An Overview of DoS and DDoS Attack Detection Techniques

Mateusz Gniewkowski(✉)

Faculty of Electronics, Wrocław University of Science and Technology, Wrocław, Poland
mateusz.gniewkowski@pwr.edu.pl

Abstract. The economic impact of (distributed) denial-of-service attacks is substantial, especially at a time when we rely on web applications more and more often. That is why, it is essential to be able to detect such threats early and therefore react before significant financial losses. In this paper, we focus on techniques, for detecting this type of attacks, that use historical data. We will discuss existing datasets, extracted features and finally the methods themselves. The solutions mentioned in this work are based on supervised learning (k-NN, MLP, DNN), unsupervised learning (mostly modified K-Means) and anomaly detection in time series analysis (ARIMA models family).

Keywords: DoS · DDoS · Anomaly detection · ARIMA · DNN · K-means · Datasets

1 Introduction

The purpose of denial-of-service (*DoS*) attacks is to prevent or disturb users from using internet applications through intentional exhaustion of a given resource (e.g. available sockets, bandwidth or computing power). The distributed version of this type of attack (*DDoS*) is different in that many computers are used to send packets. It is more difficult to perform, but it allows to consume the resource faster and makes it more difficult to react to the threat (it is necessary to filter out multiple connections which often look like normal network traffic).

A typical DDoS attack involves the creation of a so-called botnet - a set of computers on which the attackers took control. Joining new computers to such a network is often based on distributed scanning for hosts with known vulnerabilities and exploiting them, but users usually (unintentionally) install malicious software on their computers themselves (malicious email attachments, suspicious programs from the Internet). Many servers on the network are also constantly subjected to dictionary attacks [25], which also can provide access to the attacker. An interesting phenomenon is a situation in which Internet services cease to function due to naturally increasing interest (e.g. related to sales). Symptoms of such a situation may not always be distinguishable from a real attack.

W. Zamojski et al. (Eds.): DepCoS-RELCOMEX 2020, AISC 1173, pp. 233–241, 2020.
https://doi.org/10.1007/978-3-030-48256-5_23

To better understand the problem that will be addressed in this article, it is good to recall the classification of DoS attacks [17]:

1. Network Device Level – any remote attacks that involve preventing the proper functioning of network devices such as routers or switches,
2. OS level – attacks related to implementation errors of the given protocol in the operating system,
3. Application level – errors in user applications such as a poorly defined interface or buffer overflow errors,
4. Data flood – flooding a single device with a huge amount of data,
5. Protocol feature attack – all the attacks that exploiting protocol features. A good example is the *SYN flooding attack*.

The above classification gives an overview of what type of data can be used in the process of detecting DoS or DDoS attacks. This is primarily intercepted network traffic, but logs and monitoring information can also be helpful, especially when dealing with a new type of attack.

Several works [15,22] proposed a classification of DDoS attacks from both sides: the attacker and the defender perspective. Figure 1 shows the taxonomy of defence mechanisms against DDoS attacks. In this work, we will focus on "Classification by attack detection strategy", in particular on the "NBS-2" group. This group concerns methods that use historical data to detect anomalies: events that are significantly different from the others (in our case an anomaly can be understood as an attack). The main advantage of such a solution is that it allows for a certain generalisation (and therefore, detection of unknown attacks). On the other hand, it has a tendency to misidentify normal user behaviour as an attack.

In the following section, we will shortly describe few datasets that are most commonly used in denial-of-service attack detection techniques. In Sect. 3, we will describe several approaches to the problem. The last section covers conclusions.

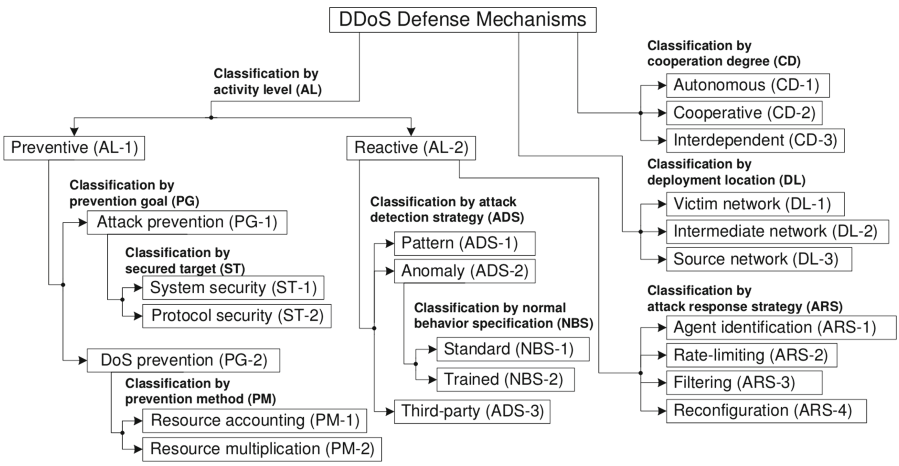

Fig. 1. Taxonomy of DDoS defense mechanisms [22]

