LSTM-DAM: Malicious Network Traffic Prediction for Cloud Manufacturing System

Longbo Zhao[1(✉)], Bohu Li[1], and Mu Gu[2]

[1] School of Automation Science and Electrical Engineering, Beihang University, Beijing, China
zlbbuaa@126.com

[2] Beijing Aerospace Smart Manufacturing Technology Development Co., Ltd., Beijing, China

Abstract. With the rapid development of Internet of Things (IoT), the applications of cloud manufacturing system are growing dramatically, resulting in increasing network heterogeneity and complexity. Network traffic prediction plays an important role in the stable operation of cloud manufacturing systems and the optimal configuration of network systems. However, existing works perform poorly confronting the data which has long time series properties and complex temporal features. To address this problem, we construct a malicious network traffic prediction model based on long and short-term memory (LSTM) neural network and dual attention mechanism. Integrated with the dual attention units of feature space and time sequence, our LSTM model can realize the dynamic correlation between malicious traffic and features series. We first obtain the weight parameters of the input data based on feature attention mechanism, and then leverage LSTM model with the attention mechanism to form a temporal attention module. These two modules strengthen the influence of key historical information. Finally, the malicious traffic prediction result of cloud manufacturing systems can be obtained from our model. The experimental results on real industrial dataset show that the prediction effect of LSTM-DAM model is better than LSTM and CNN-LSTM. Based on CIC-IDS-2017 dataset, the method also performs well in Internet malicious traffic prediction, representing great generalization ability.

Keywords: Long and short-term memory neural networks · attention mechanism · malicious traffic prediction · deep learning · cloud manufacturing system

1 Introduction

In 2009, Academician Li Bohu and other scholars and their research teams proposed the notion of "cloud manufacturing" for the first time in the world, and explained systematically the connotation system and the theoretical system and technical framework of cloud manufacturing [1]. The cloud manufacturing system has developed from the phase in which its main characteristics are networked and servitization to the current cloud manufacturing 3.0 phase with more attention to intelligence and security [2]. Due to the highly centralized management of equipment and information in the mode of cloud

Z. Xiao et al. (Eds.): ICECI 2022, LNICST 478, pp. 308–320, 2023.
https://doi.org/10.1007/978-3-031-28990-3_21

manufacturing, the security risks of system continue to rise. As the core support of the cloud manufacturing system, the intelligent cloud platform provides a full-level security protection system as an important guarantee. The research of security technology plays an important role in preventing security risks and security threats, which can effectively reduce the occurrence of security incidents such as data breaches and data corruptions of the smart manufacturing cloud platform, and provide effective security for all kinds of users of the platform while using the service.

Cloud manufacturing system relies on the network and the cloud manufacturing service platform, and invoke the manufacturing resources (manufacturing cloud) according to the customer's needs, including resource access, perception, service-oriented and other levels, highly openness. With the rapid expansion of the Internet of Things (IoT) [3–5], the industrial Internet and cloud manufacturing system technology is accelerating. The Industrial Internet promotes intelligent production and realizes inter-industry communication and resource sharing [6, 7]. The cloud manufacturing system is a service-oriented digital, networked, and intelligent organization. This design integrates advanced information and communication technologies and manufacturing science such as IoT, high-performance cloud computing [8–10] and heterogeneous network resources [10, 11]. In this mode, the system utilizes virtualizing, resource pooling and other techniques to convert manufacturing products, resources and capabilities into manufacturing cloud services. The purpose of the construction is to provide users with various intelligent services on demand through centralized management and operation of cloud manufacturing services.

However, due to the growth of the Industrial Internet community, the scale of network traffic and the complexity of network have continued to increase. The network attacks against the cloud manufacturing field have become frequent and intensive. Therefore, it is necessary to take appropriate measures to predict malicious traffic to improve the security and attack resistance of the production system. Predicting malicious traffic can avoid problems that are easy to occur on the cloud manufacturing system. It can effectively optimize and adjust network resources, and further ensure network connections of important nodes. In addition, if the malicious traffic change can be accurately predicted, then it can reduce network congestion in advance which will help system enhance network performance and block network intrusion.

