ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940
www.ijasem.org
Vol 18, Issue.2 April 2024

ABNORMAL TRAFFIC DETECTION BASED ON ATTENTION AND BIG STEP CONVOLUTION

Madhani Sai Srikari [1], Pampana Ravi Kiran [2],
E. Emmanuel Joshua [3], Dr. Pradeep Kumar[4]
[1,2,3] UG Student, Dept. of ECE, CMR Institute of Technology, Hyderabad
[4] Associate Professor, Dept. of ECE
CMR Institute of Technology, Hyderabad

ABSTRACT

Abnormal traffic detection is critical to network security and quality of service. However, the similarity of features and the single dimension of the detection model cause great difficulties for abnormal traffic detection, and thus a big-step convolutional neural network traffic detection model based on the attention mechanism is proposed. Firstly, the network traffic characteristics are analyzed and the raw traffic is preprocessed and mapped into a two-dimensional grayscale image. Then, multi-channel grayscale images are generated by histogram equalization, and an attention mechanism is introduced to assign different weights to traffic features to enhance local features. Finally, pooling-free convolutional neural networks are combined to extract traffic features of different depths, thus improving the defects such as local feature omission and overfitting in convolutional neural networks. The simulation experiment was carried out in a balanced public data set and an actual data set. Using the commonly used algorithm SVM as a baseline, the proposed model is compared with ANN, CNN, RF, Bayes and two latest models. Experimentally, the accuracy rate with multiple classifications is 99.5%. The proposed model has the best anomaly detection. And the proposed method outperforms other models in precision, recall, and F1. It is demonstrated that the model is not only efficient in detection, but also robust and robust to different complex environments.

INTRODUCTION

Internet technology is widely used in all walks of life, and has strongly contributed to the development of economy and society. However, as the current mainstream network security and defense technologies still have many shortcomings, the huge application requirements also make the security configuration of the entire network becomes particularly complex, resulting in the entire network facing the threat of extremely vulnerable to attacks. At the same time, due to the openness of the TCP/IP network architecture, computer

ISSN2454-9940
www.ijasem.org
Vol 18, Issue.2 April 2024

viruses spread more widely through disguise, which affects the normal operation of the network and causes social and economic downturn. How to take effective methods to analyze data information to predict the current network development, find abnormalities and take appropriate handling measures is of great significance to maintain network security [1]. The associate editor coordinating the review of this manuscript and approving it for publication was S. K. Hafizul Islam . Anomalous traffic detection can be achieved with the help of network traffic classification. According to its core idea there are mainly the following approaches: port-based [2], deep packet detection based [3], and machine learning based [4].

LITERATURE REVIEW

Explainable Internet Traffic Classification

The problem analyzed in this paper deals with the classification of Internet traffic. During the last years, this problem has experienced a new hype, as classification of Internet traffic has become essential to perform advanced network management. As a result, many different methods based on classical Machine Learning and Deep Learning have been proposed. Despite the success achieved by these techniques, existing methods are lacking because they provide a classification output that does not help practitioners with any information regarding the criteria that have been taken to the given classification or what information in the input data makes them arrive at their decisions. To overcome these limitations, in this paper we focus on an "explainable" method for traffic classification able to provide the practitioners with information about the classification output. More specifically, our proposed solution is based on a multi-objective evolutionary fuzzy classifier (MOEFC), which offers a good trade-off between accuracy and explainability of the generated classification models. The experimental results, obtained over two well-known publicly available data sets, namely, UniBS and UPC, demonstrate the effectiveness of our method.

Multi-Level P2P Traffic Classification Using Heuristic and Statistical-Based Techniques: A Hybrid Approach

Peer-to-peer (P2P) applications have been popular among users for more than a decade. They consume a lot of network bandwidth, due to the fact that network administrators face several issues such as congestion, security, managing resources, etc. Hence, its accurate classification will

ISSN2454-9940
www.ijasem.org
Vol 18, Issue.2 April 2024

allow them to maintain a Quality of Service for various applications. Conventional classification techniques, i.e., port-based and payload-based techniques alone, have proved ineffective in accurately classifying P2P traffic as they possess significant limitations. As new P2P applications keep emerging and existing applications change their communication patterns, a single classification approach may not be sufficient to classify P2P traffic with high accuracy. Therefore, a multi-level P2P traffic classification technique is proposed in this paper, which utilizes the benefits of both heuristic and statistical-based techniques.

