Classification of Application Layer Protocols using Deep Learning Methods

Ujjawal Modi - 21074032

Project Supervisor - Dr.Mayank Swarnkar

Department of Computer Science and Engineering
Indian Institute of Technology (BHU)
Varanasi, India

May 3,2023

Table of contents

Introduction

- Network traffic analysis is the process of capturing and analyzing data that is transmitted across a network.
- Network traffic analysis has many applications in various fields, including cybersecurity, application performance monitoring.
- A report by Cisco found that 78% of organizations experienced at least one cyber attack in 2021, with 42% experiencing more than six attacks.
- Network traffic analysis can help organizations to detect and prevent these attacks, reducing the risk of data breaches.
- The increasing use of mobile devices and cloud computing is leading to a more distributed network, making it difficult to secure network traffic.

Why Network Traffic Classification?

- Network Traffic classification is the process of categorizing network traffic into different groups.
- Important for efficient resource utilization, network security, and quality of service for customers.
- Enables network managers to monitor and manage network resources efficiently.
- Helps identify potential security issues like malware or hacking attempts, reducing the risk of data loss or harm.
- Prioritizing traffic based on its importance or sensitivity can improve the quality of service for customers and users.

What is Flow?

- A flow in networking refers to a sequence of packets that share common characteristics.
- DPI (Deep Packet Inspection) and SPI (Shallow Packet Inspection) are two different techniques for classifying network traffic.
- SPI inspects only the header information of packets, such as the source and destination IP addresses, protocol.
- DPI inspects and analyzes the contents of network packets at the application layer.
- Packet level DPI and Flow level DPI are two different approaches to DPI used in network traffic analysis.
- Packet level DPI involves examining the contents of each individual packet for classification.

What is Flow?

- Flow level DPI analyzes the behavior of each flow to determine for classification, rather than examining each individual packet.
- The transport layer is responsible for managing the transfer of packets over a network.
- The two most commonly used transport layer protocols are the Transmission Control Protocol (TCP) and the User Datagram Protocol (UDP).

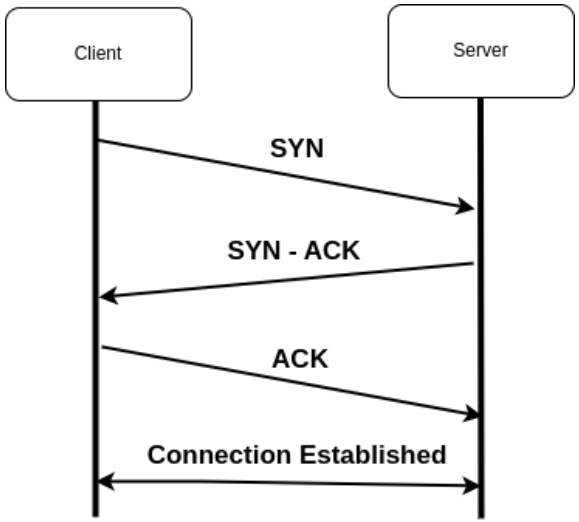

Figure 1: Establishment of a TCP flow.

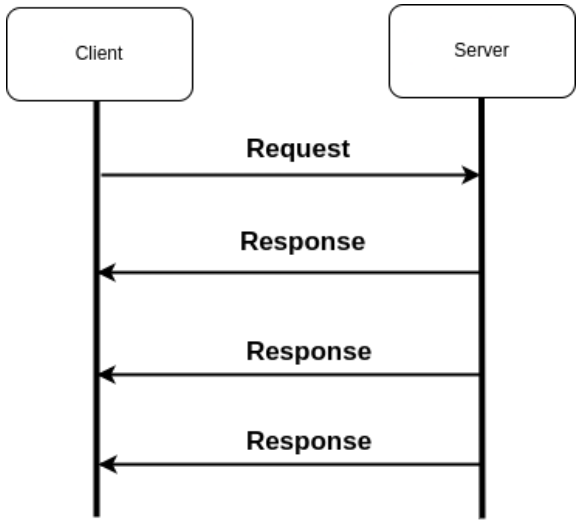

Figure 2: Establishment of a UDP flow.

