Source-side DoS Attack Detection with LSTM and Seasonality Embedding

Sungwoong Yeom
Dept. of Artificial Intelligence Convergence, Chonnam National University
Gwangju City, South Korea
yeomsw0421@gmail.com

Chulwoong Choi
Dept. of Artificial Intelligence Convergence, Chonnam National University
Gwangju City, South Korea
sentilemon02@gmail.com

Kyungbaek Kim
Dept. of Artificial Intelligence Convergence, Chonnam National University
Gwangju City, South Korea
kyungbaekkim@jnu.ac.kr

Abstract

As the denial of service attacks become sophisticated, source-side detection methods are being studied to address the limitation of target-side detection methods such as delayed detection and difficulty in tracking an attacker. Recently, some source-side detection methods are studied to use the adaptive attack detection threshold by considering seasonal behavior of network traffic. However, recent network traffic usage patterns have become irregular, and the performance of the adaptive threshold technique has deteriorated due to the increase in randomness and burstiness of the traffic.

In this paper, an LSTM(Long Short Term Memory) based source-side DoS attack detection technique is proposed in order to keep high performance under irregular seasonal traffic usage. The proposed LSTM based detection model was designed with the input feature vector consisting of an index of unit time, a normal traffic volume, and a traffic trend. Specifically, in order to make LSTM learn irregular seasonal pattern effectively, several embedding methods were proposed to embed the irregular seasonal pattern as a traffic trend of the input vector. Through extensive experiments with actual network traffic, it is observed that the proposed LSTM-based technique achieved high attack detection rate of 92% and a low false positive rate of 20% under a network with irregular and burst traffic.

CCS Concepts: • **Networks** → **Denial-of-service attacks**; • **Computing methodologies** → *Neural networks.*

Keywords: Network Security, DoS Attack, SDN IDPS, LSTM, Source-side Attack

SAC '21, March 22–26, 2021, Virtual Event, Republic of Korea

ACM ISBN 978-1-4503-8104-8/21/03…$15.00
https://doi.org/10.1145/3412841.3441987

1 INTRODUCTION

Depending on the activation of devices such as sensors, mobiles, wearables and other IoT (Internet of Things), the amount of IoT data moving over the network is exploding. These IoT devices can be abused by causing the DoS (denial of service) attack [1, 9]. The victim detection method has disadvantages such as detection delay and difficulty in tracking the attacker. Edge computing has the potential to solve these problems by moving computing locations closer to the network edge and data sources. After deploying a detection system in the source-side network, it is possible to defend against DoS attacks that are difficult to defend on the victim-side.

However, unlike the victim-side network, the amount of attack traffic observed in the source-side network is too small. Therefore, in the source-side network, attack traffic can be mixed with normal traffic easily. In order to detect the detailed attack traffic, dynamic adaptive threshold methods based on observed traffic volumes are studied [17]. However, if observed traffic is mixed with attack traffic, these methods should separate the normal traffic from the attack traffic to calculate the new threshold value. To separate attack traffic from the observed traffic, estimating the normal traffic volume by utilizing network traffic seasonality was studied [16]. This study identified the seasonal behavior through network traffic volume statistics to estimate normal traffic.

Recently, with the activation of time series deep learning neural networks, the short-term network traffic volume prediction method based on time-series neural network shows low error rate [15, 18, 20]. However, because of unexpected behaviors of network users, network traffic may have high jitter with frequent explosive traffic and it shows non-linear properties. [4–6]. In a non-linear network traffic [14], burstiness [10] and randomness are easily observed, and it degrades the performance of traffic volume prediction method as well as the DoS attack detection method. To mitigate this degradation, we need to consider the new way of training the neural network model for predicting traffic volume by using the relationship between traffic traffic states such as distinguishing different network traffic in embedding spaces. [3, 7]

In this paper, we propose LSTM based source-side DoS attack detection method. If the observed traffic in any unit time is mixed with attack traffic, the proposed method uses LSTM based traffic prediction model to estimate the normal traffic volume in the related unit time. In the proposed LSTM based traffic prediction model, three features are used as an input vector. The first feature is the changing rate of the observed traffic that represents the changing rate in the observed traffic volume in the current unit time compared to the observed traffic volume in the previous unit time. The second feature is time window index that represents the index of the observed traffic in the corresponding unit time. The third feature is traffic trend that represents a trend in the mid-to long-term changing rate in the observed traffic at consecutive unit times. This traffic trend can be expressed as a static traffic trend state and a dynamic traffic trend state. The static traffic trend state method embeds the trend state to utilize seasonal traffic patterns according to the statistics of observed traffic. On the other hand, the dynamic traffic trend state method embeds the trend state according to the changing rate of traffic observed during a unit time.

In order to verify the effectiveness of the proposed method, we evaluated the detection rate, false positive rate, and balanced accuracy of the proposed method by conducting an experiment according to the average traffic volume, jitter, and burst ratio based on actual DNS (Domain Name System) traffic data. In addition, the performance of attack detection according to the types of embedding method for the LSTM input vector was evaluated.