In the field of cloud manufacturing, the randomness of network attack behavior is high, and the impact of the input characteristics of malicious traffic on the prediction results is constantly changing over time. Therefore, the temporal correlation of the input data will also have an impact on the prediction results. Aiming at the characteristics of cloud manufacturing system malicious traffic and the problems in existing work, we propose a long and short-term memory (LSTM) neural network based on dual attention mechanism to improve the accuracy of malicious traffic prediction. With the attention mechanism, our method can learn the malicious traffic sequence of the Industrial Internet and assign greater parameter weights to key information. The design not only can fully exploit the connection before and after the malicious traffic time series, but also completely learn the overall characteristics of the malicious traffic. It can also avoid the defect that the weight of key features with a high impact factor on the accuracy of the results is diluted during the training iteration.

The remainder of this paper is organized as follows. Section 2 reviews methods and analyzes their differences in the application. Section 3 proposes an improvement of LSTM network integrated with dual attention mechanism. Section 4 describes the simulation setup and experiment results. The final section gives the conclusion of the whole paper.

2 Related Work

Network traffic prediction is a significant subfield of network traffic monitoring and analysis which is mainly focused on predicting the future of network load and its behavior [12, 13]. Network traffic prediction has become an important work in current network security research, which can support the effective defense of industrial IoT computer systems against various types of network attacks [14–16].

Traditional methods of network traffic prediction are currently divided into two main categories, linear prediction methods and nonlinear prediction methods. Linear prediction methods include the use of Markov models [17–19] and exponential smoothing [20, 21], etc. Linear methods can only be trained for network traffic features in a single dimension. Nonlinear prediction methods mainly include machine learning [22–25] and deep learning [26–29]. Particularly, deep learning has been increasingly studied on the field of time series prediction. [26] proposed a spatio-temporal convolutional network (LA-ResNet) which uses an attention mechanism to solve spatio-temporal modeling and predict wireless network traffic. [27] proposed a new network traffic prediction method based on ESN with adaptive reservoir (ESN-AR), ESN has strong nonlinear processing capability and short-term memory, which can achieve good performance in predicting nonlinear time series. [28] investigated a transfer learning strategy based on graph convolution neural network to achieve the task of large-scale traffic prediction. [29] proposed a new method using an enhanced deep reinforcement learning (EDRL) algorithm to enable intelligence-based network traffic prediction and solve network management problems.

LSTM with improved structure by Recurrent Neural Network (RNN) [30] is suitable for network traffic prediction with time-series features. [31] proposed a double LSTM structure, one of which acts as the main flow predictor, another as the detector of the time the burst flow starts at. The two LSTM units can exchange information about their internal states. [32] proposed a neural network model based on LSTM and transfer learning which can address the problem of small sample size in network traffic prediction. [33] investigated a radial kernelized LSTM-based connectionist Tversky multilayer deep structure learning (RKLSTM-CTMDSL) model to solve network traffic prediction. [34] proposed a novel hybrid prediction method ST-LSTM for such network traffic prediction, which synergistically combines the power of the Savitzky–Golay (SG) filter, the TCN and LSTM.

The above works are all traffic prediction researches in a general Internet environment. However, when the input data has long time series properties and the data characteristics are more complex, none of the existing methods can make accurate prediction. For cloud manufacturing systems [35] integrating cloud computing, IoT, virtualization and intelligent science, higher requirements are put forward for network security. Therefore, a more accurate prediction of malicious traffic in cloud manufacturing platform is eagerly needed to maintain stable operation of the system.

3 Model Design

For the malicious traffic prediction problem of the cloud manufacturing system, we propose the LSTM-DAM (Long Short-Term Memory network based on Dual Attention Mechanism) prediction model which contain the feature attention and the temporal attention modules.

3.1 Feature Attention Mechanism

The input of the feature attention mechanism is a single-step vector containing M features $x_t = [x_{1,t}, x_{2,t}, \cdots, x_{M,t}]$, the attention weight vector e_t is calculated using a single-layer neural network as:

$$e_t = \sigma\left(W_e x_t + b_e\right) \tag{1}$$

where $e_t = [e_{1,t}, e_{2,t}, \cdots, e_{M,t}]$ is the attention weights corresponding to the M input features, W_e is the weight matrix, b_e is the bias vector, and $\sigma(\cdot)$ is the Sigmoid activation function. The feature attention weights $\omega_t = [\omega_{1,t}, \omega_{2,t}, \cdots, \omega_{M,t}]$ are obtained by using the Softmax function, the subitem of ω_t is:

$$\omega_{m,t} = \frac{\exp(e_{m,t})}{\sum_{i=1}^{M} e_{i,t}} \tag{2}$$

After the attention parameter weights are acquired, they are associated with each input feature and finally input to the LSTM network for learning training.