Problem Statement

- Problem: Classification of network traffic on the basis of protocols.
- The proposed solution involves classifying network protocol using the first 32 bits of the payload
- LSTM, RNN, CNN, ANN, and GRU are used to classify the traffic
- improvement in protocol detection's precision and effectiveness

Literature Survey

Author et al	Ref No.	Work Done	Shortcomings
Zhaoyang Zhang, Yancheng Shen, Weiwei Hu	1	Uses a novel feature learning approach based on discriminative autoencoder (DAE) and support vector machine (SVM) for network traffic classification.	1. Relatively small dataset used for evaluation and the use of a single machine learning classifier. 2. Effectively 512 bytes of each packet are used ans total 32 packets are there.
Ons Aouedi, Kandaraj Piamrat, Dhruvjyoti Bagadthey	2	Uses a semi-supervised stacked autoencoder (SSAE) for network traffic classification that requires fewer labeled samples	1. The paper lacks sufficient experimental evaluation on large and complex datasets. 2.semi-supervised approach is not effective in scenarios where labeled data is scarce.
S. Jangir, S. Laxmi, M. Obaidat	3	Uses a keyword matching approach for deep packet inspection (DPI) based network traffic classification	1. the evaluation is unclear how it will perform in different datasets. 2.It assumes that the set of keywords is static and does not change over time.
Q. Liu, J. Liu, K. Lin	4	Uses a multitask learning (MTL) approach for network traffic classification that can handle multiple traffic classes.	1. Limited number of datasets were used in this method and were very specific.2.Complexity of the deep neural network architecture is high.

Literature Survey

Author et al	Ref No.	Work Done	Shortcomings
M. Ahmed, A. Bounceur, M. Ahmed	5	Uses deep convolutional recurrent autoencoder neural networks (DCRNNs) for spatial-temporal features extraction for network traffic classification	1. computationally expensive and difficult to train on large-scale datasets. 2. The paper does not compare the proposed method to state-of-the-art approaches.
F. Wang, L. Zheng, Y. Liu	6	Uses deep learning techniques to predict network flow characteristics such as packet size, interarrival time, and flow duration.	1. The paper focuses on predicting flow characteristics rather than traffic classification. 2. requires a large amount of labeled data to achieve good performance.
Y. Zhang, C. Li, Y. Wang	7	Uses a self-attention based deep learning method called SAM for online network traffic classification.	1.The SAM approach may not perform well with imbalanced data. 2.The computational complexity of the SAM approach is high.
Y. Chen, J. Hu, Y. Zhou	8	Uses a tree structural recurrent neural network (Tree-RNN) for network traffic classification.	building a tree structure can be computationally difficult. 2.tree structure used in the approach may be difficult to interpret and explain

Literature Survey

Author et al	Ref No.	Work Done	Shortcomings
C. Wang, Z. Wei, C. Liu	9	Uses byte-label joint attention learning (BL-JAL) for packet-grained network traffic classification.	1.The performance improvement of the proposed method over existing methods is not significant. 2.The paper does not compare the proposed method to state-of-the-art approaches.
C. Luo, X. Wu, Q. Zhang	10	Uses a hybrid deep learning method that combines convolutional neural networks (CNNs) and long short-term memory (LSTM) networks for network traffic classification.	1. focuses on using only two types of deep learning models (CNN and LSTM) and doesn't explore other models. 2.requires significant computational resources to train and evaluate the models.
N. Rao, G. Huang, L. Liao	11	Proposes an efficient feature selection model for network traffic classification that can reduce the number of features required for classification.	1.compares only its proposed feature selection method with one other method.2.The performance is only evaluated on a single dataset, which may not be representative of all possible network traffic scenarios.
S. Shaik, S. Laxmi	12	Use OF a multi-class Support Vector Machine (SVM) algorithm with an active learning approach for network traffic classification.	1.evaluated only on a single dataset, which may limit the generalizability of the results. 2.does not provide a comparison with state-of-the-art methods for network traffic classification.

