The University of Melbourne

COMP90007 Internet Technologies SM2, 2018

Research Project

Peiyong Wang

username peiyongw

When DDoS Attack Encounters Deep Learning
A More Sharp Spear or A Stronger Shield?

PEIYONG WANG
The University of Melbourne
October 12, 2018

1 Introduction

1.1 DDoS Attacks

Nowadays, DDoS (Distributed Denial of Service) attacks have been becoming a more and more immense threat toward the network infrastructures. A large scale DDoS attack on a website which is required to provide a low latency service(such as the website and network infrastructure of a stock trade platform) can make the entire service unavailable for hours, even days, causing great loss to the traders.

Research has been done on how the DDoS attacks work and how to detect them. DDoS attacks can roughly classified by degree of automation, exploited weakness to deny service, the validity of source address, rate dynamics of the attack, possibilities of classification, agent set persistency, type and impact of the victim. Based on different types of DDoS attacks, different means of DDoS attack defense are developed (Mirkovic & Reiher, 2004).

1.2 Artificial Intelligence and Deep Learning

Since the dawn of modern computer science, artificial intelligence and the research on learning as well as thinking machines has been an important part.

The first famous and successful application of deep learning was AlexNet in the ImageNet LSVRC-2010 contest (Krizhevsky, Sutskever, & Hinton, 2012) for image classification. Since then, the number of layers in neural networks are becoming larger and larger and the dataset on which the neural networks operates are also getting bigger and bigger(Simonyan & Zisserman, 2014), and residual blocks were introduced to neural network structures to avoid gradient vanishing or exploding(He, Zhang, Ren, & Sun, 2016). Besides image processing, deep learning is also widely applied to natural language processing, speech recognition and even solving differential equations.

2 Statistical and Deep Learning Methods for Detecting and Defending DDoS Attacks

Although there are several ways to shield network infrastructures from DDoS attacks such as shut down unnecessary services or limit SYN/ICMP traffic(Malik & Singh, 2015), the most important issue is to identify which requests of service are part of a DDoS attack and which are not(Yu, 2013).

2.1 Statistical Learning Methods for DDoS Detection and Defense: Researches and Drawbacks

Machine learning techniques, especially statistical learning methods have been applied to DDoS detection for some time. As a common classifier, naïve Bayes has been used in many classification tasks such as unbalanced text classification(Frank & Bouckaert, 2006). In recent literature, naïve Bayes has also been applied on the detection of DDoS attacks. Irfan Sofi et al.incorporated naïve Bayes with other statistical learning method such as support vector machines and decision trees to identify DDoS attacks including modern types such as HTTP flooding (Sofi, Mahajan, & Mansotra, 2017), achieving an overall accuracy of 96.89% for naïve Bayes.

Although achieved high accuracy on the dataset they collected from NIDS, during the training process, the researchers need to hand-pick the features from the dataset, which is very time consuming. This feature picking step is often referred as "feature

engineering" in statistical machine learning, which can be really hard to perform on GB or TB level data sets. Besides, traditional classification algorithms will be hard converging on large datasets and requires a lot of manual tuning. These methods and algorithms are also proved to be unable to learn deeper level of features from large amount of data (Goodfellow, Bengio, & Courville, 2016). Also, in real world cases, we often wouldn't know what the next DDoS attack will look like, so the performance of classifiers from trained on balanced data may be limited.

Network structured classification algorithms are also used in detection of DDoS attacks. Ugo Fiore et al. applied discriminative restricted Boltzmann machines, which is a semi-supervised learning algorithm to detect packets from a DDoS attack (Fiore, Palmieri, Castiglione, & De Santis, 2013). However, although shallow RBMs like the one used in this paper are easier to train than deeper ones, their ability to extract deeper level of representation of data is greatly limited (Bishop, 2006).

2.2 Deep Learning Methods for DDoS Attacks Detection and Defense

Deep learning methods, compared to traditional statistical learning methods, performs better on large datasets. Because usually the neural networks used in deep learning have more hidden layers, which can extract more latent features from the data set used both in training and testing compared to traditional statistical learning methods.

The fundamental mechanism of identifying packets from a DDoS attack, is that the normal packets and the abnormal packets will falls in different possibility distributions (sometimes after a nonlinear transformation) (Bishop, 2006). For this kind of task, normally a clustering algorithm or isolation forest will do a good job. However, learning algorithms based on statistics such as these two were found hard to converge on large datasets (datasets with tens of thousands of entries or more), which is very common for the DDoS attack datasets. Besides, such algorithms, when being trained on large datasets, are found time and resource consuming and not very easy to be accelerated by general purpose graphics processing units(GPGPU) due to the precision demands in each step of computation.