$$x'_{m,t} = x_{m,t}\omega_{m,t} \tag{3}$$

Figure 1 shows the feature attention mechanism model. The parameter weights of traditional LSTM models are shared among different features, and the output results are influenced by each feature to the same extent. However, most of the input features in practical scenarios have multi-level and multi-dimensional characteristics, and each feature has different degrees of influence on the output results. These existence will finally cause inefficient LSTM performance. Therefore, we use the feature attention mechanism to screen out the redundant information of feature sequences and thus focus on the important features.

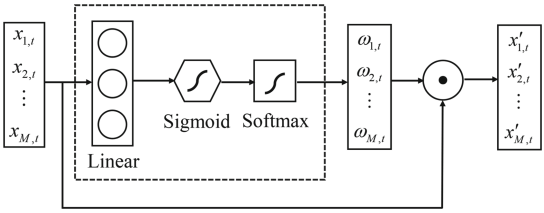


Fig. 1. Feature attention mechanism

3.2 Temporal Attention Mechanism

The input of the temporal attention mechanism is the hidden layer state $h_t = [h_{1,t}, h_{2,t}, \cdots, h_{K,t}]$ of the LSTM network at moment t. K is the length of the time of the input sequence. Define the historical temporal attention weights as:

$$s_t = ReLU(W_d h_t + b_d) \tag{4}$$

where $s_t = [s_{1,t}, s_{2,t}, \cdots, s_{K,t}]$, W_d is the weight matrix, b_d is the bias vector, and $RELU(\cdot)$ is the activation function to increase the feature difference. The temporal attention weight is $\mu_t = [\mu_{1,t}, \mu_{2,t}, \cdots, \mu_{K,t}]$, the subitem of μ_t is:

$$\mu_{k,t} = \frac{\exp(s_{k,t})}{\sum_{i=1}^{K} s_{i,t}} \tag{5}$$

We can obtain the integrated timing state h'_t as:

$$h'_t = \mu_t \otimes h_t = \sum_{k=1}^{K} \mu_{k,t} h_{k,t} \tag{6}$$

Figure 2 shows the temporal attention mechanism model. The feature attention mechanism only correlates the input features of the LSTM network before training with the target features, so the input features have the same weight at any moment. However in actual system, the correlation between cloud manufacturing system malicious traffic and feature sequences tends to change dynamically in real-time. Therefore, we introduce the temporal attention mechanism that can automatically extract the sequence data between each historical moment to further improve the information representation at important moments.

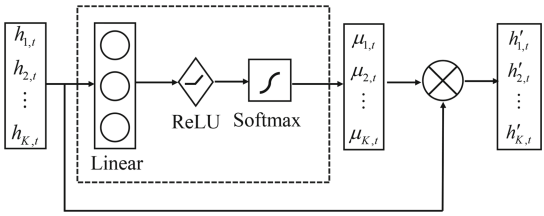


Fig. 2. Temporal attention mechanism

3.3 LSTM-DAM Model

The proposed LSTM-DAM model includes an input layer, a feature attention layer, an LSTM layer, a temporal attention layer, and a fully connected layer. Figure 3 shows the complete structure of the LSTM-DAM model.