Motivation

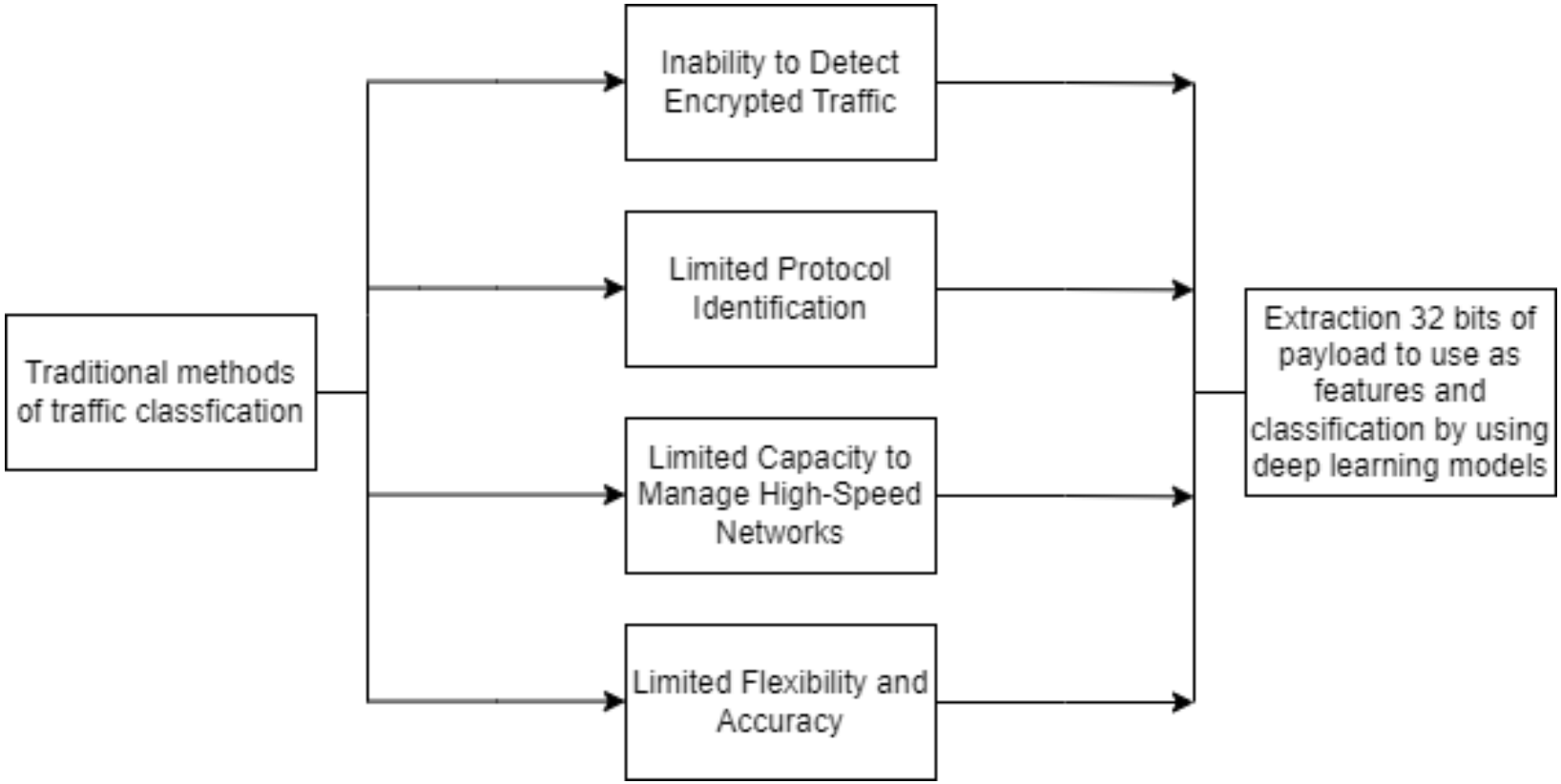

Figure 3: Flowchart representing previous shortcomings.