The rest of this paper is arranged as follows. Section 2 describes the work associated with network traffic prediction for making the DoS attack detection system efficient. Section 3 describes the proposed source-side DoS attack detection system and LSTM based traffic volume prediction model. Section 4 describes the proposed seasonality embedding method. Section 5 presents the results of the experiment, and section 6 provides conclusions and future research directions.

2 BACKGROUND AND RELATED WORK

2.1 SDN BASED DOS ATTACK DEFENSE SYSTEM

SDN (Software-Defined Networking) [12] can provide a more dynamic, manageable, and adaptive network. SDN makes networks more flexible and efficient by handling high bandwidth and modern applications which have dynamic behaviors. In a networking paradigm, a network control logic can be programmable, enabling easy configuration, holistic management, and quick network resources optimization. In addition, administrators can dynamically adjust traffic flows across networks to meet new application demands. As SDN provides a new and dynamic network architecture, it becomes easier to detect and react to DoS attacks. A SDN based DoS defense scheme can be classified based on the location of deployment: victim-side defense mechanisms using SDN and source-side defense mechanisms using SDN.

In a victim-side defense mechanism, it detects, filters, and limits malicious traffic at the routers of victim-side networks. It is straightforward to detect DoS attacks on the victim-side, because the volume of malicious network traffic is extremely high and the attack traffic can be clearly distinguished from normal traffic. Zhang.et al. [23] suggested an ARIMA (Autoregressive Integrated Moving Average) model for protecting servers from DoS attacks. Wu et al. [22] traced back to the attacker location based on traffic flow pattern matching using decision trees. However, during DoS attacks, victim resources such as network bandwidth are often overwhelmed, and this approach cannot prevent the flow beyond the routers which are near to the victim. In addition, it is not useful to detect an attack only after reaching the victim and refraining a normal client to access the victim.

The source-side defense mechanisms are deployed near to the attack sources. In this approach, malicious packets are identified while they pass a gateway of a subnet where the attack sources reside in order to prevent generating attack traffic to a victim. Detecting DoS attacks at the source is the best possible way to detect an attacker. Monge et al. [13] suggested to detect discordant behavior by detecting the participation of end-users or IoT devices on the source-side. However, the traffic flow observed in the source-side network has characteristics such as time dependency, self-similarity, and seasonality. He et al. [8] suggested characterizing the seasonality pattern of Internet user traffic. If a dynamic adaptive threshold method is used based on the observed traffic volume by utilizing these characteristics, DoS attacks which require minute observation from the source-side can also be detected well. Nguyen et al. [16] suggest an effective source-side DoS detection method with traffic seasonality aware adaptive threshold. However, as the patterns of users who generate network traffic have become irregular in recent years, characteristics such as non-linearity, randomness, and burstiness have been observed in network traffic. If we learn various patterns of traffic through time series deep learning model, we can improve the performance of the source-side DoS attack detection techniques with irregular network traffic.

2.2 PREDICTION OF NETWORK TRAFFIC VOLUME

The one of key method of a source-side DoS attack detection mechanism is predicting the volume of normal network traffic. The amount of network traffic observed in the source-side network is relatively small, and more accurate prediction of network traffic is required in order to adjust threshold in a fine tuned manner.

In general, the statistical characteristics of traffic observed on the source-side network are time dependency, self-similarity, seasonality, non-linearity, randomness and burstiness. Wang

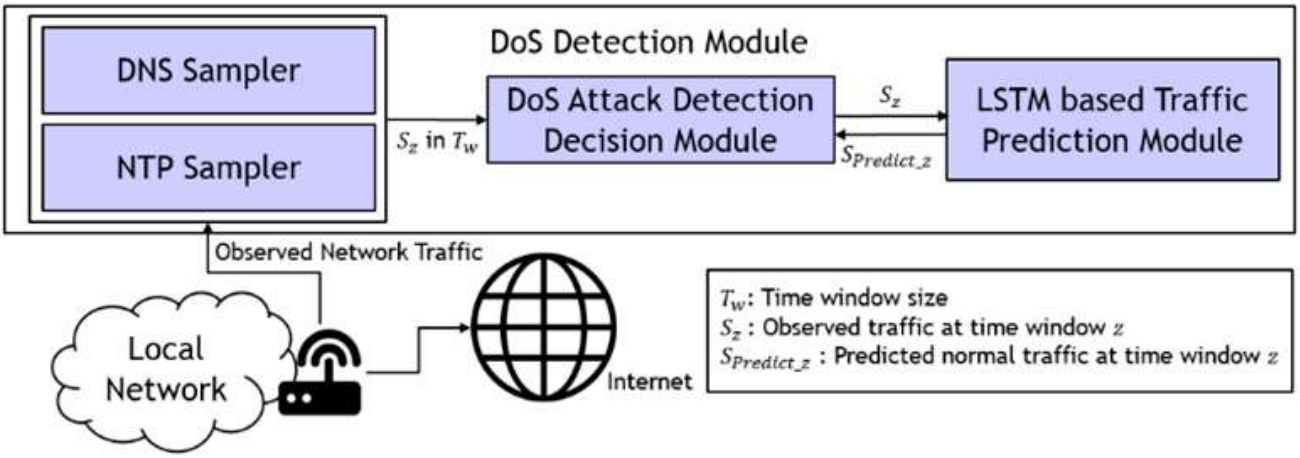

Figure 1. Overview of the proposed LSTM based source-side DoS detection system

et al. [21] suggested a flow modeling and prediction method based on ARMA(AutoRegressive Moving-Average) models according to the short-term correlation and self-similarity of traffic data to increase the accuracy of prediction and reduce overhead. However, this approach does not consider the irregular usage patterns of network users such as non-linearity, randomness, and burstiness.

