Identification of HTTP Traffic Injection for Free-Riding in Zero-Rating Data Plans

Carlos Gómez Chico, Luis de Pedro, Daniel Perdices, Jorge E. López de Vergara
Universidad Autónoma de Madrid
Madrid, Spain
carlos.gomezchico@estudiante.uam.es, {luis.depedro, daniel.perdices, jorge.lopez_vergara}@uam.es

***Abstract*—Zero-rating data plans have spurred the emergence of HTTP traffic injection apps, posing challenges for telecommunication operators offering data plan services. These apps exploit encryption to gain free Internet access, known as free-riding, which complicates detection. To combat this, network traffic classification techniques are employed to identify illegitimate traffic patterns, using Machine Learning algorithms. This paper employs heuristic methods to uncover such patterns and trains supervised algorithms such as Support Vector Machine for entropy-based free-riding detection. By analyzing characteristics from simulated zero-rating use, deeper insights into traffic generation and identification are obtained.**

I. Introduction

The rapid expansion of connectivity in Telecommunication Engineering has led to a surge in technologies improving user experience, but also introducing vulnerabilities that allow fraudulent practices. Notably, HTTP injection has emerged as a potent VPN technique, enhancing network security and privacy. However, it has been exploited for fraudulent activities such as bypassing security controls and evading Internet connection billing, known as free-riding. This study aims to tackle this issue by proposing a comprehensive approach to detect fraudulent activities via HTTP injection in metered traffic networks.

II. Problem Analysis

Internet Service Providers (ISPs) face intense competition in the industry and strive for growth by attracting subscribers from competitors' networks. One effective way to differentiate is by offering zero-rating plans [1], which provides unlimited access to certain applications without consuming data allowance, typically for mobile but also for fixed networks. The problem arises from the use of HTTP traffic injection applications and their detection, posing a significant financial challenge for ISPs, due to revenue loss from fraudulent data usage [2].

ISPs typically charge for network traffic, especially in some countries where unlimited data plans are not available. In some cases, disaggregated data plans are offered, where only certain applications have limited or unlimited data, separate from the general plan. These plans are commonly referred to as zero-rating services. By tunneling, users are able to deceive operators into believing they are browsing zero-rated applications when they are fraudulently using other services free of charge. This type of attack is known as free-riding, as depicted in fig. 1, which illustrates the problem addressed in this study. In this example, a user establishes a TLS tunnel to hide the charged traffic (e.g., `facebook.com`), indicating a forged TLS SNI of a domain included in the zero-rating service (e.g. `spotify.com`). In this way, the network operator will identify this traffic as zero-rated, and it will not charge it. To make this fraudulent use more difficult to identify, different SNI names of the zero-rated service (e.g., {`www`, `open`, `api`...}`.spotify.com`) can be provided in different connections to the TLS tunnel.

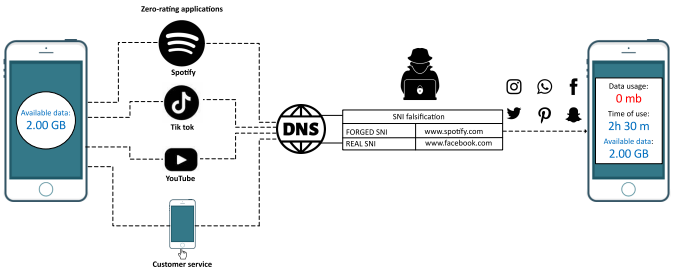

Fig. 1. Free-riding attack by injecting HTTP traffic at a zero-rating rate.

III. Anomaly Detection in Network Traffic

A. Entropy-based Anomaly Detection

Entropy-based anomaly detection techniques in network traffic are appealing, due to their simplicity and applicability in current encrypted network environments. It is important to note that entropy-based anomaly detection can be weakened due to susceptibility to deception through the addition of false data that mask the anomaly. Techniques for anomaly classification typically rely on machine learning, implying additional complexity, which may not be necessary depending on the problem being addressed [3].

In traffic anomaly detection techniques, entropy is used to represent the level of randomness in a data distribution. Changes in the data structure in a distribution obtained from the aggregation process will modify the entropy value. If the change in entropy is significant, it is considered unusual behavior in network communication, often indicating security

This research has been partially funded by the Spanish State Research Agency under the project AgileMon (AEI PID2019-104451RB-C21).