Work Done

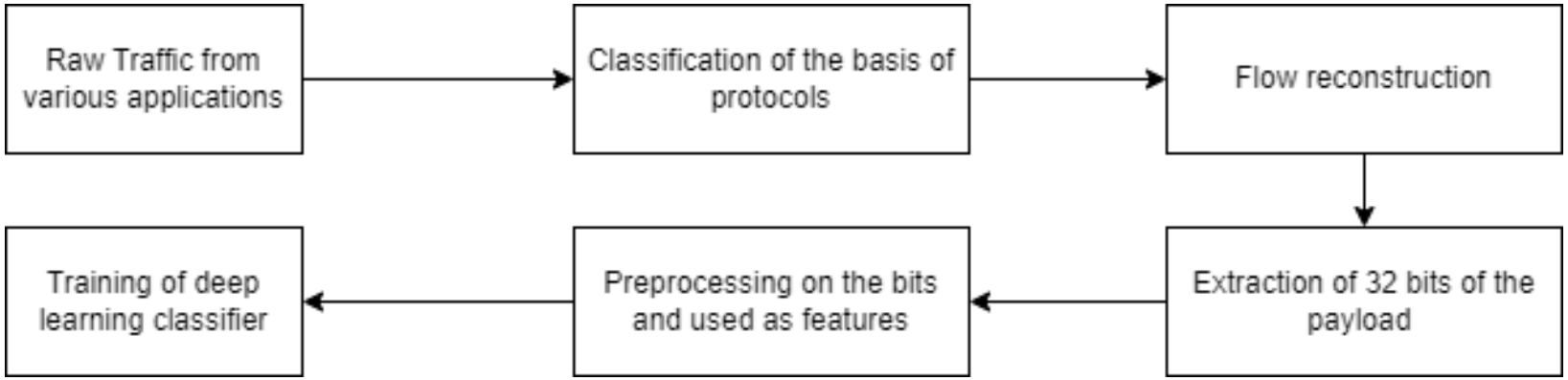

Figure 4: Architecture of the training module.

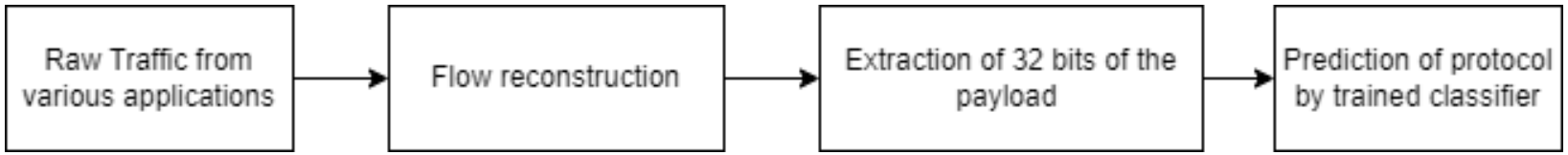

Figure 5: Architecture of the testing module.

Work Done

- Network traffic data was given in form of pcap files.
- Flow reconstruction involves reconstructing network flows from the network traffic.
- This was done by grouping packets based on source and destination IP addresses and port numbers.
- Extracting the payload involves capturing the payload data of each packet in the reconstructed flows.
- In this project, we extracted the first 32 bits of the payload as features for the deep learning models.
- After extracting the payload,we prepare the dataset by labeling each flow with the corresponding network protocol.

Work Done

- With the dataset prepared, we trained deep learning models to predict network protocols based on the extracted payload features.
- We trained 5 deep learning models to predict network protocols based on the extracted payload features.
  - CNN
  - RNN
  - LSTM
  - GRU
  - ANN
- Once the models are trained, we evaluated their performance by testing them on a separate dataset.
- Finally we calculated recall rate of each class to determine the effectiveness of the models.