2 Datasets

In the process of preparing a decision model, historical data is necessary to allow it to be trained. The problem with DoS-related datasets is that they are most often generated artificially (at least if the dataset is labelled). This requires the creation of statically correct methods of generating background network traffic (not-attacks), which, unfortunately, is a difficult task. In this section, we shortly describe a few most commonly used datasets. We try to focus on their possible criticism. If nothing is said about it, then the dataset is most likely reliable.

2.1 DARPA1998 and DARPA1999

DARPA1998 [1,19] is artificially generated and labelled dataset that contains nine weeks of network sniffing data, audit data (BSM) and full disk dumps from the three UNIX victim machines. It includes four types of attacks: DoS, R2L, U2R and Probing. One year later, a new dataset occurred called DARPA1999 [2,18] - the major differences are the addition of a Windows NT workstation as a victim and expanding the attack list. The complete lists of attacks in DARPA datasets are given in Table 1.

Table 1. DoS attack types used in the DARPA datasets

	Solaris	SunOS	NT	Linux
DARPA1998	apache2	apache2		apache2
	back	back		back
	mailbomb	land		mailbomb
	neptune	mailbomb		neptune
	ping of death	neptune		ping of death
	process table	ping of death		process table
	smurf	process table		smurf
	syslogd	smurf		teardrop
	udp-storm	udp-storm		udp-storm
DARPA1999				apache2
				arppoison
	neptune			back
	pod	arpoison	arppoison	mailbomb
	processtable	land	crashiis	neptune
	selfping	mailbomb	dosnuke	pod
	smurf	neptune	smurf	processtable
	syslogd	pod	tcpreset	smurf
	tcpreset	processtable		tcpreset
	warezclient			eardrop
				udpstorm

Although they are very popular datasets (they have been around for a long time and many researchers, wanting to compare their results with others, use them), they should not be used today. One of the reasons is that they are outdated and therefore less well suited to modern-day attacks and network traffic in general. The main cause is the wide criticism described in [20,21]. In 2000 another dataset appeared [3] in which several DDoS attacks were carried out using specific scenarios, but the background traffic is essentially the same as in 1999.

2.2 KDD1999

KDD [8] is another dataset that appears in many works related to denial-of-service attacks. It is actually a transformed DARPA1998 and should not be used for research due to the mentioned criticism. More about issues with this dataset can be found in [10,29].

2.3 CAIDA2007

The CAIDA2007 [4] is one of many datasets provided by CAIDA organisation. It contains one hour of a sequence of anonymized traffic traces (pcap files) from a real DDoS attack to one victim. Sadly, this dataset is now available only from IMPACT (https://www.impactcybertrust.org/), which means that you can legally download it only from the USA, Australia, Canada, Israel, Japan, The Netherlands, Singapore and UK.

2.4 ISCXIDS2012 and CICIDS2017

The ISCXIDS2012 [7] and CICIDS2017 [5] are two of many labelled datasets provided by University of New Brunswick. The first of them (described in [28]) consist of 7 days of real-like traffic (authors analysed real traces to create agents that generate it). Several attack scenarios were prepared, two of which were related to DoS attacks: "HTTP denial of service" and "distributed denial of service using an IRCBotnet". The second dataset (described in [27]) was generated correspondingly. It consists of 5 days of network traffic and includes following attacks: Brute Force FTP, Brute Force SSH, DoS, Heartbleed, Web Attack, Infiltration, Botnet attack and DDoS.

The datasets provided by the University of New Brunswick seem to be one of the best available to the public. At the end of 2019 another one [6] appeared, which only concerns DDoS attacks. Not many works related to it has yet appeared.

3 Methods Used in DoS and DDoS Attack Detection

In this work, three classes of methods for detection of DoS and DDoS attacks have been distinguished:

1. based on anomaly detection in time series analysis,
2. based on semi-supervised learning or unsupervised learning,
3. based on supervised learning.

This classification is not perfectly separable but allows to better understand the different approaches used in the problem. All of the listed works use at least one of the datasets from the previous section.

