more than 5 types. The more attack types used, it will be more difficult to be detected. TCP and UDP based attacks are two main attack types, including TCP SYN and TCP RST floods.

With the increase of DDoS volume, peak attack size and speed, it will be more challenging to detect and mitigate DDoS attacks.

III. Applications of Artificial Intelligence in DDoS Attack and Prevention

Xuan [1] proposes DeepDefense, a deep learning based DDoS detection approach, to improve the performance of DDoS attack detection. He formulate the DDoS detection problem as a sequence classification problem and transform the packet based detection to window based detection. The DeepDefense is composed of CNN, RNN(recurrent neural network) and fully connected layers. RNN can learn features better than other machine learning methods, especially longer historical features. LSTM and GRU are used to eliminate scaling issues when RNN is used to trace the history from previous packets. RNN also has a better performance in generalization than random forest does.

Heish [15] proposes a DDoS detection system based on neural network that is composed of five phase, packet collector, Hadoop HDFS, format converter, data processor and neural network detection module. They choose Hadoop distributed file system to store traffic data, use big data platform integrated the neural network to detect DDoS attacks by seven parameters. The detection system can analyze high velocity and volume network traffic and neural network can identify packet features efficiently.

Berral [16] extends a framework proposed by zhang in 2006 to detect and prevent DDoS flood attacks based on machine learning. All nodes in the framework have the ability of learning independently and can react according to different situations. The well known cumulated sum algorithm is used to detect huge traffic volume. Classifiers and detectors are used to distinguish pattern of normal traffic, such as Naive Bayes. Each node has the algorithm that compares the accumulated sum of means for each time unit with a characteristic threshold to classify message. The mechanism can stop and avoid DDoS flood attacks or abuse at early time.

Kiruthika [17] proposes a DDoS attack detection and mitigation model using machine learning algorithm. The model is composed of online monitoring system (OMS), spoofed traffic detection module and interface based rate limiting algorithm. OMS uses automated tools and scripts to monitor the degradation and provide DDoS impact measurements. The spoofed traffic detection module incorporates hop count inspection algorithm to check the authenticity of incoming packet. He constructs legitimate records with IP and hop count to detect potential attacks. Hop count inspection algorithm is to check the authenticity of packet. HCF-SVM is trained and updated with source IP and respective hop count. The performance of the model is better than random forest and decision tree when classifying instance.

Zhao [18] develops a DDoS detection system based on neural network and implements in Apache Hadoop cluster and HBase system. The system has a neural network architecture that has the ability of adapting to new types of DDoS attacks. A Hadoop and HBase cluster is setup to process huge traffic, then a neural network model is designed to detect DDoS attacks. The neural network selects parameters from Hadoop and HBase cluster module, such as CPU usage, packet size and total number of TCP connections. He chooses the multi-factor detection approach instead of single factor detection approach to detect DDoS attack, which can improve the performance of detection.

Meitei [19] design a model of system based on ANN and the packet header statistical information to detect DDoS DNS amplification attack. They classify the DNS traffic using machine learning classification algorithms, including decision tree, multi layer perception (MLP), Naive Bayes and support vector machine(SVM). Then choose decision tree as machine learning classification models for its best

performance. The selection approach is attributed based, optimal features are extracted from attributed selection algorithms like information gain, gain ratio and chi square. The feature parameters selected are inter packet arrival time, probability of occurrence of one IP address, answer, additional and authority of resource record, minimum packet size, average packet size and maximum packet size.

Ndibwile [20] proposes a simple network architecture that makes use of real web server, Bait server, and Decoy web servers to distinguish DDoS traffic from normal traffic. The architecture use a customized Intrusion Prevention System(IPS) at the network gateway that use rules generated by random tree machine learning algorithm through supervised learning. Decision tree is chosen to classify malicious traffic from normal traffic. Random tree machine learning algorithm using labeled datasets is used to avoid false positive traffic.