978-3-903176-64-5

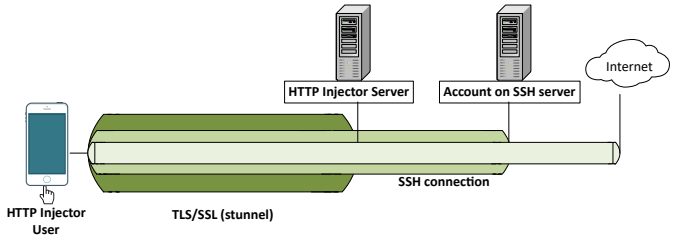

Fig. 2. HTTP Injector connection diagram: SSL/TLS→SSH

threats. In many cases, the well-known Shannon entropy (1) is used:

$$H(X) = -\sum_{i} p(x_i) \cdot \log_2(p(x_i)). \tag{1}$$

Some authors claim that parameterized Rényi (2) and Tsallis (3) entropies outperform Shannon entropy in terms of better detection of peaks or tails in feature distributions [4], [5]. These conclusions are closely related to the detection methods, the data, and the features of the experiments. Consequently, this conclusion cannot be simply generalized. This study investigates and compares identification results using both methods.

$$H_\alpha(X) = \frac{1}{1-\alpha} \log_b \left(\sum_{i=1}^{N} p(x_i)^\alpha \right). \tag{2}$$

$$H_q(X) = \frac{1}{q-1} \left(1 - \sum_{i=1}^{N} p(x_i)^q \right), \text{ where } q \geq 0 \text{ and } q \neq 1. \tag{3}$$

B. Detection through Machine Learning

The approach for detection depends on the dataset you are working with and the objective you want to achieve. Despite the advent of Deep Learning, we employed a Support Vector Machine (SVM) classifier due to its superior trade-off between computational cost and performance, as network operators need to perform this classification for millions of users. In any case, it is important to avoid producing biases due to data leakage in training data when applying machine learning.

IV. Solution Design and Development

HTTP Injector, an Android application, presents itself as a professional VPN application, enabling secure and private Internet browsing [6]. This app can be employed to tunnel an SSH connection over an HTTPS flow, using an arbitrary SNI domain (or a set of domains) to access any destination from the SSH server that acts as a proxy. Fig. 2 depicts this configuration. The user connects to the HTTP Injector Server using a SNI that fools the zero-rating data plan identification of the network operator so that the connection is then tunneled to the SSH server, where a SOCKS proxy allows you to access another destination on the Internet.

Throughout the development of this work, we considered *fraudulent traffic* the emulation of a free-riding attacks, as described above. We consider *legitimate traffic* when actually accessing a concrete zero-rating service.

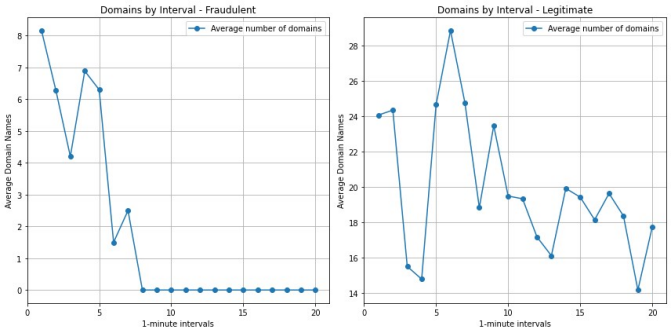

Fig. 3. Domains per 1-minute time interval for fraudulent (left) and legitimate (right) traffic.

To evaluate our solution, 50 captures of each traffic type were obtained. Each one consists of 20 minutes of traffic, resulting in a total of 1,000 minutes of traffic for each type. The generated traffic is diverse, allowing models to generalize and avoid overfitting. The dataset is available at GitHub.[1]

V. Integration, Testing and Results

In this section, we use our heuristic calculations for the feature selection and training process.

A. Domain analysis by time interval

The number of domains per time interval was examined to verify if this feature followed any traffic pattern and, therefore, if it could be subsequently used as a feature. For this purpose, a time interval of 1 minute was chosen, and the number of seen domains in the SNI TLS field throughout the entire network traffic capture was studied. Fig. 3 shows the domains seen in fraudulent (left) and legitimate (right) traffic patterns. Although they behave different, sometimes it is difficult to distinguish them based on this number. Thus, we have abandoned this approach, and focused on the set of IP addresses seen in both scenarios, as shown below.