Dataset Description

Protocol	TCP/UDP	Flows for Training	Size of Training (MB)	Flows for Testing	Size of Testing (MB)
BACnet	UDP	9	0.009	11	0.004
BJNP	UDP	34	0.0025	38	0.003
Bootp	UDP	177	0.714	172	4.4
DNS	UDP	59015	19.5	59605	18.1
Kerberos	UDP	670	1.6	673	1.9
NBNS	UDP	1272	1.6	1271	8.7
NBSS	TCP	366	2.7	364	3.9
NTP	UDP	403	1.8	401	0.78
QUIC	UDP	126	0.014	92	0.01
SSH	TCP	1132	6.2	1134	6.2

Table 1: Description of dataset used for training and evaluating model performance.

CNN Results

In each bracket count of testing flows is mentioned. Example if BACnet has 11 testing flows and they are 100 % matched with BACnet than 11 was written in BACnet and 0 in remaining cells. If BJNP has 15 flows and out of them 2 testing flows are mismatched with DNS then 13 was wriiten in (BJNP to BJNP) cell and 2 in (BJNP to DNS) cell.

	BACnet (11)	BJNP (38)	Bootp (172)	DNS (59605)	Kerberos (673)	NBNS (1271)	NBSS (364)	NTP (401)	QUIC (92)	SSH (1134)
BACnet	11	0	0	0	0	0	0	0	0	0
BJNP	0	38	0	0	0	0	0	0	0	0
Bootp	0	0	172	0	0	0	0	0	0	0
DNS	0	0	0	59605	0	8	0	0	6	0
Kerberos	0	0	0	0	670	1	2	0	0	0
NBNS	0	0	0	0	0	1262	0	0	0	0
NBSS	0	0	0	0	3	0	362	0	0	0
NTP	0	0	0	0	0	0	0	401	0	0
QUIC	0	0	0	0	0	0	0	0	86	0
SSH	0	0	0	0	0	0	0	0	0	1134

Table 2: Confusion Matrix

Protocol	Recall Rate(%)
BACnet	100
BJNP	100
Bootp	100
DNS	100
Kerberos	99.55
NBNS	99.29
NBSS	99.45
NTP	100
QUIC	93.47
SSH	100

Table 3: Recall rates (%).

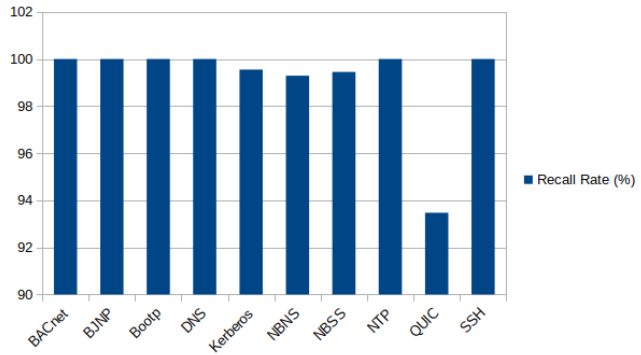

Figure 6: Plot for recall rates (%).

RNN Results

In each bracket count of testing flows is mentioned. Example if BACnet has 11 testing flows and they are 100 % matched with BACnet than 11 was written in BACnet and 0 in remaining cells. If BJNP has 15 flows and out of them 2 testing flows are mismatched with DNS then 13 was wriiten in (BJNP to BJNP) cell and 2 in (BJNP to DNS) cell.

	BACnet (11)	BJNP (38)	Bootp (172)	DNS (59605)	Kerberos (673)	NBNS (1271)	NBSS (364)	NTP (401)	QUIC (92)	SSH (1134)
BACnet	5	0	0	0	0	0	0	0	0	0
BJNP	0	38	0	0	0	0	0	0	1	0
Bootp	0	0	172	0	0	0	0	0	0	0
DNS	0	0	0	59602	14	8	0	0	5	0
Kerberos	0	0	0	0	364	0	0	0	1	0
NBNS	6	0	0	3	1	1263	0	0	4	0
NBSS	0	0	0	0	4	0	359	0	1	0
NTP	0	0	0	0	207	0	3	401	12	0
QUIC	0	0	0	0	83	0	2	0	68	0
SSH	0	0	0	0	0	0	0	0	0	1134

Table 4: Confusion Matrix

Protocol	Recall Rate(%)
BACnet	45.45
BJNP	100
Bootp	100
DNS	99.99
Kerberos	54.08
NBNS	99.37
NBSS	98.62
NTP	100
QUIC	73.91
SSH	100

Table 5: Recall rates (%).

