Research Plan

Student Name: Xinnan SHEN

Student ID: 1051380

1. Motivation for research question

My research question is: Is the Support Vector Machine (SVM) algorithm effective in predicting distributed denial of service (DDoS) attacks in telephony systems?

Nowadays, cyber security is a relatively common topic. There are many Internet attacks which may affect the server and the client, and one of the most serious attacks is DDoS attacks, which has bad influence on our computer and network security [1]. A distributed denial of service (DDoS) attack refers to a type of Internet attack that many attackers use multiple computers to attack the target machine simultaneously, which has a severe impact on users and the host machine [2]. If a telephony system gets attacked, the system will become unavailable and users will not be able to use the system to make phone calls [3-4]. This type of attack differs from other attacks in that it uses distributed techniques to attack target machines, which is much more harmful to the machine [5].

To avoid this type of serious Internet attack, we need to take action to detect and predict DDoS attacks. Although the traditional method can help detect DDoS attacks to some extent, the results are far from satisfactory. Besides, it is difficult to predict DDoS attacks only by using traditional methods [5]. Thus, more advanced techniques should be used to predict DDoS attacks, such as machine learning methods. In this project, I will use a machine learning algorithm called SVM to predict the DDoS attacks and find out whether it is suitable to predict the DDoS attacks. The research plan is organized as follows. First, section 2 presents the identification of research question. Then, section 3 shows the investigation method used in this project. Besides, section 4 demonstrates the analysis method of the result. Finally, section 5 mentions the contribution of this research and summarizes this research plan.

2. Identification for research question

Many related studies have been done to defend against DDoS attacks. Some of them have focused on DDoS attack detection, while few of the studies have mentioned the methods of predicting DDoS attacks. When dealing with DDoS attacks, most people use traditional methods to detect it. This method is useful to some extent, but it may cause some problems. The data of DDoS attacks are quite similar to that of ordinary programs, and it is difficult to distinguish between them only by using traditional methods [5-6]. Thus, some people have chosen to use machine learning algorithms to detect the DDoS attacks so that the detection process will become much more efficient. Previous studies have shown that machine learning algorithms are effective in detecting DDoS attacks [7]. However, few people have used machine learning methods to predict DDoS attacks so far.

In my research, I will focus on predicting DDoS attacks in telephony systems by using the SVM algorithm to solve my research question. The telephony system is a widely used system and people can make phone calls with the help of it. If a telephony system gets attacked, both the client and the server will be affected [8]. Thus, it is essential to protect the telephony system from DDoS attacks. While traditional methods can help detect DDoS attacks, using machine learning algorithms like SVM algorithm is much more effective in detecting and predicting DDoS attacks.

3. Investigation Method

In order to enhance the effectiveness of predicting DDoS attacks, I will use the Support Vector Machine (SVM) algorithm to analyze the possible features before the attack happens. SVM is a type of machine learning algorithm that can help classify data [9]. The features of ordinary dataflow and those of DDoS attack dataflow are quite similar, so we need to use machine learning algorithms like SVM to help distinguish between them. By using SVM, we are more likely to tell the difference between them based on the features we have observed.

Specifically, I will use the simulation method to detect the effectiveness of the SVM algorithm. In this project, I will first collect data from the existing DDoS attacks which have already happened before. The data that will be used in the experiment include features collected before DDoS attacks occur and ordinary program features. These data will be used as the training data to train the machine learning model. After completing the training process, I will then use extra data to evaluate the model. The data for evaluation are gathered from the result of simulation. I will simulate a DDoS attack on a virtual platform and get the features before the attack occurs. Meanwhile, the features of the ordinary programs will also be collected to evaluate the model. These data will be used to show whether the SVM algorithm can help predict DDoS attacks. The evaluation process is just using the gathered data as the input of the model and observe whether the prediction is correct. The prediction of the model refers to whether the dataflow is DDoS attack data.

Some previous studies have shown that the investigation methods mentioned above are appropriate. In [10], the author has used the SVM algorithm to help detect DDoS attacks, which shows that the SVM algorithm can distinguish the dataflow between ordinary programs and DDoS attack data. Furthermore, Daneshgadeh Salva et al. [11] have utilized the simulation method to collect data that are similar to those in real DDoS attacks. They have built a virtual environment to collect the most useful features of the DDoS attack and the ordinary programs, and the results are quite satisfactory. Thus, using the way of simulation and SVM algorithm is suitable in this project to conduct the experiment.

4. Analysis Method

After using the SVM algorithm to analyze the features of dataflow, the model will have the prediction result of the given data. To evaluate the model, I will classify the result of prediction into four different classes, namely true positives, false positives, true negatives and false negatives. True positives refer to the correctly predicted DDoS attacks. False positives refer to the occasion when the dataflow is predicted as a DDoS attack while actually it isn’t. True negatives refer to the correctly predicted normal dataflow. False negatives refer to the occasion when the real DDoS attacks have not been predicted [12].

Based on the four classes of results, I will use some standards to evaluate the model. The first standard is the accuracy of DDoS attack prediction, which refers to the ratio of correctly predicted dataflow over the total amount of dataflow [12]. In addition, the precision value is also an important standard. It can be calculated as the percentage of true positives over the total number of predicted DDoS attacks [12]. Furthermore, I will also use recall as an evaluation metric, and its value is the proportion of true positives over the number of real DDoS attacks [12]. These three values can be helpful in model evaluation, but in some cases models with high precision have a low recall and vice versa. Thus, we need to consider both the value of precision and recall. A good way to do that is calculating the weighted average of them, and the value is called F1 score. F1 score is a good standard to evaluate whether this system can effectively predict DDoS attacks, while accuracy can be used for the entire system evaluation [12].