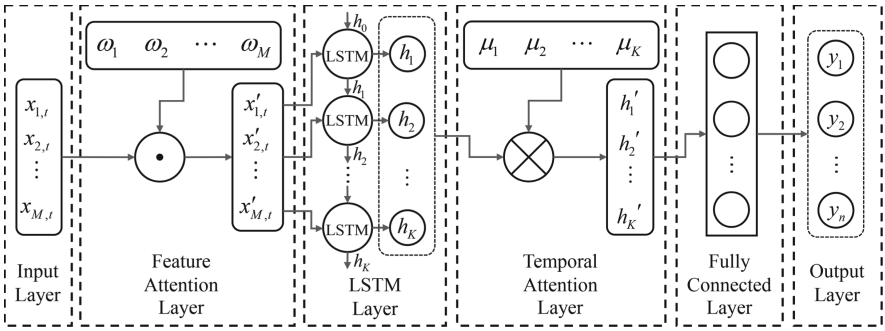


Fig. 3. LSTM-DAM Model Structure

The LSTM-DAM input layer sends the processed network traffic data history sequences combined with the feature sequences to the feature attention layer. Then the model uses the feature attention mechanism to perform dynamic weight assignment for the features to achieve the focus of key features. The LSTM layer performs learning calculations on the input sequences, which can obtain the hidden layer state h. After that, the temporal attention layer performs the periodic trend feature of the feature sequences learning to realize the weight assignment of temporal attention, thus improving the network expression capability. Finally, the temporal attention module inputs the global hidden layer state to the fully connected layer to complete the final malicious traffic prediction work.

4 Experiments Evaluation

In this section, we use recent datasets to evaluate our LSTM-DAM model. Firstly, we describe the detail of the datasets and the pre-processing method. Secondly, we show the metrics to evaluate our method's performance. Afterward, we demonstrate the key experiment setup. Finally, we justify our method in predicting malicious traffic with superior performance over other comparative methods.

4.1 Datasets

Cloud Manufacturing System. The source of the dataset is a real Industrial Internet and cloud manufacturing platform. Data collection is performed at the sampling point every minute from 0:00 on April 5, 2022 to 24:00 on April 23. Then, the private dataset with multiple labels including timestamp, historical traffic value and historical malicious traffic value is formed. After data cleaning, the dataset has 27342 valid samples.

CIC-IDS-2017. The CIC-IDS-2017 dataset is often used by researchers as experimental data for network intrusion detection, which contains both benign Internet data and common abnormal data. Similar to real-world data (PCAPs). The dataset contains various attack types such as brute force FTP, brute force SSH, DoS, Heartbleed, penetration, botnet and DDoS. We utilize four labels to clean the dataset: timestamp, flow duration, flow bytes/s and label. Finally, 1966 valid data are obtained.

Datasets preprocessing includes two phases: division and converting: firstly, the datasets are randomly divided into training data and test data in a ratio of 9:1. Then, we convert the time series data into sequence data containing pairs of input and output. For a given data, the converting method is as follows. We copy the data column and move the replicated data of the column forward or backward by N times. The data gaps generated after the movement will be filled with NaN. After the above steps, the lag value data column is completed, and a data format with supervised learning attributes is obtained. The purpose of converting data is to transform the problem into a supervised learning problem. The processing process is shown in Fig. 4.

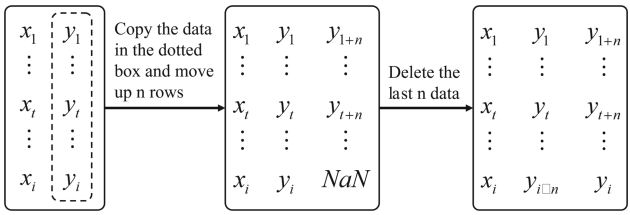


Fig. 4. Dataset processing

4.2 Metrics

Concerning malicious traffic prediction tasks in cloud manufacturing system, the results can be correct or incorrect. Therefore, the evaluation of the performance of LSTM-DAM is grounded on the accuracy of the prediction results. Related study generally uses the error-index to verify the performance of the model. The larger the error value, the lower the prediction accuracy, which also means the worse the performance of the experimental model. All results correspond to the following three outcomes:

1) Mean Absolute Error (MAE): It is the average value of absolute error. The index value can intuitively reflect the prediction error. The value range of MAE is [0, +

∞], 0 means the predicted value is consistent with the actual value. The larger the error, the larger MAE.

$$MAE = \frac{1}{K}\sum_{i=1}^{k}\left|x_i - x_i'\right|^2 \tag{7}$$

2) Root Mean Square Error (RMSE): It is also known as standard error, its range is [0, $+\infty$], the same as MAE, 0 means that there is no error in the model result. The larger the error, the larger RMSE.

$$RMSE = \sqrt{\frac{1}{K}\sum_{i=1}^{k}\left|x_i - x_i'\right|^2} \tag{8}$$

3) R-squared: It is the accuracy to convert the prediction result into a standard, and its value is [0, 1]. It can be used to determine which type of prediction problem the model is more suitable for. The larger the value of R^2, the better performance of the proposed model to fit the real value, when $R^2 = 1$, it is a perfect model.