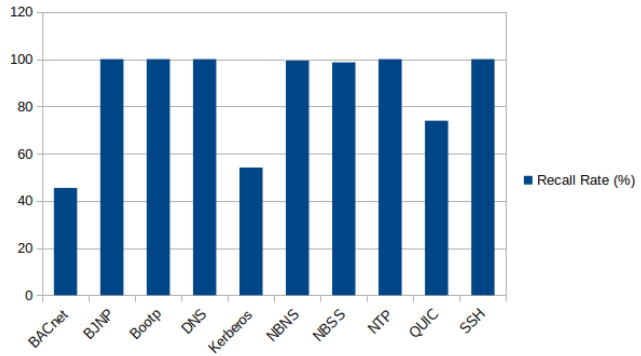

Figure 7: Plot for recall rates (%).

LSTM Results

In each bracket count of testing flows is mentioned. Example if BACnet has 11 testing flows and they are 100 % matched with BACnet than 11 was written in BACnet and 0 in remaining cells. If BJNP has 15 flows and out of them 2 testing flows are mismatched with DNS then 13 was wriiten in (BJNP to BJNP) cell and 2 in (BJNP to DNS) cell.

	BACnet (11)	BJNP (38)	Bootp (172)	DNS (59605)	Kerberos (673)	NBNS (1271)	NBSS (364)	NTP (401)	QUIC (92)	SSH (1134)
BACnet	11	0	0	0	0	0	0	0	0	0
BJNP	0	38	0	0	0	5	0	0	0	0
Bootp	0	0	172	0	0	0	0	0	0	0
DNS	0	0	0	59601	0	29	0	0	10	0
Kerberos	0	0	0	4	668	0	3	0	9	0
NBNS	0	0	0	0	0	1233	0	0	1	0
NBSS	0	0	0	0	5	3	360	0	0	0
NTP	0	0	0	0	0	0	1	401	1	0
QUIC	0	0	0	0	0	1	0	0	71	0
SSH	0	0	0	0	0	0	0	0	0	1134

Table 6: Confusion Matrix.

Protocol	Recall Rate(%)
BACnet	100
BJNP	100
Bootp	100
DNS	99.99
Kerberos	99.25
NBNS	97.01
NBSS	98.9
NTP	100
QUIC	77.17
SSH	100

Table 7: Recall rates (%).

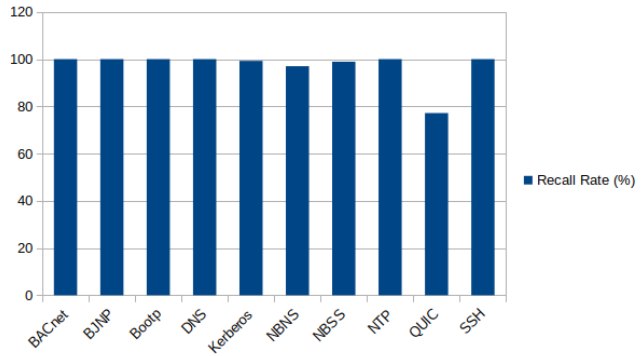

Figure 8: Plot for recall rates (%).

GRU Results

In each bracket count of testing flows is mentioned. Example if BACnet has 11 testing flows and they are 100 % matched with BACnet than 11 was written in BACnet and 0 in remaining cells. If BJNP has 15 flows and out of them 2 testing flows are mismatched with DNS then 13 was wriiten in (BJNP to BJNP) cell and 2 in (BJNP to DNS) cell.