When the evaluation metrics have already been calculated, the system can be evaluated. If the F1 score is high enough, it means the SVM algorithm can help predict DDoS attacks. However, if the F1 score is incredibly low, it means the SVM algorithm is not useful in the field of predicting DDoS attacks. In previous studies [13-15], the authors have also used the F1 score to detect the effectiveness of the DDoS prediction model, which shows that the evaluation standard is satisfactory.

5. Research contribution

This research has focused on predicting DDoS attacks by using the SVM algorithm. Most previous studies have only focused on the detection of DDoS attacks, and few of them have used machine learning techniques to help us defend against DDoS attacks in the telephony system. In my research, I have used machine learning methods to help predict DDoS attacks, which can help us a lot in defending against DDoS attacks in telephony systems. If we use the technique to predict the occurrence of DDoS attacks, we will be more likely to protect the client and server from the serious DDoS attacks. By using the SVM algorithm, the DDoS attacks can be easily predicted in an early stage. Therefore, people can take action to protect the machine and minimize the influence caused by the DDoS attacks. In general, this research has narrowed the gap of predicting DDoS attacks by using the SVM algorithm.

References

[1] Idhammad Mohamed; Afdel Karim; Belouch Mustapha. (2018). Semi-supervised machine learning approach for DDoS detection. APPLIED INTELLIGENCE, 48(10), 3193-3208, SPRINGER. doi:10.1007/s10489-018-1141-2.

[2] Zargar S.T.; Joshi J.; Tipper D. (2013). A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks. IEEE Communications Surveys & Tutorials, 15(4), 2046-2069, IEEE. doi:10.1109/SURV.2013.031413.00127.

[3] Cadet Frantz; Fokum Daniel T. (2016). Coping with denial-of-service attacks on the IP telephony system. SoutheastCon 2016, 1-7, IEEE. doi:10.1109/SECON.2016.7506691.

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

[5] Bing Wang; Yao Zheng; Wenjing Lou; Hou Y.T. (2014). DDoS Attack Protection in the Era of Cloud Computing and Software-Defined Networking. 2014 IEEE 22nd International Conference on Network Protocols, 624-629, IEEE, doi:10.1109/ICNP.2014.99.

[6] Sahi, A.; Lai, D.; Li, Y.; Diykh, M. (2017). An Efficient DDoS TCP Flood Attack Detection and Prevention System in a Cloud Environment. IEEE Access, 6036-6048, IEEE. doi:10.1109/ACCESS.2017.2688460.

[7] Pelloso Mateus; Vergutz Andressa; Santos Aldri; Nogueira Michele. (2018). A Self-Adaptable System for DDoS Attack Prediction Based on the Metastability Theory. 2018 IEEE Global Communications Conference (GLOBECOM), 1-6, IEEE. doi:10.1109/GLOCOM.2018.8647934.

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

[9] Cortes, Corinna; Vapnik, Vladimir N. (1995). "Support-vector networks" (PDF). Machine Learning. 20 (3): 273–297. CiteSeerX 10.1.1.15.9362. doi:10.1007/BF00994018.

[10] Khuphiran Panida; Leelaprute Pattara; Uthayopas Putchong; Ichikawa Kohei; Watanakeesuntorn Wassapon. (2018). Performance Comparison of Machine Learning Models for DDoS Attacks Detection. 2018 22nd International Computer Science and Engineering Conference (ICSEC), 1-4, IEEE. doi:10.1109/ICSEC.2018.8712757.

[11] Daneshgadeh Salva; Ahmed Tarem; Kemmerich Thomas; Baykal Nazife. (2019). Detection of DDoS Attacks and Flash Events Using Shannon Entropy, KOAD and Mahalanobis Distance. 2019 22nd Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN), 222-229, IEEE. doi:10.1109/ICIN.2019.8685891.

[12] Chenguang Wang; Jing Zheng; Xiaoyong Li. (2017). Research on DDoS Attacks Detection Based on RDF-SVM. 2017 10th International Conference on Intelligent Computation Technology and Automation (ICICTA), 161-165, IEEE. doi:10.1109/ICICTA.2017.43.

[13] Alsadhan Abeer Abdullah; Hussain Abir; Alani Mohammed M. (2018). Detecting NDP Distributed Denial of Service Attacks Using Machine Learning Algorithm Based on Flow-Based Representation. 2018 11th International Conference on Developments in eSystems Engineering (DeSE), 134-140, IEEE. doi:10.1109/INES.2018.8523851.

[14] Zhang Boyang; Zhang Tao; Yu Zhijian. (2017). DDoS detection and prevention based on artificial intelligence techniques. 2017 3rd IEEE International Conference on Computer and Communications (ICCC), 1276-1280, IEEE. doi:10.1109/CompComm.2017.8322748.

[15] Hoon Kian Son; Yeo Kheng Cher; Azam Sami; Shunmugam Bharanidharan; De Boer

Friso. (2018). Critical review of machine learning approaches to apply big data analytics in DDoS forensics. 2018 International Conference on Computer Communication and Informatics (ICCCI), 1-5, IEEE. doi:10.1109/ICCCI.2018.8441286.