$$R^2 = 1 - \frac{\sum_{i=1}^{k}(x_i - x_i')^2}{\sum_{i=1}^{k}(x_i - \overline{x_i})^2} \tag{9}$$

4.3 Experiment Setup

The LSTM-DAM model uses the Adam algorithm to adjust parameters for units, epochs, and batch size. Units represent the output dimension, which is the number of hidden neurons in the feedforward neural network in the LSTM neural network. Epochs represent the total number of rounds of training. We introduce the early stop mechanism to effectively avoid the result overfitting and long training time caused by manually inputting the epochs value. Batch size indicates the number of samples used in each batch for gradient descent in model training. The gradient descent is calculated to optimize the objective function when each batch sample is trained. After multiple rounds of testing, the parameters are set that the units value is 32, the stop training condition is 20 rounds and the batch size value is 32.

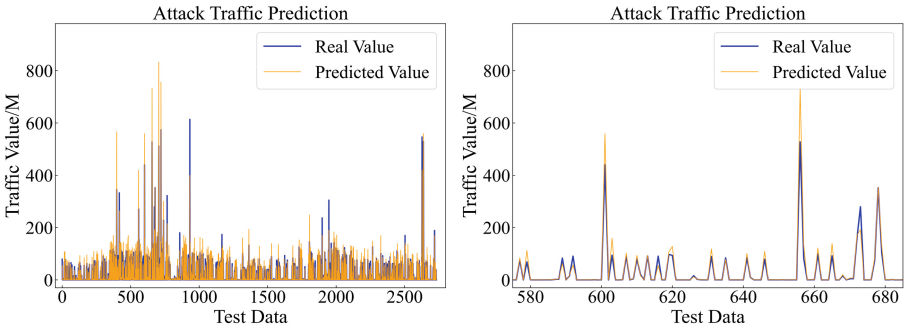
Under the situation of the above parameter values, set the time step parameter N in the interval [2, 7]. The prediction results with different values of N on the cloud manufacturing system dataset are shown in Table 1. Note that in the conducted experiments, when the time step N is 5, the MAE and RMSE results achieve the best 0.872 and 1.685 respectively. Therefore, the step parameter N is set to be 5.

Table 1. Time-step Tuning Result.

Time step	MAE	RMSE
n = 2	0.862	1.731
n = 3	0.853	1.725
n = 4	0.846	1.694
n = 5	0.827	1.685
n = 6	0.881	1.752
n = 7	0.832	1.690

4.4 Result Evaluation

To verify the performance improvement of the LSTM-DAM model, we use the traditional LSTM model and the CNN-LSTM model as the experimental comparison model. The prediction results on the cloud manufacturing system dataset in the proposed method and related research works are represented in Figs. 5, 6 and 7.