Recently, there are few studies which uses LSTM to predict network traffic volume. Vinayakumar et al. [20] used GEANT backbone network data to learn and predict non-linear characteristics of traffic through LSTM neural networks. Azzouni et al. [2] presented LSTM based model for high volume traffic prediction by deploying the model in SDN to predict the volume of traffic between nodes and train the model using real data. Lu et al. [11] suggested a real-time network traffic prediction model based on LSTM to cope with network traffic burstiness and uncertainty. Nguyen et al. [15] proposed LSTM-based network traffic volume estimation method by learning the traffic seasonality observed at the source-side. In these studies, a network traffic trend is distinguished from other network traffic trend, and the training model of network traffic trend in LSTM can be efficiently customized by applying different embedding spaces [3, 7]. In this paper, we propose a LSTM-based source-side DoS detection method with various embedding spaces for representing irregular traffic seasonal patterns.

3 LSTM BASED SOURCE-SIDE DoS ATTACK DETECTION SYSTEM

3.1 SYSTEM DESIGN

Figure 1 shows an overview of the proposed LSTM-based source-side DoS detection system. The proposed system can be deployed on a gateway of a target subnet to detect DoS attack traffics in advance which flow to a victim. This DoS detection system captures network traffic from the gateway through samplers such as a DNS sampler and a NTP sampler by using SDN [17]. In every given time unit, called as a time window with a given constant size T_w, the network traffic is captured, and the volume of the observed traffic in the z_{th} time window is defined as S_z. Whenever the network traffic is captured, adaptive threshold is applied to S_z in order to determine whether the observed traffic contains malicious attack traffic or not. The adaptive threshold is dynamically adjusted by predicting the traffic volume of the next time window by using an exponential smoothing algorithm. In an exponential smoothing function, the predicted traffic volume of the next time window, $\bar{S}_{z+1}$, is needed. $\bar{S}_{z+1}$ is calculated by the equation 1.

$$\bar{S}_{z+1} = \alpha * S_z + (1 - \alpha) * \bar{S}_z \tag{1}$$

Then, the adaptive attack detection threshold, θ_{z+1}, is set dynamically by adding margin, δ, to the predicted traffic volume, $\bar{S}_{z+1}$, as shown in the equation 2.

$$\theta_{z+1} = (1 + \delta) * \bar{S}_{z+1} \tag{2}$$

If the traffic volume observed in the z_{th} time window, S_z, is bigger than the detection threshold of the z_{th} time window, θ_z, it is judged that observed traffic is mixed with attack traffic. In this case, the volume of attack traffic should separate from the volume of the observed traffic for the new adaptive detection threshold. To separate attack traffic from observed traffic, the predicted volume of normal traffic of the z_{th} time window, $S_{predict_z}$, is required. $S_{predict_z}$ is set by multiplying observed traffic volume in the previous $(z-1)_{th}$ time window, S_{z-1}, to the traffic changing rate in the z_{th} time window, Δ_z, as shown in the equation 3.

$$S_{predict_z} = S_{z-1} * \Delta_z \tag{3}$$

By using predicted volume of normal traffic, $S_{predict_z}$, the traffic volume for setting attack detection threshold about the

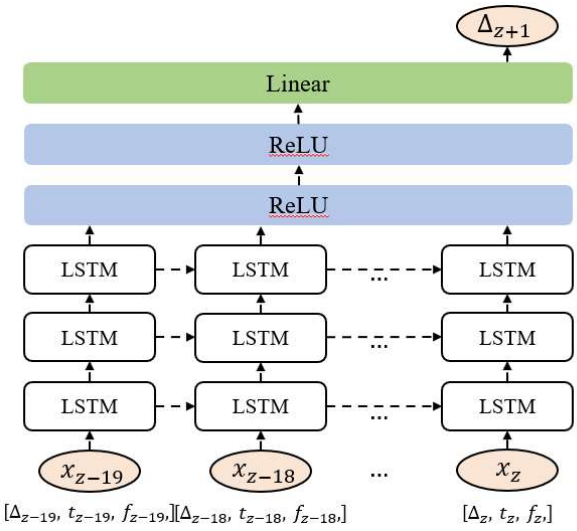

Figure 2. Architecture of LSTM based traffic volume prediction model

$(z+1)_{th}$ time window is calculated as shown in the equation 4.

$$\bar{S}_{z+1} = \alpha * S_{predict_z} + (1-\alpha) * \bar{S}_z \tag{4}$$

With this, the adaptive threshold for the $(z+1)_{th}$ time window is calculated with the equation 2.