3.1 Anomaly Detection in Time Series Analysis

ARIMA [11] is one of the most widely used models for time-series forecasting. Due to the fact that network traffic can be presented as a series of values over time, it is possible to use such a method on it [30]. An anomaly (and thus a potential attack) is a situation in which the predicted value significantly differs from the actual value. Now the question is: How can the input data for the ARIMA algorithm be obtained from the captured network traffic and how exactly can we detect the anomaly?

Probably one of the first works using the ARIMA algorithm is [32]. The authors predict the flow (calculated in MB) of packets in each second. If a certain threshold is exceeded, an alarm is raised. The authors test their solution only for TCP flooding and UDP flooding attacks on their own generated data (they attack their system themselves). It makes it difficult to compare their solution with others, which is a common problem in this type of papers. Relying solely on the flow is also able to detect only a narrow group attacks.

Another example of time series analysis in DDoS detection problem is shown in [13]. Apart from the fact that the AR algorithm was used for forecasting (ARIMA is an AR generalization), the classification method as DDoS traffic has been changed. The solution is based on Lyapunov exponent [31] and [14]. Lyapunov exponent can be defined as follows:

$$\lambda_k \approx \ln(\frac{\Delta x_k}{\Delta x_0}) * \frac{1}{t_k},$$

where Δx_k is the difference between the real and predicted value and t_k is a time range. The researchers state that if $\lambda_k < 0$ the traffic might be a DDoS attack. To evaluate the results, the authors selected three days from the DARPA2000 dataset. The usage of the above equation gave them 71.84% of true positives (positive means anomaly). To improve the results, they trained a back-propagation neural network, which made it possible to achieve the result of 93.75%. The authors, unfortunately, do not give the number of false alarms. Another problem is using a dataset that doesn't have a good reputation.

The latest work [23] based on a similar idea is a work in which the authors managed to achieve results at the level of 98% (sensitivity, but the entire matrix

of confusion is also given in this paper). The tests were performed on the fifth Friday of DARPA1998 dataset. The authors argue that this dataset was chosen because the result could be compared with others. The algorithm analysed two time series (the number of packets in one minute and the number of packets in one minute divided by the number of IP source addresses).

The works based on time series analysis are not very precise, because they usually lack accurate evaluation tests (no datasets other than DARPA and selection of only a part of the dataset may be biased). Methods based mainly on deep learning are becoming more and more popular, but it could be worth to ensure whether, for certain specific DoS or DDoS problem, time series analysis algorithms do not perform better.

3.2 Semi-supervised Learning and Unsupervised Learning

Many of DDoS attack detection methods are based on unsupervised learning. Data is usually divided into two clusters where one is designated for regular network traffic and the other for anomalies. Most of the available solutions are based on the classic K-means algorithm. For example, in [26] authors modified the algorithm so that it also iteratively adds and removes additional clusters. This method should allow it to better handle non-spherical distributions. The features were extracted from network traffic using a sliding window algorithm with constant size. Authors use nine of them, among others: number of traffic from the same source IP, number of traffic with "SYN" flag, number of traffic with the same protocol etc. The choice was not justified. For the purpose of evaluation, the DARPA1998 dataset was used. The method obtained 99% of precision and 1.2% of FPR. Apart from the quality of the dataset and the unclear method of testing, the result is quite high and therefore it should be verified. It is worth to notice that this work (and most of the following) does not focus only on DDoS attacks but also classifies all the others available in the dataset.

Another interesting example of work using the K-means algorithm and a hybrid of SVM and ELM algorithms is [9]. Authors separate the training dataset into five categories related to attack types in the dataset (Normal, DoS, Probe, R2L, and U2R). After that, a slightly modified version of K-means algorithm is used in every category to obtain new training datasets. Then the SVM or ELM algorithm is trained on each of the newly received datasets. The prediction process is carried out as shown in Fig. 2. The work, unfortunately, uses the discredited dataset KDD99. Let us remind that this dataset contains already extracted features. The overall performance of the proposed algorithm achieved 95.2% of precision and 1.9% of FPR, also 99.6% of DoS attacks were recognised correctly.

The newer idea based on semi-supervised learning is presented in [16]. The authors attached great importance to the selection of initial features and, based on a review of related works, they selected nine entropy-based features. The proposed algorithm evaluates and selects a subset of those features for a given dataset and performs another version of modified K-Means algorithm. In this work, the initial positions of centroids depend on a labelled sample of data. For

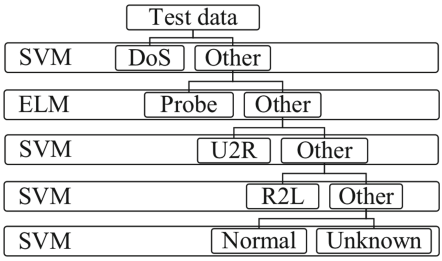

Fig. 2. Multi-level hybrid SVM and ELM [9]

evaluation purposes, they used four different datasets: DARPA2000 (for comparative purposes), CAIDA2007, CICIDS2017 and "Real-world dataset" (their own experiment). They achieved over 99% of precision for each of them.