However, deep learning algorithms, especially deep neural networks outperform

traditional methods on large datasets, can be trained on multiple GPGPUs and easily deployed for distributed computing tasks. Li et al. applied a bidirectional recurrent neural networks to DDoS attack detection in an OpenFlow based software defined network (Li et al., 2018). Bi-RNNs were firstly put forward by Mike Schuster and Kuldip K. Paliwal in 1997 (Schuster & Paliwal, 1997) and used in predicting temporal dynamic systems. Later this was applied to machine translation and used as generative models to fill in the gaps in time series data (Berglund et al., 2015).

In Li et al's work, they firstly classified the twenty feature segments into three types of fields. After some preprocess for this three kinds of fields, the raw data was transformed into a 2-D matrix, which was cut into several continuous time frames and labeled with tag values indicating whether the tagged time frame belongs to a normal packet or a DDoS attack packet. After the preprocessing, they fed the data into a deep neural network mainly consists of a bi-RNN and several fully connected layers (Figure 2.1)(Li et al., 2018).

After the model was trained, the integrated the neural network to a OpenFlow based SDN and tested it with real-time DDoS attacks. In the training process the achieved an accuracy of 98% and in the test phase their model can effectively clean the network traffic from DDoS attacks.

Also. Yuan et al. applied RNNs to DDoS attack detection and reduced the error from 7.517% to 2.103% compared to conventional methods (Yuan, Li, & Li, 2017).

However, the data for DDoS attacks may not be as balanced as the data people used for lab researches. The number of "bad" packets is probably a lot smaller than the normal packets, maybe there is no DDoS attack packets for model training at all. In such cases, anomaly detection algorithms usually perform better than common classification algorithms. Sadly, common anomaly detection algorithms based on statistical learning methods do not easily converge on large datasets. Luckily, Chalapathy et al came up with a one-class neural network (Chalapathy, Menon, & Chawla, 2018), which combined deep neural networks for latent representation extraction from the data and a SVM like loss function. Their model performs well on sequential data. Hopefully this will be apply to DDoS detection in future researches.

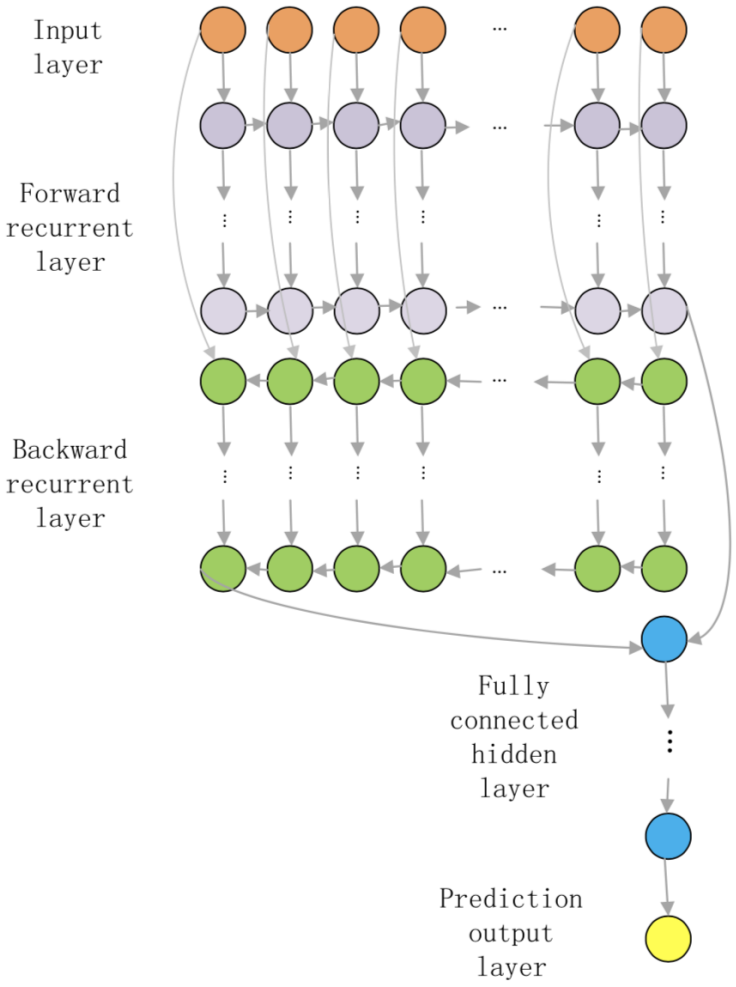

Figure 2.1: NN Architecture from Li et al's work

3 Another Side of the Coin: Threats from GAN and DRL

While we are enjoying the convenience brought to DDoS defense by deep learning, we cannot neglect the potential threats posed by it, especially GAN (Generative Adversarial Nets) and DRL (Deep Reinforcement Learning).