3.2 LSTM BASED TRAFFIC VOLUME PREDICTION MODEL

The important knob of the performance of predicting the volume of network traffic is how to estimate Δ_Z in z_{th} time window accurately, and the proposed approach uses LSTM which is a time series neural network model for this estimation. Figure 2 depicts the architecture of the proposed LSTM based traffic volume prediction model. This model consists of three LSTM layers and three dense layers. Among these three dense layers, the first two layers use ReLU (Rectified Linear Unit) activation function and the last layer uses Linear activation function. ReLU activation function usually improve the performance of LSTM based prediction model [19]. Each layer includes 20 nodes.

For training and testing the proposed LSTM based traffic volume prediction model, it is necessary to manage observed traffic in a well-organized manner. The observed traffic during a day y is one set of traffic data represented as d_y. During one day, the network traffic is observed k times at a given constant time interval, and the volume of observed network traffic at k_{th} time interval is represented as S_k as shown in equation 5.

$$d_y = \{S_1, S_2, ..., S_k\} \tag{5}$$

With this collection of observed data, the input vector for the proposed LSTM model is generated in order to ensure the stability of the prediction. The input vector consists of three features including the changing rate of observed traffic volume Δ_z, the time window index t_z, and the traffic trend f_z which are corresponding to the z_{th} time window. The input vector for the z_{th} time window represents as I_z=(Δ, t_z, f_z).

The proposed LSTM model trains with the input vectors to predict the traffic changing rate. In detail, the proposed LSTM model uses input vector $I_{z-19}, I_{z-18}, ..., I_z$ as 20 continuous observed traffic information from $(z-19)_{th}$ to z_{th}, and it provides output as the traffic changing rate at the $(z+1)_{th}$ time window Δ_{z+1}.

The first domain of the input vector is the traffic changing rate Δ_z, which is the changing rate between two continuous observed traffic volumes. Because the observed network traffic is managed in a daily manner, the traffic changing rate is also managed in a daily manner. The set of traffic changing for a day y is represented as c_y like the following equation.

$$c_y = \{\Delta_1, \Delta_2, ..., \Delta_k\} \tag{6}$$

Accordingly, the traffic changing rate of the z_{th} time window of the y_{th} day can be calculated as the following equation.

$$c_y[\Delta_z] = d_y[S_z]/d_y[S_{z-n}] - 1 \tag{7}$$

However, in the case of Δ_1 which is corresponding to the first observed traffic of a day, the calculation of traffic changing rate should consider the change in traffic between the last observed traffic of the previous day and the first observed traffic of the current day as the following equation.

$$c_y[\Delta_1] = d_y[S_z]/d_{y-1}[S_k] - 1 \tag{8}$$

The second domain of the input vector is the time window index for the traffic changing rate. In this paper, the duration of time window T_w sets to 1 minutes, and the length of d_y and c_y becomes 1440. Accordingly, the time window index t_z has values between 1 and 1440.

The third domain of the input vector is a network traffic trend, which represents the characteristics of network traffic including seasonality and randomness. The traffic trend is categorized by considering various characteristics of network traffic such as changing rate and seasonality. The categorized traffic trend is embedded to f_z through one-hot encoding as following equation where i is the number of possible categories of the network traffic trend.

$$f_z = \{v_1, v_2, ..., v_i\}, v \in \{0, 1\} \tag{9}$$

4 TRAFFIC SEASONALITY EMBEDDING

The network traffic trend is necessary to express the relationship between the seasonal patterns of traffic observed per unit time in order to clearly improve the performance of

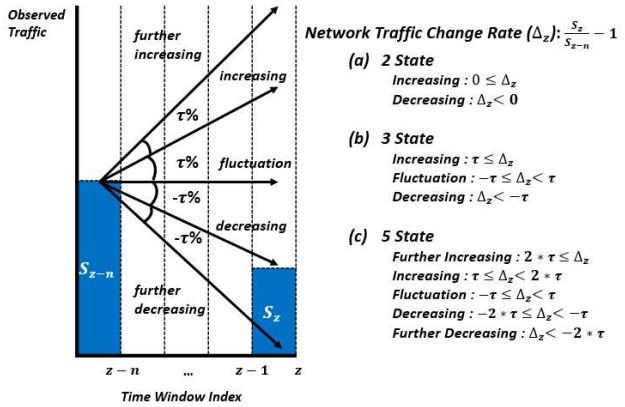

Figure 3. Trend state selection mechanism based on the traffic changing rate

the LSTM model. This embedding technique can be defined as a static embedding technique and a dynamic embedding technique.

4.1 STATIC EMBEDDING

The static embedding method categorizes the traffic trend into a given number of states in a static manner. That is, a day is divided into a given number of time zone by considering the seasonal pattern of network traffic. For example, usually network traffic increases between 8:00 am and 11:00 am, it fluctuates between 11:00 am and 7:00 pm and it decreases between 7:00 pm and 11:00 pm. Then, the network trend can be categorized into three distinguished trend such as increasing trend, decreasing trend and fluctuating trend. Once the trend category is determined by using the time index, the traffic trend corresponding to the z_{th} time window of a day is embedded with one-hot vector encoding as $f_z = \{v_1, v_2, v_3\}$, $v \in \{0, 1\}$. For example, {1,0,0} denotes the increasing trend, {0,1,0} denotes the fluctuating trend and {0,0,1} denotes the decreasing trend.