3.3 Supervised Learning

One of the simplest examples of a supervised learning algorithm is k-NN. It was used among the others in the article [24] in order to classify the network status (normal, pre-attack, attack) rather than the traffic itself. As a distance measure, a weighted cosine formula was applied. The authors used DARPA2000 dataset and mostly entropy-based set of features. In this problem, they managed to achieve 92% of accuracy, but the results are hard to compare with others.

In [10] authors applied decision tree algorithm to classify attacks in KDD99 and they managed to correctly specify 97.1% of DoS attacks. An important contribution that this work brought is that KDD99 is not an appropriate transformation of DARPA1998 dataset, making R2L attacks difficult to classify. The authors introduced a few conditions, that might prevent information losses.

Newer methods often benefit from deep learning and do not bother with elaborated feature extraction. In the example from [12], authors trained two channels CNN network (packet and traffic features) and achieved 98.87% of accuracy for CICIDS2017 dataset and 98.54% for KDD99 dataset. Authors in [33] prepared few variants of LSTM neural network to predict the label for the last packet in a window. They evaluate their work on two days from ISCX2012 dataset and achieved 97.996% and 98.410% of accuracy respectively.

4 Conclusion

In this paper, we shortly discussed several datasets and methods used in DDoS detection problem. Four conclusions are drawn from the overview. First of all, there is a problem with comparing results. Most researchers must refer to the outdated and discredited DARPA dataset. What is more, many of the methods have never been verified on newer datasets. This mainly applies to those based on time series analysis. An interesting and understandable phenomenon is testing

solutions on one's servers. However, the results of such an experiment are difficult to evaluate. Maybe researchers should standardise the method of conducting such experiments? This is not an easy task, but it is not impossible. Finally, not many researchers are concerned about the time complexity of their solutions, which may be important, especially for larger networks.

References

1. The 1998 DARPA intrusion detection evaluation dataset. https://www.ll.mit.edu/r-d/datasets/1998-darpa-intrusion-detection-evaluation-dataset. Accessed 05 Dec 2019
2. The 1998 DARPA intrusion detection evaluation dataset. https://www.ll.mit.edu/r-d/datasets/1999-darpa-intrusion-detection-evaluation-dataset. Accessed 05 Dec 2019
3. 2000 DARPA intrusion detection scenario specific datasets. https://www.ll.mit.edu/r-d/datasets/2000-darpa-intrusion-detection-scenario-specific-datasets. Accessed 05 Dec 2019
4. The CAIDA UCSD DDoS attack 2007 dataset. http://www.caida.org/data/passive/ddos-20070804_dataset.xml. Accessed 05 Dec 2019
5. The CICIDS DDoS attack 2017 dataset. https://www.unb.ca/cic/datasets/ids-2017.htm. Accessed 05 Dec 2019
6. DDoS evaluation dataset (CICDDoS 2019). https://www.unb.ca/cic/datasets/ddos-2019.html. Accessed 05 Dec 2019
7. Intrusion detection evaluation dataset (ISCXIDS 2012). https://www.unb.ca/cic/datasets/ids.html. Accessed 05 Dec 2019
8. KDD CUP 1999 data. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html. Accessed 05 Dec 2019
9. Al-Yaseen, W.L., Othman, Z.A., Nazri, M.Z.A.: Multi-level hybrid support vector machine and extreme learning machine based on modified k-means for intrusion detection system. Expert Syst. Appl. **67**, 296–303 (2017)
10. Bouzida, Y., Cuppens, F.: Detecting known and novel network intrusions. In: IFIP International Information Security Conference, pp. 258–270. Springer (2006)
11. Brockwell, P.J., Davis, R.A.: Introduction to Time Series and Forecasting. Springer, Cham (2016)
12. Chen, J., Yang, Y.T., Hu, K.K., Zheng, H.B., Wang, Z.: DAD-MCNN: DDoS attack detection via multi-channel CNN. In: Proceedings of the 2019 11th International Conference on Machine Learning and Computing, pp. 484–488. ACM (2019)
13. Chen, Y., Ma, X., Wu, X.: DDoS detection algorithm based on preprocessing network traffic predicted method and chaos theory. IEEE Commun. Lett. **17**(5), 1052–1054 (2013)
14. Chonka, A., Singh, J., Zhou, W.: Chaos theory based detection against network mimicking DDoS attacks. IEEE Commun. Lett. **13**(9), 717–719 (2009)
15. Douligeris, C., Mitrokotsa, A.: DDoS attacks and defense mechanisms: classification and state-of-the-art. Comput. Netw. **44**(5), 643–666 (2004)
16. Gu, Y., Li, K., Guo, Z., Wang, Y.: Semi-supervised k-means DDoS detection method using hybrid feature selection algorithm. IEEE Access **7**, 64351–64365 (2019)
17. Karig, D., Lee, R.: Remote denial of service attacks and countermeasures. Princeton University Department of Electrical Engineering Technical report CE-L2001-002 **17** (2001)