Fouladi [21] proposes a stand-alone frequency analysis method to detect DDoS attack. They use both Discrete Wavelet Transform (DWT) and Discrete Fourier Transform (DFT) to extract features from large volumes of traffic. The experiment shows that DWT has a higher accuracy and resolution information. Then Naive Bayes classification is used to classify DDoS attack traffic and normal traffic, which has a better performance than a simple thresholding classifier. The accuracy of detection can be increased when features extracted from DWT and DFT are combined.

Ramadhan [22] designs a TCP flood DDoS detection system which uses Artificial Immune System(AIS). The system is composed of two main component, collection data and analysis data. In the AIS, there are many algorithms based on human immune functions, principles and models can be applied to detect attacks, such as Dendritic Cell Algorithm(DCA). The four phase of DCA are pre-processing and initialization phase, detection phase, context assessment phase, and classification phase. The system presents DDoS attack by danger signals. Danger signals has been predefined as danger, safe, PAMA and inflammation. PAMA is a confident indicator of an abnormality and different signals indicate different kinds of attacks.

Peraković [23] develops a detection and classification model system based on artificial neural network(ANN) architecture to detect DDoS attack. In the developed ANN model, traffic are classified as four kinds, class-DNS DDoS attack traffic, chargen DDoS attack traffic, UDP DDoS attack traffic and normal traffic. Parameters used in detection of DDoS are source IP address, destination IP address, protocol and packet length. Because of the correspondence of the features of UDP DDoS attack and those of normal traffic, the accuracy in detection and classification of UDP DDoS attacks is a little lower.

Anomaly based detection technique models the behavior of normal traffic to distinguish attack traffic from normal while the signature based detection uses pattern matching to compare data instance with the signature already stored in the database.

Machine learning based classifiers are experts in finding out patterns in the dataset with the help of features used to describe the data. Machine learning techniques can provide decision aids for the analysts and can automatically generate rules to be used for network intrusion detection system.

Classifiers are tools that classify data based on specified features or patterns present in that data. Some of the worth noticing works in the field of DDoS detection includes the work of Gil and Poletto, in which they assume that packet rates between two hosts are proportional during normal operation. The work make use of a dynamic tree structure known as Multi Level Tree for Online packet statistics structure for monitoring packet rates for each IP address.

Robinson [24] uses different kinds of machine learning algorithms to classify DDoS attacks and ranks the performance of these algorithms, such as Naive Bayes, RBF network, multi layer perception, Bayesnet, IBK, J48, voting, Bagging+Random Forest, Random Forest, and Adaboost+Random Forest. The datasets used are DARPA 2000 intrusion detection scenario-specific datasets, CAIDA dataset 2007 and LLS-DDoS 1.0 scenario one. The ranking includes four phases, feature extraction, normalization, classification and evaluation, ranking of algorithms. In all three datasets, Adaboost with Random forest performs best while the random forest is much better than other algorithms as the single classifier category.

TABLE I. APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN DDOS DETECTION

Authors	Approaches used	Published Year
Berral [16]	machine learning, Naive Bayes	2008
Kiruthika [17]	machine learning, SVM	2014
Zhao [18]	neural network, Hadoop	2015
Ndibwile [20]	machine learning	2015
Robinson [24]	machine learning	2015
Heish [15]	neural network, Hadoop	2016
Meitei [19]	ANN	2016
Fouladi [21]	machine learning, Naive Bayes	2016
Ramadhan [22]	artificial immune system	2016
Peraković [23]	ANN	2016
Xuan [1]	deep learning, CNN, RNN	2017

From Table I, we can know exactly which kind of approaches these published papers used and their published year. And we can learn that more researchers have been using ANN in 2016 and 2017, while they mainly use machine learning algorithms before 2015. The change of DDoS attacks trend and the maturity of ANN may contribute most to the change. There are also some papers that use both neural network and Hadoop. Because of the larger volumes and peak size of DDoS attacks, Hadoop is necessary to analyze and process large volume of

information efficiently. To detect and prevent DDoS attacks efficiently, combination of ANN, Hadoop and other emerging technologies may be the best choice.