4.2 DYNAMIC EMBEDDING

The dynamic embedding method categorizes the traffic trend into a given number of states in a dynamic manner by considering the traffic changing rate of a given time window. Figure 3 shows the state selection mechanism based on the traffic changing rate, and this mechanism calculates the changing rate between observed traffic volumes of z_{th} and $(z-n)_{th}$ time window. In here, n means how long the traffic trend is embedded, and if n is bigger, the embedded traffic trend represents longer period of time. Through the degree of the changing rate, the state of traffic trend can be categorized into five states such as Further Increasing, Increasing, Fluctuation, Decreasing, and Further Decreasing.

With this five state, we may use different embedding space such as $f_z = \{v_1, v_2\}$ (Increasing, Decreasing), $f_z = \{v_1, v_2, v_3\}$ (Increasing, Fluctuating, Decreasing), and $f_z = \{v_1, v_2, v_3, v_4, v_5\}$ (Further Increasing, Increasing, Fluctuating, Decreasing, Further Decreasing). The length of embedding space depends on the granularity of representation of traffic trend, and it may affect to the performance of LSTM based network traffic prediction. For two states embedding, the traffic trend is distinguished by checking whether the changing rate Δ_t is positive or negative. For three state embedding, the traffic trend line τ is used to categorize the traffic trend. If the changing rate Δ_t is larger than τ, the traffic trend is considered as Increasing state. If the changing rate Δ_t smaller than τ, the traffic trend is considered as Decreasing state. Otherwise, the traffic trend is considered as Fluctuation state. For five state embedding, two traffic trend lines, τ and 2τ, are considered to categorize the traffic trend.

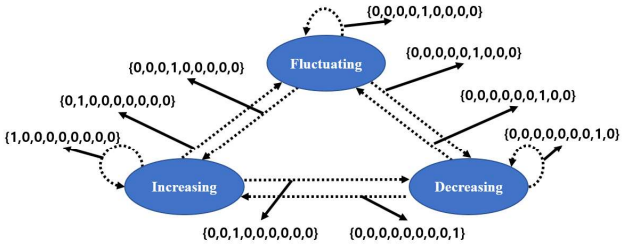

Figure 4. State Diagram with 3 Traffic Trends

4.3 CONSIDERING STATE TRANSITIONS

When preparing an input vector, the traffic trend is embedded as an encoded hot-vector. If the traffic trend is categorized in three states, the length of the encoded hot-vector is three. In here, we may focus state transitions rather than the state itself. The state transition means that the state changes between the previous time window and the current time window. For example, Figure 4 shows the state diagram of three traffic trends (Increasing, Decreasing, Fluctuation), and this diagram has nine distinguished trend state transitions. That is, if there are two, three and five states, we can consider four, nine and twenty five state transitions are considered, respectively.

Through this traffic trend state embedding method, the overall performance of the LSTM-based source-side DoS detection method can be changed. It is because that the granularity of trend embedding may affect to the performance of LSTM based model in accordance with the degree of randomness and burstiness of network traffic.

5 EVALUATION

5.1 DATASETS AND EVALUATION OVERVIEW

In order to evaluate the proposed LSTM based source-side DoS attack detection system (*LBAT*), we compare the performance of the proposed system with the seasonality-aware adaptive threshold based source-side DoS attack detection

Table 1. Datasets of DNS Request traffic

Dataset	Avg. Volume	Avg. Jitter	Burst ratio
2018 Korea	11073	0.2232	0.0004
2020 Brazil	13993	0.3585	0.0465
2020 Mexico	3537	0.5595	0.1582
2020 Slovakia	220	0.6269	0.1757

system (*STBAT*) [16], under the network traffic containing high randomness and high burstiness. The proposed LSTM based method, *LBAT*, is specified by considering different seasonality embedding methods. Through the evaluation, the LSTM based detection system with the static embedding method and the dynamic embedding method is represented as $LBAT\_S(N)$ and $LBAT\_D(N)$, respectively, where N is the number of traffic states. When we use the dynamic embedding method, n and τ set to 1 and 1%, respectively.

We evaluated the proposed method with the real-world DNS request traffic collected from DNS-STAT:Hedgehog which is operated by ICANN (Internet Corporation for Assigned Names and Numbers). As shown in Table 1, we collect DNS request traffic from four different site: Korea, Brazil, Mexico and Slovakia. Each dataset has DNS request traffic for first 10 days of year of 2018 or 2020. Table 1 shows the average of traffic volume, the average of traffic jitter and the burst ratio of traffic for each DNS request traffic. Among the datasets, 2018 Korea and 2020 Brazil have relatively low jitter and low burst ratio, but 2020 Mexico and 2020 Slovakia have high jitter and high burst ratio. That is, the former two traffic behave more likely in the linear and seasonal manner, but the latter two traffic behave more likely in the non-linear, random and bursty manner.