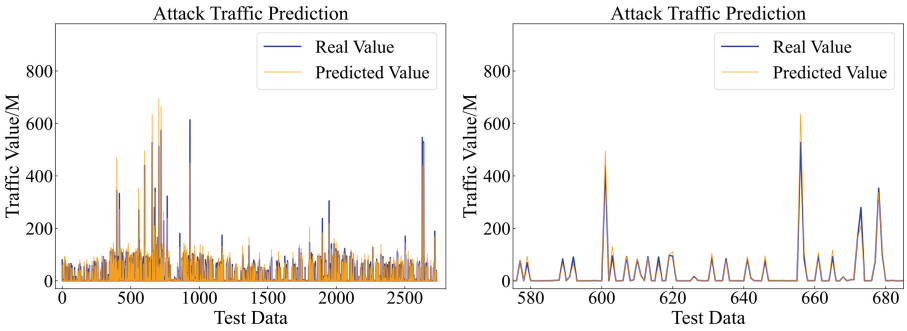


(a) Full Results (b) Partial zoom results

Fig. 5. LSTM model results for Cloud manufacturing system dataset

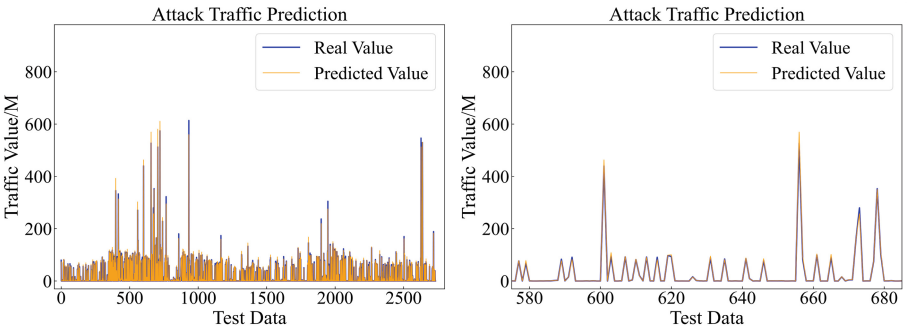
The results show that the LSTM-DAM model after the introduction of the dual attention mechanism has higher prediction accuracy than the traditional LSTM model. Although The trend of the prediction curve of the CNN-LSTM model is consistent with the actual value curve, the error of the prediction result is large, especially at the peak. Therefore, the performance of the prediction model is not ideal, and the overall accuracy of the prediction value is poorer than the LSTM-DAM model. In terms of the results, it shows that the accuracy and stability of the proposed LSTM-DAM model has been significantly improved. Moreover, the prediction results at the peak of the curve are more outstanding than the CNN-LSTM model.

To further demonstrate the outperformance of the LSTM-DAM model, we calculate the metrics of different models. The results are shown in Table 2. We can observe that the LSTM-DAM model has the highest prediction accuracy. Specifically, the CNN-LSTM



(a) Full Results (b) Partial zoom results

Fig. 6. CNN-LSTM model results for Cloud manufacturing system dataset



(a) Full Results (b) Partial zoom results

Fig. 7. LSTM-DAM model results for Cloud manufacturing system dataset

model extracts the input feature weights through the CNN network, and then combines them with LSTM to predict the Industrial Internet malicious traffic. The results show that the prediction accuracy is significantly higher than the traditional LSTM model after the feature attention layer is introduced. Compared with LSTM, the error of the LSTM-DAM model proposed in this paper is greatly reduced. The MAE and RMSE are decreased by 0.749 and 10.429 respectively, and the R-squared is 8.2% higher than that of LSTM. Compared with CNN-LSTM, LSTM-DAM model improves by 4% on R-squared, respectively drops 0.445 and 10.241 on MAE and RMSE.

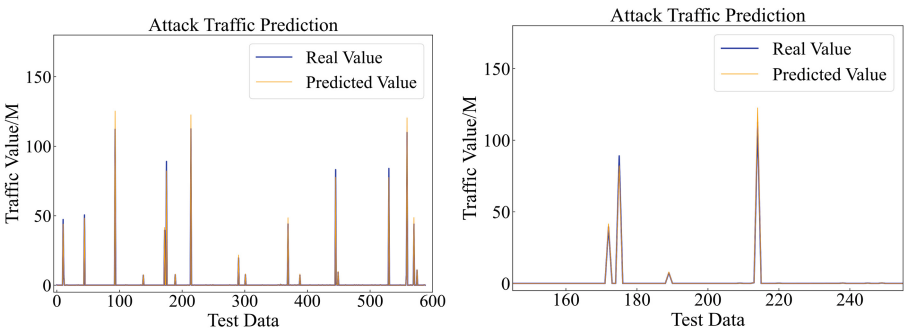
Table 2. Result Comparison.

Prediction Model	R-Squared	MAE	RMSE
LSTM	90.4%	1.576	12.114
CNN-LSTM	94.6%	1.272	11.926
LSTM-DAM	98.6%	0.827	1.685

In terms of prediction performance, LSTM-DAM has high accuracy. The prediction results are significantly better than LSTM and CNN-LSTM models. The proposed model can achieve tremendous capability in the field of Industrial Internet malicious traffic prediction.

4.5 Generalization Analysis

Concerning the Internet malicious traffic has high similarity to the malicious traffic in cloud manufacturing system, we consider using the CIC-IDS-2017 dataset to verify the generalization ability of the proposed LSTM-DAM model. The results are shown in Fig. 8. The predicted value of the LSTM-DAM model is consistent with the actual value. Meanwhile, the predicted curve is stable without large fluctuations and the error value is small at the peak value.