GAN, which was proposed by Goodfellow et al, as well as its variations, is often used to generate data which follows a possibility distribution that extremely resembles the distribution followed by training data(Goodfellow et al., 2014). This technology has already been used to generate fake pictures which look a lot like real ones. Although can be used to generate DDoS attack data in the training process for a DDoS detection model when lacking abnormal packets, GAN can also be used by perpetrators to generate DDoS attack traffic which shares great similarities with legitimate network traffic, making it much more harder to detect and defense such attacks.

DRL, which was proposed by Google DeepMind based on traditional Q-learning algorithms (Mnih et al., 2015) for decision making and gaming, demonstrated its power when a go program built on it defeated the best human go player Lee Sedol (Silver et al., 2016) and later defeated all human on nearly all board games (Silver et al., 2017) without prior knowledge from human being. Based on this, we can assume that people can train a deep reinforcement neural network, which uses a deep neural network as the action-value function, and interact with an environment consisting of nowadays DDoS detection and defense algorithms. By using how many attack packets bypass the defense system as the reward function in this DRL system, given enough training time, we may get a DDoS attack system that can bypass nearly all the defense systems, which will be a great threat to our network infrastructures.

4 Conclusion and Future Work

Compared to conventional machine learning algorithms, deep learning shows great potential in the detection and defense against modern DDoS attacks. However, the potential threats brought by deep learning can also not be neglected.

Future work should concentrate on two aspects: one is to apply deep learning anomaly

detection method to network abnormalities detection, and the other is to investigate how GAN and DRL will shape the form of future DDoS attacks.

References

Berglund, M., Raiko, T., Honkala, M., Kärkkäinen, L., Vetek, A., & Karhunen, J. T. (2015). Bidirectional recurrent neural networks as generative models. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, & R. Garnett (Eds.), *Advances in neural information processing systems 28* (pp. 856–864). Curran Associates, Inc.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Chalapathy, R., Menon, A. K., & Chawla, S. (2018). Anomaly Detection using One-Class Neural Networks. *arXiv:1802.06360 [cs, stat]*. (arXiv: 1802.06360)

Fiore, U., Palmieri, F., Castiglione, A., & De Santis, A. (2013). Network anomaly detection with the restricted Boltzmann machine. *Neurocomputing*, *122*, 13–23. doi: 10.1016/j.neucom.2012.11.050

Frank, E., & Bouckaert, R. R. (2006). Naive bayes for text classification with unbalanced classes. In *Proceedings of the 10th european conference on principle and practice of knowledge discovery in databases* (pp. 503–510). Berlin, Heidelberg: Springer-Verlag. doi: 10.1007/11871637_49

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., … Bengio, Y. (2014). Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems 27* (pp. 2672–2680). Curran Associates, Inc.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *CVPR* (pp. 770–778). IEEE Computer Society.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012, June). Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems 25* (pp. 1097–1105). Curran Associates, Inc.

Li, C., Wu, Y., Yuan, X., Sun, Z., Wang, W., Li, X., & Gong, L. (2018). Detection and defense of DDoS attack-based on deep learning in OpenFlow-based SDN. *International Journal of Communication Systems*, *31*(5), e3497. doi: 10.1002/dac.3497

Malik, M., & Singh, D. Y. (2015, June). A Review: DoS and DDoS Attacks. *IJCSMC*, *4*,

260 – 265.

Mirkovic, J., & Reiher, P. (2004, April). A taxonomy of DDoS attack and DDoS defense mechanisms. *ACM SIGCOMM Computer Communication Review*, *34*(2), 39. doi: 10.1145/997150.997156

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... Hassabis, D. (2015, February). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533.

Schuster, M., & Paliwal, K. (1997, November). Bidirectional recurrent neural networks. *Trans. Sig. Proc.*, *45*(11), 2673–2681. doi: 10.1109/78.650093

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., ... Hassabis, D. (2016, January). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484–489. doi: 10.1038/nature16961

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., ... Hassabis, D. (2017, October). Mastering the game of go without human knowledge. *Nature*, *550*, 354.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *CoRR*, *abs/1409.1556*.

Sofi, I., Mahajan, A., & Mansotra, V. (2017, June). Machine Learning Techniques used for the Detection and Analysis of Modern Types of DDoS Attacks. *International Research Journal of Engineering and Technology*, *04*(06), 8.

Yu, S. (2013). *Distributed denial of service attack and defense*. Springer Publishing Company, Incorporated.

Yuan, X., Li, C., & Li, X. (2017, May). DeepDefense: Identifying DDoS Attack via Deep Learning. In *2017 IEEE International Conference on Smart Computing (SMARTCOMP)* (pp. 1–8). Hong Kong, China: IEEE. doi: 10.1109/SMARTCOMP.2017.7946998