The traffic of the first 8 days of each dataset is used to train both of *STBAT* and *LBAT*, and the following 2days of each dataset are used as a test set for evaluating how effectively each system detect the attack traffic mixed in legitimate traffic. For evaluation, we select 180 time window indices randomly from the test set, and infuse the attack traffic by increasing the volume of the traffic up to 10%. For each dataset, we evaluate the performance of *STBAT* and *LBAT* in 5 times and summarize the result in average.

We use a detection rate, false positive rate and balanced accuracy to evaluate each method. The detection rate is the percentage of detected attack when an attack is granted, and it is considered as sensitivity. The higher the detection rate, the higher the performance of the method. The false positive rate is the percentage of falsely detected attacks when the attack does not granted. The higher the false positive rate, the lower the performance of the method. The balanced accuracy is the arithmetic mean of the detection rate (sensitivity) and true negative rate (specificity). Here, the true negative rate is calculated by subtract the false positive rate from 1. The main reason of using balanced accuracy rather than just accuracy is that the number of attack traffic is proportionally far less than the number of legitimate traffic. In the source-side network traffic, the attack traffic is much less than the legitimate traffic, and the accuracy is highly dependent on the specificity. To prevent this skewed evaluation, we use the balanced accuracy considering both of sensitivity and specificity. If the balance accuracy is high, the accuracy considering the sensitivity and specificity of the method is high.

5.2 PERFORMANCE COMPARISON

Figure 5 shows the performance comparison with different dataset. As we described earlier, datasets for Korea and Brazil has low burstiness and others has high burstiness. In Figure 5, *LB* and *HB* stands for low and high burstiness, respectively. Each detection method, we use the same value of margin, δ, as 4 which is used for adjusting detection threshold, and it is represented as $M(4)$ suffix.

It is observed that the overall detection rate decreases and the false positive rate increases as the burst ratio increases. In Korea dataset, the detection rate is around 96%, but it drops down to 65% in Slovakia dataset. For the false positive rate, it achieves around 6% in Korea dataset but it increases up to 35%. Especially, the performance degradation happens significantly when the network traffic has high burstiness and randomness.

Though the burstiness and randomness cause the performance degradation, the proposed LSTM based detection method mitigates this degradation. In Figure 5, while *STBAT* drops its detection rate from 95% down to 65%, *LBAT* keeps its detection rate around 75% and more. This endurance against the burstiness and randomness is observed in the balanced accuracy result as well. In the case of *LB*, *STBAT* and *LBAT* achieves similar balanced accuracy. But, in the case of *HB*, *LBAT* shows better performance than *STBAT*.

In order to understand the impact of burstiness to the proposed method in detail, we evaluate the performance measures with various margin for both of *STBAT* and *LBAT*. Figure 6 shows the evaluation results for 2018 Korea DNS request traffic which has low burst ratio, and Figure 7 shows the evaluation results for 2020 Mexico DNS request traffic which has high burst ratio. The result of 2020 Brazil is similar to 2018 Korea and the result of 2020 Slovakia is similar to 2020 Mexico, and we omitted these results.

In Figure 6, the detection rate and the false positive rate of *STBAT*, $LBAT\_S(9)$, and $LBAT\_D(9)$ are similar to each other, and when the margin is 4%, they are around 95% and 5%, respectively. Also, the balanced accuracy of *STBAT*, $LBAT\_S(9)$, and $LBAT\_D(9)$ are similar to each other, and when the margin is 4% each method achieves the highest balanced accuracy. Through this evaluation results, both of *STBAT* and *LBAT* properly detect attack traffic under linear and seasonal network traffic.

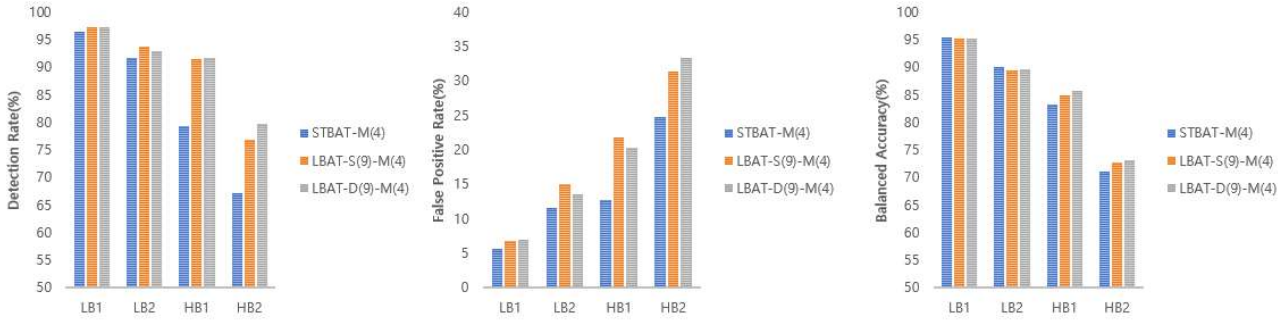

Figure 5. Performance Comparison under different burst ratio (LB1:2018 Korea, LB2:2020 Braziil, HB1:2020 Mexico, HB2:2020 Slovakia)