(a) Full Results (b) Partial zoom results

Fig. 8. LSTM-DAM model results for CIC-IDS-2017 dataset

Table 3 illustrates the result of CIC-IDS-2017 dataset. R-squared of the LSTM-DAM model is 98.5%, showing a high fitness. MAE and RMSE are 0.13 and 0.26, which represent the overall error is small and stable. The experimental results justify that our method is capable of predicting malicious network traffic in a different network environment.

Table 3. Prediction Result.

Prediction Model	R-Squared	MAE	RMSE
LSTM-DAM	98.5%	0.13	0.26

5 Conclusion

Security is the life of cloud manufacturing, and it is also the basic needs of every provider, users (including both manufacturing enterprise and manufacturing product) and operators of cloud manufacturing service. The real-time security status' monitoring and

prediction of cloud manufacturing systems in the network environment has important practical significance. For industrial network malicious traffic mostly concentrated in the field of detection, there is a shortage of research for the prediction of malicious traffic values. We propose a LSTM network malicious traffic prediction model based on dual attention mechanism for the problem that existing methods cannot handle complex data features under long time sequences in the cloud manufacturing system. The experimental results show that the R-squared, MAE and RMSE performance indexes are significantly improved when compared with LSTM and CNN-LSTM models. Our method can fit the relationship between historical data and network malicious traffic well, which can improve the prediction accuracy of future malicious traffic. The generalization ability of the model is also verified using two different data sets, which further proves the effectiveness and superiority of this paper's model of predicting malicious traffic in cloud manufacturing system.

References

1. Li, B., Zhang, L., Wang, S., et al.: Cloud manufacturing: a new service-oriented networked manufacturing model. Comput. Integr. Manuf. Syst. **16**(01), 1–7+16 (2010)
2. Li, B., Chai, X., Hou, B., et al.: Cloud manufacturing system 3.0——new intelligent manufacturing system in era of intelligence +. Comput. Integr. Manuf. Syst. **25**(12), 2997–3012 (2019)
3. Jiang, H., Xiao, Z., Li, Z., et al.: An energy-efficient framework for internet of things underlaying heterogeneous small cell networks. IEEE Trans. Mob. Comput. **21**(1), 31–43 (2022)
4. Dai, X., Xiao, Z., Jiang, H., et al.: Task co-offloading for D2D-assisted mobile edge computing in industrial internet of things. IEEE Trans. Ind. Inform. 1 (2022)
5. Jiang, H., Dai, X., Xiao, Z., et al.: Joint task offloading and resource allocation for energy-constrained mobile edge computing. IEEE Trans. Mob. Comput. 1 (2022)
6. Hu, Z., Zeng, F., Xiao, Z., et al.: Computation efficiency maximization and QoE-provisioning in UAV-enabled MEC communication systems. IEEE Trans. Netw. Sci. Eng. **8**(2), 1630–1645 (2021)
7. Zhang, W., Zhou, S., Yang, L., et al.: WiFiMap+: high-level indoor semantic inference with WiFi human activity and environment. IEEE Trans. Veh. Technol. **68**(8), 7890–7903 (2019)
8. Xiao, Z., Chen, Y., Jiang, H., et al.: Resource management in UAV-assisted MEC: state-of-the-art and open challenges. Wirel. Netw. **28**(7), 3305–3322 (2022)
9. Ali, T.A.A., Xiao, Z., Sun, J., et al.: Optimal design of IIR wideband digital differentiators and integrators using salp swarm algorithm. Knowl. Based Syst. **182**, 104834 (2019)
10. Xiao, Z., Li, F., Jiang, H., et al.: A joint information and energy cooperation framework for CR-enabled macro–femto heterogeneous networks. IEEE Internet Things J. **7**(4), 2828–2839 (2020)
11. Zeng, F., Li, Q., Xiao, Z., et al.: A price-based optimization strategy of power control and resource allocation in full-duplex heterogeneous macrocell-femtocell networks. IEEE Access **6**, 42004–42013 (2018)
12. Lohrasbinasab, I., Shahraki, A., Taherkordi, A., et al.: From statistical- to machine learning-based network traffic prediction. Trans. Emerg. Telecommun. Technol. **33**(4) (2022)
13. Long, W., Xiao, Z., Wang, D., et al.: Unified spatial-temporal neighbor attention network for dynamic traffic prediction. IEEE Trans. Veh. Technol. 1–15 (2022)