	BACnet (11)	BJNP (38)	Bootp (172)	DNS (59605)	Kerberos (673)	NBNS (1271)	NBSS (364)	NTP (401)	QUIC (92)	SSH (1134)
BACnet	11	0	0	0	0	0	0	0	0	0
BJNP	0	38	0	0	0	2	0	0	0	0
Bootp	0	0	172	0	0	0	0	0	0	0
DNS	0	0	0	59602	0	21	1	0	3	0
Kerberos	0	0	0	1	670	7	3	0	6	0
NBNS	0	0	0	1	0	1240	0	0	4	0
NBSS	0	0	0	0	3	1	360	0	0	0
NTP	0	0	0	0	0	0	0	401	0	0
QUIC	0	0	0	1	0	0	0	0	79	0
SSH	0	0	0	0	0	0	0	0	0	1134

Table 8: Confusion Matrix.

Protocol	Recall Rate(%)
BACnet	100
BJNP	100
Bootp	100
DNS	99.99
Kerberos	99.55
NBNS	97.56
NBSS	98.9
NTP	100
QUIC	85.86
SSH	100

Table 9: Recall rates (%).

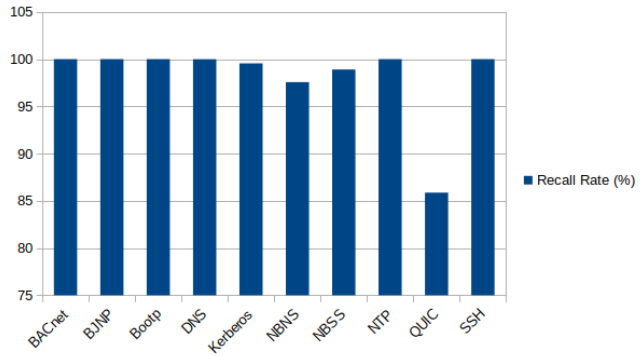

Figure 9: Plot for recall rates (%).

ANN Results

In each bracket count of testing flows is mentioned. Example if BACnet has 11 testing flows and they are 100 % matched with BACnet than 11 was written in BACnet and 0 in remaining cells. If BJNP has 15 flows and out of them 2 testing flows are mismatched with DNS then 13 was wriiten in (BJNP to BJNP) cell and 2 in (BJNP to DNS) cell.

	BACnet (11)	BJNP (38)	Bootp (172)	DNS (59605)	Kerberos (673)	NBNS (1271)	NBSS (364)	NTP (401)	QUIC (92)	SSH (1134)
BACnet	11	0	0	0	0	0	0	0	0	0
BJNP	0	38	0	0	0	0	0	0	0	0
Bootp	0	0	172	0	0	0	0	0	0	0
DNS	0	0	0	59605	0	12	0	0	0	0
Kerberos	0	0	0	0	672	0	2	0	0	0
NBNS	0	0	0	0	0	1259	0	0	0	0
NBSS	0	0	0	0	1	0	362	0	0	0
NTP	0	0	0	0	0	0	0	401	1	0
QUIC	0	0	0	0	0	0	0	0	91	0
SSH	0	0	0	0	0	0	0	0	0	1134

Table 10: Confusion Matrix.

Protocol	Recall Rate(%)
BACnet	100
BJNP	100
Bootp	100
DNS	100
Kerberos	99.85
NBNS	99.05
NBSS	99.45
NTP	100
QUIC	98.91
SSH	100

Table 11: Recall rates (%).

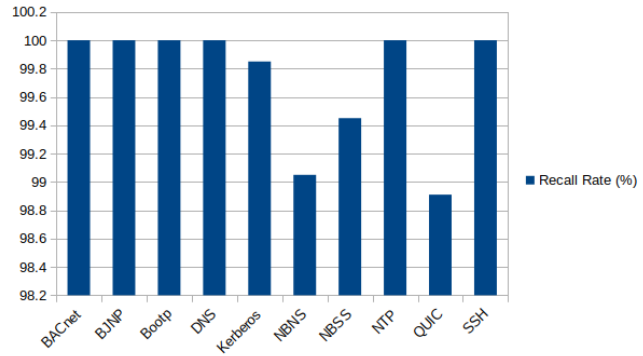

Figure 10: Plot for recall rates (%).