B. Optimal parameter search for Rényi and Tsallis entropies

The average entropies of source and destination IP addresses in the user packets, and their combination, are calculated for each of the legitimate and fraudulent datasets (see fig. 4 and fig. 5 for Rényi and Tsallis entropies, respectively). These values are plotted to study at which optimal values of each entropy there is no overlap. The standard deviation is plotted to account for the variability. The optimal value for Rényi is $\alpha = 0.25$ and Tsallis is $q = 0.5$.

C. SVMs for Shannon, Rényi and Tsallis entropies

Table I shows the results when using Shannon entropy for classification. This SVM model demonstrates a solid performance, with execution times of just 1.09 milliseconds. Table II gives the classification results for Rényi and Tsallis entropies. This SVM model is highly effective in classification, with

[1] https://github.com/CarlosGChico/TFM

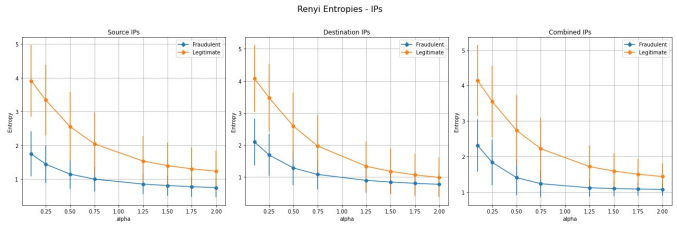

Fig. 4. Variation of the Rényi entropy for packet IP addresses according to the parameter α, for source (left), destination (center) and combined (right).

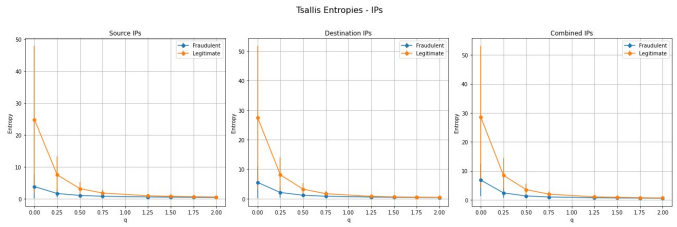

Fig. 5. Variation of the Tsallis entropy for packet IP addresses according to the parameter q, for source (left), destination (center) and combined (right).

execution times of just 10.59 milliseconds. The time to obtain the classification is very important for network operators, as they have millions of users to analyze every minute to tag their traffic as fraudulent or legitimate.

D. Detection of fraudulent traffic based on supervised algorithms for entropy-based network traffic

From the test dataset, a random 20-minute user connection is extracted, divided into 1-minute intervals, to observe how the algorithms detect traffic during a connection in near real-time. The objective is to determine how many attempts in 1 minute intervals the operator would need to ensure that the detected traffic is fraudulent.

Figs. 6 and 7 show the results of this temporal classification for different entropies. As shown, when using the second approach, less false positives (legitimate traffic classified as fraudulent) and negatives (fraudulent traffic classified as legitimate) are obtained in the studied timeframe.

VI. Conclusion

The main aim of this work has been to find models to detect fraudulent HTTP traffic injections. Using entropies of packet IP addresses per each user over time intervals was considered as a simpler and more manageable option, suggesting multiple traffic patterns for operators to utilize in detection. We have confirmed that SVM algorithms using traffic entropies are effective and easy to implement, and they show better performance for Rényi and Tsallis entropies. We also tested other approach using enriched network flows [7] for classification, but it did not work as expected. While flows are intriguing, other classification methods may be more effective in this scope. In summary, supervised algorithms based on traffic entropies are useful for classifying HTTP injection traffic and could be implemented in real-time, studying different time intervals to find which is optimal for fraudulent usage detection.

TABLE I
Classification Report - SVM Model for Shannon Entropies with Optimal Parameters

	Precision	Sensitivity	f1-score	Support
Legitimate Traffic	0.91	0.87	0.89	210
Fraudulent Traffic	0.87	0.92	0.90	210
Accuracy	0.89	420		
Macro avg	0.89	0.89	0.89	420
Weighted avg	0.89	0.89	0.89	420

TABLE II
Classification Report - SVM Model for Rényi and Tsallis Entropies with Optimal Parameters

	Precision	Recall	f1-score	Support
Legitimate Traffic	0.93	0.94	0.94	210
Fraudulent Traffic	0.94	0.93	0.94	210
Precision	0.94	420		
Macro avg	0.94	0.94	0.94	420
Weighted avg	0.94	0.94	0.94	420