Edge Computing Intelligence Using Robust Feature Selection for Network Traffic Classification in Internet-of-Things

Internet-of-Things (IoT) devices are massively interconnected, which generates a massive amount of network traffic. The concept of edge computing brings a new paradigm to monitor and manage network traffic at the network's edge. Network traffic classification is a critical task to monitor and identify Internet traffic. Recent traffic classification works suggested using statistical flow features to classify network traffic accurately using machine learning techniques. The selected classification features must be stable and can work across different spatial and temporal heterogeneity. This paper proposes a feature selection mechanism called Ensemble Weight Approach (EWA) for selecting significant features for Internet traffic classification based on multi-criterion ranking and selection mechanisms.

On Internet Traffic Classification: A Two-Phased Machine Learning Approach

Traffic classification utilizing flow measurement enables operators to perform essential network management. Flow accounting methods such as NetFlow are, however, considered inadequate for classification requiring additional packet-level information, host behaviour analysis, and specialized hardware limiting their practical adoption. This paper aims to overcome these challenges by proposing two-phased machine learning classification mechanism with NetFlow as input. The individual flow classes are derived per application through -means and are further used to train a C5.0 decision tree classifier.

CTTGAN: Traffic Data Synthesizing Scheme Based on Conditional GA

ISSN2454-9940
www.ijasem.org
Vol 18, Issue.2 April 2024

Most machine learning algorithms only have a good recognition rate on balanced datasets. However, in the field of malicious traffic identification, benign traffic on the network is far greater than malicious traffic, and the network traffic dataset is imbalanced, which makes the algorithm have a low identification rate for small categories of malicious traffic samples. This paper presents a traffic sample synthesizing model named Conditional Tabular Traffic Generative Adversarial Network (CTTGAN), which uses a Conditional Tabular Generative Adversarial Network (CTGAN) algorithm to expand the small category traffic samples and balance the dataset in order to improve the malicious traffic identification rate. The CTTGAN model expands and recognizes feature data, which meets the requirements of a machine learning algorithm for training and prediction data. The contributions of this paper are as follows: first, the small category samples are expanded and the traffic dataset is balanced; second, the storage cost and computational complexity are reduced compared to models using image data; third, discrete variables and continuous variables in traffic feature data are processed at the same time, and the data distribution is described well. The experimental results show that the recognition rate of the expanded samples is more than 0.99 in MLP, KNN and SVM algorithms. In addition, the recognition rate of the proposed CTTGAN model is better than the oversampling and undersampling schemes.

Network traffic classification: Techniques, datasets, and challenges

In network traffic classification, it is important to understand the correlation between network traffic and its causal application, protocol, or service group, for example, in facilitating lawful interception, ensuring the quality of service, preventing application choke points, and facilitating malicious behavior identification. In this paper, we review existing network classification techniques, such as port-based identification and those based on deep packet inspection, statistical features in conjunction with machine learning, and deep learning algorithms. We also explain the implementations, advantages, and limitations associated with these techniques. Our review also extends to publicly available datasets used in the literature. Finally, we discuss existing and emerging challenges, as well as future research directions.