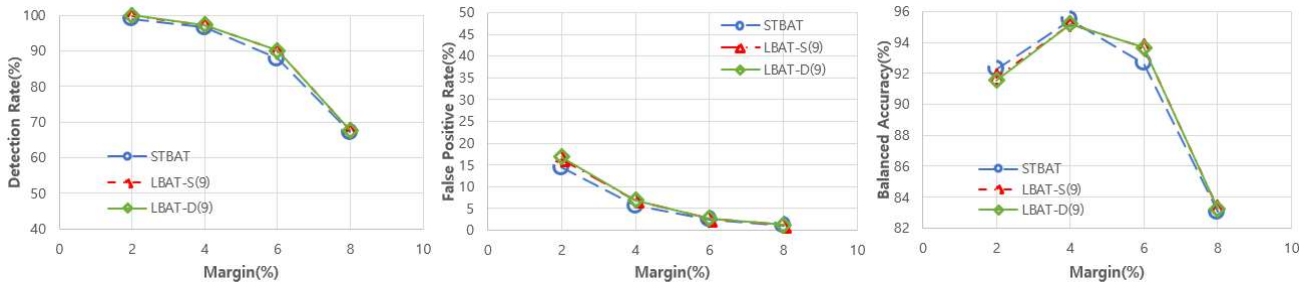

Figure 6. Performance Comparison of different methods with 2018 Korea dataset having low burst ratio

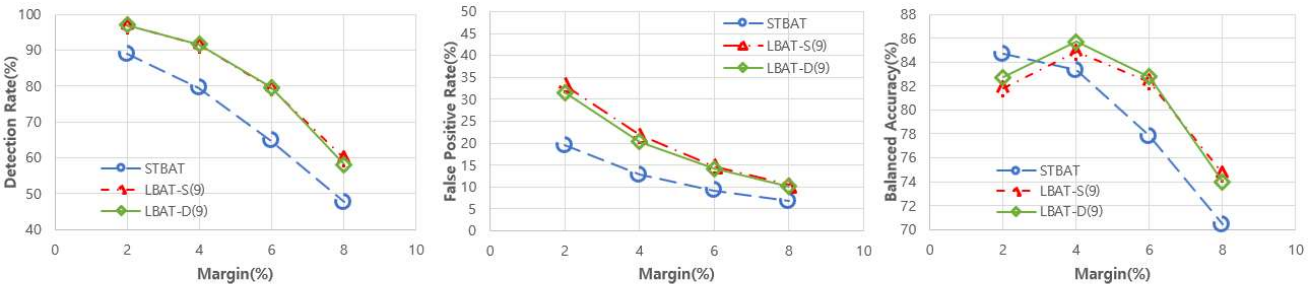

Figure 7. Performance Comparison of different methods with 2020 Mexico dataset having high burst ratio

In Figure 7, the detection rate of *STABT* is lower than both of *LBAT_S*(9), and *LBAT_D*(9). However, the false positive rate of *STABT* is also lower then both of *LBAT_S*(9), and *LBAT_D*(9). That is, *LBAT* aggressively adjust the detection threshold to detect attack traffic based on the fine grained traffic trend embedding, and it also increases the false positive rate as a side effect. However, we observed that the balanced accuracy of *LBAT* is higher than *STBAT* when the margin is more than 4%. Especially, *LBAT_D*(9) achieves the highest balanced accuracy when the margin is 4%. According to these results, we confirmed that the LSTM based source-side DoS attack detection with dynamic seasonality embedding is well suited for the network traffic with non-linear behavior, high burstiness and high randomness.

6 CONCLUSION

In this paper, we propose a LSTM-based source-side DoS attack detection method which changes the attack detection threshold adaptively even under non-linear network traffic. Through the real DNS network traffic based evaluation, the dynamic embedding of network traffic for LSTM input vectors is effective to improve the performance of LSTM based source-side DoS attack detection method under non-linear network traffic with high burstiness and randomness. Even though the burstiness of traffic is relatively low, the proposed method achieve similar performance of previous approach. Additionally, with well-adjusted margin, the proposed method keeps higher balanced accuracy than previous approach.

In the future, we are going to apply the STL (Seasonal and Trend decomposition using Losses) technique to the LSTM-based source-side DoS attack detection method in order to use more detail knowledge of traffic for training the traffic volume prediction model.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2017R1A2B4012559). This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2020-2016-0-00314) supervised by the IITP(Institute for Information & communications Technology Planning& Evaluation).

References

[1] Manos Antonakakis, Tim April, Michael Bailey, Matt Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric, J Alex Halderman, Luca Invernizzi, Michalis Kallitsis, et al. 2017. Understanding the mirai botnet. In *26th {USENIX} security symposium ({USENIX} Security 17)*. 1093–1110.

[2] A. Azzouni and G. Pujolle. 2018. NeuTM: A neural network-based framework for traffic matrix prediction in SDN. In *NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium*. 1–5.

[3] Manish Bhanu, João Mendes-Moreira, and Joydeep Chandra. 2020. Embedding Traffic Network Characteristics Using Tensor for Improved Traffic Prediction. *IEEE Transactions on Intelligent Transportation Systems* (2020).

[4] Massimo Candela, Valerio Luconi, and Alessio Vecchio. 2020. Impact of the COVID-19 pandemic on the Internet latency: a large-scale study. *arXiv preprint arXiv:2005.06127* (2020).