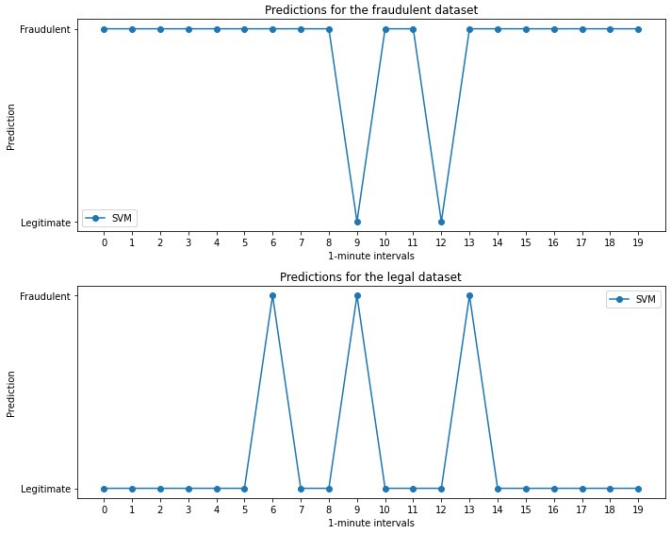

Fig. 6. 1-minute interval Shannon-based SVM model predictions for fraudulent (above) and legitimate (below) traffic.

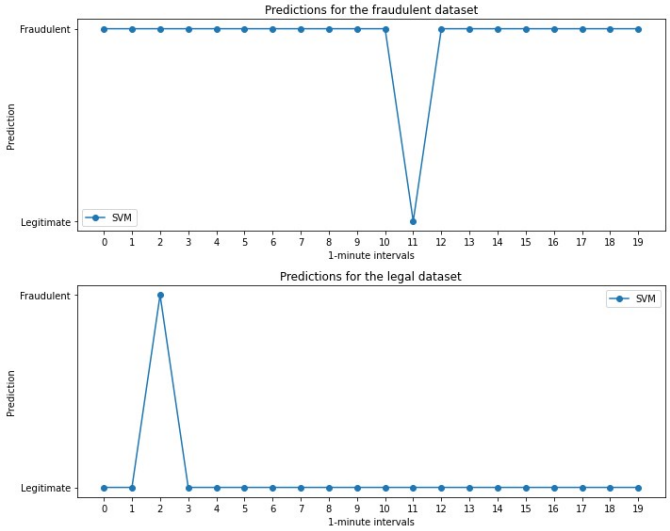

Fig. 7. Predictions of the SVM model based on Rényi and Tsallis per 1-minute interval for fraudulent (above) and legitimate (below) traffic.

REFERENCES

[1] "Sandvine: Global Internet Phenomena Spotlight: Zero-Rating Fraud," 2017. [Online]. Available: https://www.sandvine.com/hubfs/Sandvine_Redesign_2019/Downloads/Internet%20Phenomena/sandvine-spotlight-zero-rating-fraud.pdf. [Last accessed: 02 10 2024].

[2] M. Di Martino, P. Quax, and W. Lamotte, "Knocking on IPs: Identifying HTTPS Websites for Zero-Rated Traffic," Hindawi, p. 14, 2020.

[3] R. D. Triviño, A. A. Franco, and R. L. Ochoa, "Zero Rating Effects in South American Countries," in *2023 Ninth International Conference on eDemocracy & eGovernment (ICEDEG)*, Quito, Ecuador, 2023, 2023.

[4] J. Ibrahim and S. Gajin, "Entropy-based network traffic anomaly classification method resilient to deception," *Computer Science and Information Systems*, pp. 87-116, 2019.

[5] P. Bereziński, B. Jasiul, and M. Szpyrka, "An Entropy-Based Network Anomaly Detection Method," *Entropy*, vol. 17, no. 4, pp. 2367-2048, 2015.

[6] Evozi, "HTTP Injector (SSH/V2R/DNS) VPN," Evozi, [Online]. Available: https://play.google.com/store/apps/details?id=com.evozi.injector&hl=es_419&gl=US&pli=1. [Last accessed: 10 02 2024].

[7] A Habibi Lashkari, "CICFlowMeter v4. 0," [Online]. Available: https://github.com/ahlashkari/CICFlowMeter. [Last accessed: 17 04 2024].