ISSN2454-9940
www.ijasem.org
Vol 18, Issue.2 April 2024

An ICS Traffic Classification Based on Industrial Control Protocol Keyword Feature Extraction Algorithm

Industrial control protocol feature extraction is an important way to improve the accuracy and speed of industrial control protocol traffic classification. This paper firstly proposes a keyword feature extraction method for industrial control protocol, and then designs and implements an industrial control system (ICS) traffic classification based on this method. The proposed method utilizes the characteristics of the relatively fixed format of the industrial control protocol and the periodicity of the protocol traffic in ICS. The keyword features of the industrial control protocol can be accurately extracted after data preprocessing, data segmentation, redundant data filtering, and feature byte mining. A feature dataset is then formed. The designed ICS traffic classifier adopts decision tree and is trained with the feature dataset. Experiments are carried out on the open-source dataset. The results show that the proposed method achieves 99.99% classification accuracy, and the classification precision and classification recall rate reach 99.98% and 99.93%, respectively. The training time and predicting time of classifier are 0.34 s and 0.264 s, respectively, which meets the requirements of high precision and low latency of industrial control system.

EXISTING SYSTEM

Shi et al. [16] proposed a cost-sensitive SVM (CMSVM) for the network traffic imbalance problem. The model uses a multi-class SVM with an active learning algorithm to solve the imbalance problem for different applications by adaptive weights. Cao et al. [17] proposed a real-time network classification model with SPPSVM. The model uses the feature selection method of principal component analysis (PCA) to reduce the dimensionality of the original data and uses an improved particle swarm optimization algorithm to obtain the optimal parameters. The classification accuracy is higher compared to the traditional SVM model. Farid et al. [18] combined naive bayes and decision trees for anomalous traffic detection while eliminating redundant attributes of the traffic data. The proposed algorithm improves the detection rate. Machine learning based classification methods usually require manual feature design and selection, which cannot cope with the evolution of networks nowadays.

ISSN2454-9940
www.ijasem.org
Vol 18, Issue.2 April 2024

Disadvantages

- An existing system is not implemented hybrid deep learning or an efficient ml model detection policy to improve the efficiency and effectiveness of Abnormal Traffic Detection Generation.
- An existing system never used Attention and Big Step Convolutional Neural Network (ABS-CNN) model which is more accurate and efficient.

Proposed System

• In this paper, we propose an Attention and Big Step Convolutional Neural Network (ABS-CNN) model based on the attention mechanism [11]. To solve the problems such as similar features leading to worse classification results, the attention mechanism is invited to assign attention weights to data sequences to distinguish subtle features. To solve the problems such as similar features leading to worse classification results, the attention mechanism is invited to assign attention weights to data sequences to distinguish subtle features. Experiments show that the model with enhanced features has higher classification accuracy and better robustness.

Advantages

- An input layer, three convolutional layers, a fully connected layer and an output layer are set in the ABS-CNN model, and a convolutional attention mechanism is introduced to enhance the ability of convolution to extract traffic features.
- In the proposed system, the ablation study is performed by removing each component in turn from the proposed ABS-CNN and comparing it with the ABS-CNN of the complete pair to verify the impact of each component on the model. To examine the effects of attention mechanism, histogram equalization, and large-step convolution on model performance.

CONCLUSION

To address the difficulties caused by similar features and single model structure on abnormal traffic detection, this paper proposes a detection model based on attention and big-step convolution. Experiments were conducted on both publicly available dataset and real environment crawls dataset. The efficiency

ISSN2454-9940
www.ijasem.org
Vol 18, Issue.2 April 2024

of the model is seen through performance analysis.

• ABS-CNN is the highest in accuracy, precision, recall and F1-Score when compared with traditional models. It is proved that ABS-CNN achieves high accuracy and prediction with good detection effect. And from the confusion matrix of various types of traffic, the classification accuracy of multiple traffic is 100%, which reflects the high sensitivity of ABS-CNN in abnormal traffic detection.

• Compared with different variants of CNN models, ABS-CNN has a shorter training time as well as testing time and runs efficiently. And ABS-CNN shows unparalleled advantages in accuracy, precision, recall and F1-Score with the best classification results.

REFERENCES

[1] O. Salman, I. H. Elhajj, A. Kayssi, and A. Chehab, ‘‘A review on machine learning–based approaches for internet traffic classification,’’ Ann. Telecommun., vol. 75, nos. 11–12, pp. 673–710, Dec. 2020.