IV. CONCLUSION

DDoS attacks have been the major threats for the Internet and can bring great loss to companies and government. With the development of emerging technologies, such as cloud computing, Internet of things, artificial intelligence techniques, attackers can launch DDoS attacks with a low cost, and it becomes much harder to detect and prevent DDoS attacks. Some artificial intelligence techniques like machine learning algorithms have been used to classify DDoS attack traffic and detect DDoS attack, such as Naive Bayes and Random forest tree. In the paper, we survey on the latest progress on the DDoS attack detection using artificial intelligence techniques. Features that can be used to detect DDoS attack, such as number of packets, average of packet size, time interval variance, packet size variance, number of bytes, packet rate and bit rate. Among those artificial intelligence techniques, we recommend that random forest tree and Naive Bayes are used to classify malicious traffic and normal traffic for their better performance. Multi machine algorithms can be combined to detect DDoS attacks, which will have a better accuracy and performance.

REFERENCES

[1] X. Yuan, C. Li and X. Li, "DeepDefense: Identifying DDoS Attack via Deep Learning," 2017 IEEE International Conference on Smart Computing (SMARTCOMP), Hong Kong, 2017, pp. 1-8. doi: 10.1109/SMARTCOMP.2017.7946998

[2] M. Guri, Y. Mirsky and Y. Elovici, "9-1-1 DDoS: Attacks, Analysis and Mitigation," 2017 IEEE European Symposium on Security and Privacy (EuroS&P), Paris, France, 2017, pp. 218-232. doi: 10.1109/EuroSP.2017.23

[3] R. F. Fouladi, C. E. Kayatas and E. Anarim, "Frequency based DDoS attack detection approach using naive Bayes classification," 2016 39th International Conference on Telecommunications and Signal Processing (TSP), Vienna, 2016, pp. 104-107. doi: 10.1109/TSP.2016.7760838

[4] C. J. Hsieh and T. Y. Chan, "Detection DDoS attacks based on neural-network using Apache Spark," 2016 International Conference on Applied System Innovation (ICASI), Okinawa, 2016, pp. 1-4. doi: 10.1109/ICASI.2016.7539833

[5] Zejun Ren, Xiangang Liu, Runguo Ye, Tao Zhang, "Security and Privacy on Internet of Things," 2017IEEE 7th International Conference on Electronics Information and Emergency Communication (ICEIEC 2017) Macau&Shenzhen, 2017.

[6] B. S. Kiruthika Devi, G. Preetha, G. Selvaram and S. Mercy Shalinie, "An impact analysis: Real time DDoS attack detection and mitigation using machine learning," 2014 International Conference on Recent Trends in Information Technology, Chennai, 2014, pp. 1-7. doi: 10.1109/ICRTIT.2014.6996133

[7] Irom Lalit Meitei, Khundrakpam Johnson Singh, and Tanmay De. 2016. Detection of DDoS DNS Amplification Attack Using Classification Algorithm. In Proceedings of the International Conference on Informatics and Analytics (ICIA-16). ACM, New York, NY, USA, Article 81, 6 pages. DOI: https://doi.org/10.1145/2980258.2980431

[8] G. Ramadhan, Y. Kurniawan and Chang-Soo Kim, "Design of TCP SYN Flood DDoS attack detection using artificial immune systems," 2016 6th International Conference on System Engineering and Technology (ICSET), Bandung, 2016, pp. 72-76. doi: 10.1109/ICSEngT.2016.7849626

[9] C. J. Hsieh and T. Y. Chan, "Detection DDoS attacks based on neural-network using Apache Spark," 2016 International Conference on Applied System Innovation (ICASI), Okinawa, 2016, pp. 1-4. doi: 10.1109/ICASI.2016.7539833

[10] Xiuquan Li and Tao Zhang, "An exploration on artificial intelligence application: From security, privacy and ethic perspective," 2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), Chengdu, 2017, pp. 416-420. doi: 10.1109/ICCCBDA.2017.7951949

[11] Rish, I. "An empirical study of the naive Bayes classifier." Journal of Universal Computer Science 1.2(2001):127.