18. Lippmann, R., Haines, J.W., Fried, D.J., Korba, J., Das, K.: Analysis and results of the 1999 DARPA off-line intrusion detection evaluation. In: International Workshop on Recent Advances in Intrusion Detection, pp. 162–182. Springer (2000)
19. Lippmann, R.P., Fried, D.J., Graf, I., Haines, J.W., Kendall, K.R., McClung, D., Weber, D., Webster, S.E., Wyschogrod, D., Cunningham, R.K., et al.: Evaluating intrusion detection systems: the 1998 DARPA off-line intrusion detection evaluation. In: Proceedings DARPA Information Survivability Conference and Exposition. DISCEX 2000. vol. 2, pp. 12–26. IEEE (2000)
20. Mahoney, M.V., Chan, P.K.: An analysis of the 1999 DARPA/Lincoln laboratory evaluation data for network anomaly detection. In: International Workshop on Recent Advances in Intrusion Detection, pp. 220–237. Springer (2003)
21. McHugh, J.: Testing intrusion detection systems: a critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by lincoln laboratory. ACM Trans. Inf. Syst. Secur. (TISSEC) **3**(4), 262–294 (2000)
22. Mirkovic, J., Reiher, P.: A taxonomy of DDoS attack and DDoS defense mechanisms. ACM SIGCOMM Comput. Commun. Rev. **34**(2), 39–53 (2004)
23. Nezhad, S.M.T., Nazari, M., Gharavol, E.A.: A novel DoS and DDoS attacks detection algorithm using arima time series model and chaotic system in computer networks. IEEE Commun. Lett. **20**(4), 700–703 (2016)
24. Nguyen, H.V., Choi, Y.: Proactive detection of DDoS attacks utilizing k-NN classifier in an anti-DDoS framework. Int. J. Electr. Comput. Syst. Eng. **4**(4), 247–252 (2010)
25. Pinkas, B., Sander, T.: Securing passwords against dictionary attacks. In: Proceedings of the 9th ACM Conference on Computer and Communications Security, pp. 161–170 (2002)
26. Pramana, M.I.W., Purwanto, Y., Suratman, F.Y.: DDoS detection using modified k-means clustering with chain initialization over landmark window. In: 2015 International Conference on Control, Electronics, Renewable Energy and Communications (ICCEREC), pp. 7–11. IEEE (2015)
27. Sharafaldin, I., Lashkari, A.H., Ghorbani, A.A.: Toward generating a new intrusion detection dataset and intrusion traffic characterization. In: ICISSP, pp. 108–116 (2018)
28. Shiravi, A., Shiravi, H., Tavallaee, M., Ghorbani, A.A.: Toward developing a systematic approach to generate benchmark datasets for intrusion detection. Comput. Secur. **31**(3), 357–374 (2012)
29. Tavallaee, M., Bagheri, E., Lu, W., Ghorbani, A.A.: A detailed analysis of the KDD CUP 99 data set. In: 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, pp. 1–6. IEEE (2009)
30. Vafeiadis, T., Papanikolaou, A., Ilioudis, C., Charchalakis, S.: Real-time network data analysis using time series models. Simul. Model. Pract. Theory **29**, 173–180 (2012)
31. Wolf, A., Swift, J.B., Swinney, H.L., Vastano, J.A.: Determining lyapunov exponents from a time series. Physica D **16**(3), 285–317 (1985)
32. Yaacob, A.H., Tan, I.K., Chien, S.F., Tan, H.K.: Arima based network anomaly detection. In: 2010 Second International Conference on Communication Software and Networks, pp. 205–209. IEEE (2010)
33. Yuan, X., Li, C., Li, X.: DeepDefense: identifying DDoS attack via deep learning. In: 2017 IEEE International Conference on Smart Computing (SMARTCOMP), pp. 1–8. IEEE (2017)