[5] Thomas Favale, Francesca Soro, Martino Trevisan, Idilio Drago, and Marco Mellia. 2020. Campus traffic and e-Learning during COVID-19 pandemic. *Computer Networks* (2020), 107290.

[6] Anja Feldmann, Oliver Gasser, Franziska Lichtblau, Enric Pujol, Ingmar Poese, Christoph Dietzel, Daniel Wagner, Matthias Wichtlhuber, Juan Tapidor, Narseo Vallina-Rodriguez, et al. 2020. The Lockdown Effect: Implications of the COVID-19 Pandemic on Internet Traffic. *arXiv preprint arXiv:2008.10959* (2020).

[7] Hyeokmin Gwon, Chungjun Lee, Rakun Keum, and Heeyoul Choi. 2019. Network Intrusion Detection based on LSTM and Feature Embedding. *arXiv preprint arXiv:1911.11552* (2019).

[8] Hao He and Niklas Karlsson. 2019. Identification of seasonality in Internet traffic to support control of online advertising. In *2019 American Control Conference (ACC)*. IEEE, 3835–3840.

[9] Stephen Herwig, Katura Harvey, George Hughey, Richard Roberts, and Dave Levin. 2019. Measurement and Analysis of Hajime, a Peer-to-peer IoT Botnet.. In *NDSS*.

[10] Hao Jiang and Constantinos Dovrolis. 2005. Why is the Internet Traffic Bursty in Short Time Scales? *(SIGMETRICS '05)*. Association for Computing Machinery, New York, NY, USA, 241–252. https://doi.org/10.1145/1064212.1064240

[11] H. Lu and F. Yang. 2018. Research on Network Traffic Prediction Based on Long Short-Term Memory Neural Network. In *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*. 1109–1113.

[12] Antonio Manzalini, Roberto Saracco, Cagatay Buyukkoc, Prosper Chemouil, Slawomir Kuklinski, Andreas Gladisch, Masaki Fukui, E Dekel, D Soldani, M Ulema, et al. 2013. Software-defined networks for future networks and services. In *White Paper based on the IEEE Workshop SDN4FNS*.

[13] Marco Antonio Sotelo Monge, Borja Lorenzo Fernandez, Diego Maestre Vidal, Guillermo Rius Garcia, Andres Herranz Gonzalez, and Jorge Maestre Vidal. 2018. Source-side DDoS Detection on IoT-enabled 5G Environments. In *2018 International Workshop on Secure Internet of Things (SIoT)*. IEEE, 28–37.

[14] Somenath Mukherjee, Rajdeep Ray, Rajkumar Samanta, Mofazzal H Khondekar, and Goutam Sanyal. 2017. Nonlinearity and chaos in wireless network traffic. *Chaos, Solitons & Fractals* 96 (2017), 23–29.

[15] Giang-Truong Nguyen, Van-Quyet Nguyen, Huu-Duy Nguyen, and Kyungbaek Kim. 2018. LSTM based Network Traffic Volume Prediction. In *Proceedings of 2018 KIPS Spring Conference*.

[16] Giang-Truong Nguyen, Van-Quyet Nguyen, Sinh-Ngoc Nguyen, and Kyungbaek Kim. 2019. Traffic Seasonality aware Threshold Adjustment for Effective Source-side DoS Attack Detection. *KSII Transactions on Internet & Information Systems* 13, 5 (2019).

[17] Sinh-Ngoc Nguyen, Van-Quyet Nguyen, Giang-Truong Nguyen, JeongNyeo Kim, and Kyungbaek Kim. 2018. Source-Side Detection of DRDoS Attack Request with Traffic-Aware Adaptive Threshold. *IEICE Transactions on Information and Systems* 101, 6 (2018), 1686–1690.

[18] Nipun Ramakrishnan and Tarun Soni. 2018. Network traffic prediction using recurrent neural networks. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 187–193.

[19] Sachin S Talathi and Aniket Vartak. 2015. Improving performance of recurrent neural network with relu nonlinearity. *arXiv preprint arXiv:1511.03771* (2015).

[20] R Vinayakumar, KP Soman, and Prabaharan Poornachandran. 2017. Applying deep learning approaches for network traffic prediction. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE, 2353–2358.

[21] Yuqing Wang, Dingde Jiang, Liuwei Huo, and Yong Zhao. 2019. On Reconstruction and Prediction of Network Traffic in Software Defined Networking. In *2019 IEEE International Conference on Industrial Internet (ICII)*. IEEE, 98–102.

[22] Yi-Chi Wu, Huei-Ru Tseng, Wuu Yang, and Rong-Hong Jan. 2009. DDoS detection and traceback with decision tree and grey relational analysis. In *2009 Third International Conference on Multimedia and Ubiquitous Engineering*. IEEE, 306–314.

[23] Guoxing Zhang, Shengming Jiang, Gang Wei, and Quansheng Guan. 2009. A prediction-based detection algorithm against distributed denial-of-service attacks. In *Proceedings of the 2009 international conference on wireless communications and mobile computing: Connecting the World wirelessly*. 106–110.