[12] Ahmad, Iftikhar, A. B. Abdullah, and A. S. Alghamdi. "Artificial neural network approaches to intrusion detection: a review." Wseas International Conference on Telecommunications and Informatics World Scientific and Engineering Academy and Society (WSEAS), 2009:200-205.

[13] Tong, Simon, and D. Koller. "Support vector machine active learning with applications to text classification." Journal of Machine Learning Research 2.1(2001):45-66.

[14] Verisign. VERISIGN DISTRIBUTED DENIAL OF SERVICE TRENDS REPORT(2017). https://www.verisign.com/assets/report-ddos-trends-Q12017.pdf

[15] C. J. Hsieh and T. Y. Chan, "Detection DDoS attacks based on neural-network using Apache Spark," 2016 International Conference on Applied System Innovation (ICASI), Okinawa, 2016, pp. 1-4. doi: 10.1109/ICASI.2016.7539833

[16] Josep L. Berral, Nicolas Poggi, Javier Alonso, Ricard Gavaldà, Jordi Torres, and Manish Parashar. 2008. Adaptive distributed mechanism against flooding network attacks based on machine learning. In Proceedings of the 1st ACM workshop on Workshop on AISec (AISec '08). ACM, New York, NY, USA, 43-50. DOI=http://dx.doi.org/10.1145/1456377.1456389

[17] B. S. Kiruthika Devi, G. Preetha, G. Selvaram and S. Mercy Shalinie, "An impact analysis: Real time DDoS attack detection and mitigation using machine learning," 2014 International Conference on Recent Trends in Information Technology, Chennai, 2014, pp. 1-7. doi: 10.1109/ICRTIT.2014.6996133

[18] T. Zhao, D. C. T. Lo and K. Qian, "A Neural-Network Based DDoS Detection System Using Hadoop and HBase," 2015 IEEE 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems, New York, NY, 2015, pp. 1326-1331. doi: 10.1109/HPCC-CSS-ICESS.2015.38

[19] Irom Lalit Meitei, Khundrakpam Johnson Singh, and Tanmay De. 2016. Detection of DDoS DNS Amplification Attack Using Classification Algorithm. In Proceedings of the International Conference on Informatics and Analytics (ICIA-16). ACM, New York, NY, USA, Article 81 , 6 pages. DOI: https://doi.org/10.1145/2980258.2980431

[20] J. D. Ndibwile, A. Govardhan, K. Okada and Y. Kadobayashi, "Web Server Protection against Application Layer DDoS Attacks Using Machine Learning and Traffic Authentication," 2015 IEEE 39th Annual Computer Software and Applications Conference, Taichung, 2015, pp. 261-267. doi: 10.1109/COMPSAC.2015.240

[21] R. F. Fouladi, C. E. Kayatas and E. Anarim, "Frequency based DDoS attack detection approach using naive Bayes classification," 2016 39th International Conference on Telecommunications and Signal Processing (TSP), Vienna, 2016, pp. 104-107. doi: 10.1109/TSP.2016.7760838

[22] G. Ramadhan, Y. Kurniawan and Chang-Soo Kim, "Design of TCP SYN Flood DDoS attack detection using artificial immune systems," 2016 6th International Conference on System Engineering and Technology (ICSET), Bandung, 2016, pp. 72-76. doi: 10.1109/ICSEngT.2016.7849626

[23] D. Peraković, M. Periša, I. Cvitić and S. Husnjak, "Artificial neuron network implementation in detection and classification of DDoS traffic," 2016 24th Telecommunications Forum (TELFOR), Belgrade, 2016, pp. 1-4. doi: 10.1109/TELFOR.2016.7818791

[24] R. R. R. Robinson and C. Thomas, "Ranking of machine learning algorithms based on the performance in classifying DDoS attacks," 2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS), Trivandrum, 2015, pp. 185-190. doi: 10.1109/RAICS.2015.7488411.