Entropy-based Image Preprocessing Method for Encrypted Network Traffic

Haeseong Cha[1], Seungyong Park[1], Byoungjin Seok[1], Erzhena Tcydenova[1], Jungsuk Song[2]*, and Kiwook Sohn[1]

[1] Seoul National University of Science and Technology, Seoul, Republic of Korea
{haeseongcha, sypark25, bjseok, etcydenova, kiwook}@seoultech.ac.kr
[2] Korea Institute of Science and Technology Information, Republic of Korea
song@kisti.re.kr

Abstract

Given the rise of cyberattacks exploiting encrypted traffic, traditional classification methods lack the capability to effectively detect such threats. This study introduces a preprocessing technique converting encrypted traffic into RGB images using entropy differences between encrypted and unencrypted data. This offers a richer representation than grayscale methods. Using Resnet-9, we achieved a 98.04% accuracy, a roughly 1%p improvement over prior techniques. Further research will focus on consistency and wider application.

Keywords: Encrypted Network Traffic, Deep Packet, Deep Learning, Classification, Preprocessing, Entropy

1 Introduction

In the current landscape of cyberattacks, encrypted traffic is increasingly being used to hide malicious activities, challenging traditional detection methods [1]. This study aims to address this by introducing a novel preprocessing technique that transforms packet data into RGB images, leveraging both packet byte values and entropy information. This approach promises a richer representation compared to conventional grayscale methods and seeks to enhance malware detection performance when combined with computer vision models.

2 Proposed Method

Encrypted packets predominantly consist of a header and a payload, distinguished by differences in entropy. The header, which contains basic packet information, exhibits relatively lower entropy. In contrast, due to encryption, the payload, encapsulating the main content of the packet, has higher entropy. This research proposes a novel preprocessing method that leverages this entropy difference to transform encrypted network traffic into RGB images. This method is visually represented in Figure 1. The packet image preprocessing method proposed in this poster is divided into two stages: packet normalization and the generation of RGB images enriched with entropy information. The Packet Normalization step ensures that packets adopt a standardized format and that specific packet details don't skew the classification results. This process comprises the Ethernet Header Removal,the Mask IP, the UDP Header Padding,and the Concatenate Header with Payload. The Generation of RGB Images with Entropy Information

The 7th International Conference on Mobile Internet Security (MobiSec'23), Dec. 19-21, 2023, Okinawa, JAPAN, Article No.P-31
*Corresponding author: Korea Institute of Science and Technology Information, Seoul, 02456, Republic of Korea, Email: song@kisti.re.kr

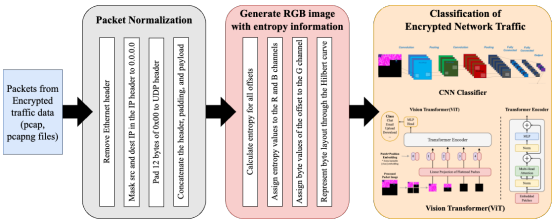

Figure 1: Design of Entropy-based encrypted network traffic preprocessing to image method

stage distinguishes between the header and the payload more clearly and enhances computer vision models' ability to glean additional insights from the packet. This step includes the Calculate Entropy, Assign Entropy Value, the Assign Byte Value, and the Byte Layout to Hilbert Curve.

3 Experimental Results

In our experiments, we compared the preprocessing methods of [2] and [3], which generate grayscale images to classify encrypted network traffic, with the preprocessing technique proposed in this study. These were evaluated using the Resnet-9 model, and the results are presented in Table 1.

	Accuracy	F1-score
[2]	0.93	0.93
[3]	0.9698	0.9004
This Study	**0.9804**	**0.9655**

Table 1: Result of Experiment

4 Conclusion

In this study, we analyzed network traffic at the packet level, achieving an accuracy of 98.04%, an improvement over existing works. Diverging from traditional grayscale imaging, we visualized packet and entropy information as RGB images. This approach enhanced our deep learning model's capability to distinguish between encrypted and unencrypted traffic. Future works will aim to further validate this performance and explore its broader application.

References

[1] Cisco. Enterprise Network Security - Encrypted Traffic Analytics (ETA). https://www.cisco.com/c/en/us/solutions/enterprise-networks/enterprise-network-security/eta.html, last viewed 2023-10-20.

[2] Mohammad Lotfollahi, Mahdi Jafari Siavoshani, Ramin Shirali Hossein Zade, and Mohammdsadegh Saberian. Deep packet: a novel approach for encrypted traffic classification using deep learning. *Soft Computing*, 24(3):1999–2012, February 2020.

[3] Hyun-Kyo Lim, Ju-Bong Kim, Joo-Seong Heo, Kwihoon Kim, Yong-Geun Hong, and Youn-Hee Han. Packet-based Network Traffic Classification Using Deep Learning. In *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, pages 046–051, February 2019.