Conclusion

- The development of accurate and reliable protocol classification systems is a crucial area of research that has many practical applications.
- This project aimed to investigate the use of deep learning methods for network traffic classification.
- The results of the project showed that these models were able to achieve high levels of accuracy in identifying various network protocols, including NBNS, SSH, and DNS.
- As network technology continues to evolve, the importance of effective network traffic classification will increase.
- We hope this project has provided some valuable insights into the methods and approaches that can be used to classify network traffic efficiently.

References

[1] Zhao, L., Cai, L., Yu, A., Xu, Z. and Meng, D., 2020, March. A novel network traffic classification approach via discriminative feature learning. In Proceedings of the 35th annual ACM symposium on applied computing (pp. 1026-1033).

[2] Khandait, P., Hubballi, N. and Mazumdar, B., 2020, January. Efficient keyword matching for deep packet inspection based network traffic classification. In 2020 International Conference on COMmunication Systems NETworkS (COMSNETS) (pp. 567-570). IEEE.

[3] Rezaei, S. and Liu, X., 2020, August. Multitask learning for network traffic classification. In 2020 29th International Conference on Computer Communications and Networks (ICCCN) (pp. 1-9). IEEE.

[4] D'Angelo, G. and Palmieri, F., 2021. Network traffic classification using deep convolutional recurrent autoencoder neural networks for spatial–temporal features extraction. Journal of Network and Computer Applications, 173, p.102890.

[5] Xie, G., Li, Q., Jiang, Y., Dai, T., Shen, G., Li, R., Sinnott, R. and Xia, S., 2020, August. Sam: Self-attention based deep learning method for online traffic classification. In Proceedings of the Workshop on Network Meets AI ML (pp. 14-20).

[6] Ren, X., Gu, H. and Wei, W., 2021. Tree-RNN: Tree structural recurrent neural network for network traffic classification. Expert Systems with Applications, 167, p.114363.

References

[7] Mao, K., Xiao, X., Hu, G., Luo, X., Zhang, B. and Xia, S., 2021, June. Byte-Label Joint Attention Learning for Packet-grained Network Traffic Classification. In 2021 IEEE/ACM 29th International Symposium on Quality of Service (IWQOS) (pp. 1-10). IEEE.

[8] Sarhangian, F., Kashef, R. and Jaseemuddin, M., 2021, April. Efficient traffic classification using hybrid deep learning. In 2021 IEEE International Systems Conference (SysCon) (pp. 1-8). IEEE.

[9] Alam, F., Kashef, R. and Jaseemuddin, M., 2021, April. Enhancing The Performance of Network Traffic Classification Methods Using Efficient Feature Selection Models. In 2021 IEEE International Systems Conference (SysCon) (pp. 1-6). IEEE.

[10] Jin, Z., Liang, Z., Wang, Y. and Meng, W., 2021. Mobile network traffic pattern classification with incomplete a priori information. Computer Communications, 166, pp.262-270.

[11] Dong, S., 2021. Multi class SVM algorithm with active learning for network traffic classification. Expert Systems with Applications, 176, p.114885.

[12] Eom, W.J., Song, Y.J., Park, C.H., Kim, J.K., Kim, G.H. and Cho, Y.Z., 2021, April. Network traffic classification using ensemble learning in software-defined networks. In 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) (pp. 089-092). IEEE.

References

[13] Aouedi, O., Piamrat, K. and Parrein, B., 2021, June. Performance evaluation of feature selection and tree-based algorithms for traffic classification. In 2021 IEEE International Conference on Communications Workshops (ICC Workshops) (pp. 1-6). IEEE.

[14] Wei, W., Gu, H., Deng, W., Xiao, Z. and Ren, X., 2022. ABL-TC: A lightweight design for network traffic classification empowered by deep learning. Neurocomputing, 489, pp.333-344.

Thank You