14. Mohammadi, M., Al-Fuqaha, A., Sorour, S., et al.: Deep learning for IoT big data and streaming analytics: a survey. IEEE Commun. Surv. Tutor. **20**(4), 2923–2960 (2018)
15. Zhang, X.Y., Wu, Z.J., Zhang, J.W., et al.: An adaptive network traffic prediction approach for LDoS attacks detection. Int. J. Commun. Syst. **31**(5) (2018)
16. Zhao, P., Jiang, H., Li, J., et al.: Synthesizing privacy preserving traces: enhancing plausibility with social networks. IEEE/ACM Trans. Netw. **27**(6), 2391–2404 (2019)
17. Chen, Z.T., Wen, J.Y., Geng, Y.H.: Predicting future traffic using hidden markov models. In: 2016 IEEE 24th International Conference on Network Protocols (ICNP) (2016)
18. Tian, Z.D.: Network traffic prediction method based on wavelet transform and multiple models fusion. Int. J. Commun. Syst. **33**(11) (2020)
19. Guarino, I., Nascita, A., Aceto, G., et al.: Mobile network traffic prediction using high order Markov chains trained at multiple granularity, pp. 394–399 (2021)
20. Tran, Q.T., Hao, L., Trinh, Q.K.: Cellular network traffic prediction using exponential smoothing methods. J. Inf. Commun. Technol. Malays. **18**(1), 1–18 (2019)
21. Andrysiak, T., Saganowski, L., Kiedrowski, P.: Predictive Abuse Detection for a PLC Smart Lighting Network Based on Automatically Created Models of Exponential Smoothing. Security and Communication Networks (2017)
22. Wang, Q.-M., Fan, A., Shi, H.: Network traffic prediction based on improved support vector machine. Int. J. Syst. Assur. Eng. Manag. **8**(3s), 1976–1980 (2017)
23. Wang, Y., Nakachi, T.: Prediction of network traffic through light-weight machine learning. IEEE Open J. Commun. Soc. **1**, 1919–1933 (2020)
24. Szostak, D.: Machine learning ensemble methods for optical network traffic prediction, pp. 105–115 (2021)
25. Ke, G., Chen, R.-S., Ji, S., et al.: Network traffic prediction based on least squares support vector machine with simple estimation of Gaussian kernel width. Int. J. Inf. Comput. Secur. **18**(1/2), 1–11 (2022)
26. Li, M., Wang, Y., Wang, Z., et al.: A deep learning method based on an attention mechanism for wireless network traffic prediction. Ad Hoc Netw. **107**, 102258 (2020)
27. Zhou, J., Wang, H., Xiao, F., et al.: Network traffic prediction method based on echo state network with adaptive reservoir. Softw. Pract. Exp. **51**(11), 2238–2251 (2021)
28. Zhou, X., Zhang, Y., Li, Z., et al.: Large-scale cellular traffic prediction based on graph convolutional networks with transfer learning. Neural Comput. Appl. **34**(7), 5549–5559 (2022)
29. Balamurugan, N.M., Adimoolam, M., Alsharif, M.H., et al.: A novel method for improved network traffic prediction using enhanced deep reinforcement learning algorithm. Sensors **22**(13), 5006 (2022)
30. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
31. Huang, L., Wang, D., Liu, X., et al.: Double LSTM structure for network traffic flow prediction, pp. 380–388 (2020)
32. Wan, X., Liu, H., Xu, H., et al.: Network traffic prediction based on LSTM and transfer learning. IEEE Access **10**, 86181–86190 (2022)
33. Govindarajan, M., Chandrasekaran, V., Anitha, S.: Network traffic prediction using radial kernelized-tversky indexes-based multilayer classifier. Comput. Syst. Sci. Eng. **40**(3), 851–863 (2022)
34. Bi, J., Zhang, X., Yuan, H.T., et al.: A hybrid prediction method for realistic network traffic with temporal convolutional network and LSTM. IEEE Trans. Autom. Sci. Eng. **19**(3), 1869–1879 (2022)
35. Liao, Y.X., Panetto, H., Stadzisz, P.C., et al., A notification-oriented solution for data-intensive enterprise information systems - a cloud manufacturing case. Enterp. Inf. Syst. **12**(8–9), 942–959 (2018)