[2] A. Madhukar and C. Williamson, ‘‘A longitudinal study of P2P traffic classification,’’ in Proc. 14th IEEE Int. Symp. Modeling, Anal., Simulation, Monterey, CA, USA, Sep. 2006, pp. 179–188, doi: 10.1109/MASCOTS.2006.6.

[3] S. Sen, O. Spatscheck, and D. Wang, ‘‘Accurate, scalable in-network identification of P2P P2P traffic using application signatures,’’ in Proc. 13th Int. Conf. World Wide Web, New York, MY, USA, May 2004, pp. 512–521.

[4] L. Ding, J. Liu, T. Qin, and H. Li, ‘‘Internet traffic classification based on expanding vector of flow,’’ Comput. Netw., vol. 129, pp. 178–192, Dec. 2017.

[5] T. Liu, Y. Sun, and L. Guo, ‘‘Fast and memory-efficient traffic classification with deep packet inspection in CMP architecture,’’ in Proc. IEEE 5th Int. Conf. Netw., Archit., Storage, Macau, China, Jul. 2010, pp. 208–217, doi: 10.1109/NAS.2010.43.

[6] N. Cascarano, L. Ciminiera, and F. Risso, ‘‘Optimizing deep packet inspection for high-speed traffic analysis,’’ J. Netw. Syst. Manage., vol. 19, no. 1, pp. 7–31, Mar. 2011.

[7] G. Aceto, A. Dainotti, W. de Donato, and A. Pescape, ‘‘PortLoad: Taking the

ISSN2454-9940
www.ijasem.org
Vol 18, Issue.2 April 2024

best of two worlds in traffic classification,'' in Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM), San Diego, CA, USA, Mar. 2010, pp. 1–5, doi: 10.1109/INFCOMW.2010.5466645.

[8] L. Vu, C. T. Bui, and Q. U. Nguyen, ''A deep learning based method for handling imbalanced problem in network traffic classification,'' in Proc. 8th Int. Symp. Inf. Commun. Technol., Dec. 2017, pp. 333–339.

[9] P. Wang, F. Ye, X. Chen, and Y. Qian, ''Datanet: Deep learning based encrypted network traffic classification in SDN home gateway,'' IEEE Access, vol. 6, pp. 55380–55391, 2018.

[10] J. H. Shu, J. Jiang, and J. X. Sun, ''Network traffic classification based on deep learning,'' J. Phys., Conf. Ser., vol. 1087, Sep. 2018, Art. no. 062021.

[9] Karne, R. K. ., & Sreeja, T. K. . (2023). PMLC- Predictions of Mobility and Transmission in a Lane-Based Cluster VANET Validated on Machine Learning. International Journal on Recent and Innovation Trends in Computing and Communication, 11(5s), 477–483. https://doi.org/10.17762/ijritcc.v11i5s.7109

[10] Radha Krishna Karne and Dr. T. K. Sreeja (2022), A Novel Approach for Dynamic Stable Clustering in VANET Using Deep Learning (LSTM) Model. IJEER 10(4), 1092-1098. DOI: 10.37391/IJEER.100454.

[11] Reddy, Kallem Niranjan, and Pappu Venkata Yasoda Jayasree. "Low Power Strain and Dimension Aware SRAM Cell Design Using a New Tunnel FET and Domino Independent Logic." International Journal of Intelligent Engineering & Systems 11, no. 4 (2018).

[12] Reddy, K. Niranjan, and P. V. Y. Jayasree. "Design of a Dual Doping Less Double Gate Tfet and Its Material Optimization Analysis on a 6t Sram Cells."

[13] Reddy, K. Niranjan, and P. V. Y. Jayasree. "Low power process, voltage, and temperature (PVT) variations aware improved tunnel FET on 6T SRAM cells." Sustainable Computing: Informatics and Systems 21 (2019): 143-153.

[14] Reddy, K. Niranjan, and P. V. Y. Jayasree. "Survey on improvement of PVT aware variations in tunnel FET on SRAM cells." In 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC), pp. 703-705. IEEE, 2017