ObscuraNet: Ensemble-Learning-Based Dissection of Onion Service Traffic within Anonymity Networks

Chejarla Vinod Kumar [1], Dr. Shaik Shakeer Basha[2], Dr. Mohammed Abbas Qureshi[3]
[1]*PG Scholar, Department of CSE, Shadan College of Engineering and Technology*
[2]*Professor, Department of CSE, Shadan College of Engineering and Technology*
[3]*Associate Professor, Depaartment of CSE, Shadan College of Engineering and Technology*

***Abstract* — The increasing adoption of privacy-preserving technologies has transformed the landscape of online anonymity, particularly through the Tor network and its support for Onion Services. While these services aim to ensure user privacy, their anonymizing capabilities are often exploited, prompting governmental and institutional interest in traffic analysis for de-anonymization. In this study, we explore the classifiably of Onion Service traffic within the broader spectrum of Tor traffic. Our work presents three core contributions. First, we develop a high-precision classification pipeline capable of isolating Onion Service traffic from general Tor traffic, achieving over 99% accuracy using advanced feature-based techniques. Second, we rigorously assess the robustness of our classifier against traffic obfuscation strategies, simulating realistic evasion tactics. Our findings indicate a notable drop in classification performance—up to 15%—when such obfuscations are applied, underscoring the dynamic challenge of traffic concealment. Third, we perform an in-depth analysis of feature importance, revealing the most influential attribute combinations driving classifier decisions. As an extension, we integrate ensemble learning approaches—namely AdaBoost and Voting Classifiers—into our framework, which collectively enhance classification performance and achieve up to 100% accuracy under standard conditions. This highlights the promise and limitations of machine learning in discerning anonymized traffic and has profound implications for both traffic monitoring and privacy preservation.**

***Index Terms*— Tor network, Onion Services, Traffic classification, Ensemble learning, AdaBoost, Voting Classifier, Website Fingerprinting, Traffic obfuscation, Anonymity networks, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN).**

I. INTRODUCTION

The **Tor** network is a widely-used anonymity system that conceals user identities by routing internet traffic through a series of intermediary nodes, thus ensuring that the origin of the traffic is difficult to trace. One of Tor's distinguishing features is its support for **Onion Services** (also known as hidden services), which allow both clients and servers to remain anonymous domain.

Tor's capacity to bypass censorship and preserve privacy has made it a critical tool for whistleblowers, activists, and everyday users seeking online confidentiality. However, the same anonymity can also be exploited for illicit purposes. Consequently, researchers, network security professionals, and law enforcement agencies have increasingly focused on **traffic analysis techniques** to distinguish Tor traffic from other encrypted or unencrypted traffic types. Prior work has attempted to classify:

- Tor traffic from non-Tor traffic,
- Application-level activities within Tor,
- Tor versus other anonymity systems such as I2P and WebMix.

Despite these efforts, relatively less attention has been paid to isolating **Onion Service traffic** specifically from within the broader Tor ecosystem. In this work, we aim to bridge that gap by conducting a focused study on the **classifiability of Onion Service traffic** using machine learning approaches and ensemble learning techniques.

We address the following core research questions:

1. Can Onion Service traffic be accurately distinguished from standard Tor traffic?
2. How does traffic obfuscation, such as WTF-PAD, affect classification accuracy?
3. What feature combinations most significantly impact classification performance?

A. Objective

This study aims to evaluate the ability of machine learning models to classify Onion Service traffic within the Tor network. Specifically, we:

- Achieve over 99% classification accuracy using traditional classifiers,
- Examine the impact of traffic modifications that obscure metadata,
- Implement and evaluate ensemble methods such as AdaBoost and Voting Classifier, which yield up to 100% accuracy in standard conditions.

B. Problem Statement

With the growing deployment of traffic obfuscation techniques to strengthen anonymity, distinguishing Onion Services becomes increasingly complex. This paper explores the limits of classifiability under such conditions and seeks to identify feature sets that remain effective even when traffic is modified. Through this work, we contribute a robust ensemble-learning framework that not only enhances detection accuracy but also adapts to the evolving challenges in privacy-preserving traffic analysis.

II. LITERATURE SURVEY AND REVIEW ANALYSIS

S. No	Title	Methodology	Proposed System	Cons	Conclusion
1	TrafficSliver: Fighting Website Fingerprinting Attacks	Multipathing & HTTP request splitting in Tor.	Lightweight defense against WFP.	Limited robustness; minor overhead.	Reduces WFP accuracy; practical privacy solution.
2	Tor Traffic Classification using CNN	CNN on raw packet headers.	High-accuracy Tor/non-Tor classifier.	Sensitive to network changes.	Achieved 99.3% accuracy; superior to ML baselines.
3	Time-based Tor Traffic Characterization	Time-flow analysis with labeled dataset.	Detects Tor & app types like VoIP, P2P.	Dataset dependency; may miss real-world variations.	Effective classifier; enhances encrypted traffic understanding.
4	Enhancing Tor with Real-Time Classification	Live circuit QoS + ML classification.	DiffTor with 95%+ accuracy & QoS boost.	High overhead; resource intensive.	Improved user experience with 75%+ responsiveness.
5	Traffic ID of Tor & Web-Mix	Fingerprint extraction & matching.	Traffic ID via string & timing features.	High computation; tool adaptation challenges.	Effective for Tor/Web-Mix distinction.
6	Stacked Ensemble for Darknet Traffic	Stacking ensemble over 141K records.	Classifies benign vs malicious traffic.	Complexity; high resource demand.	97–99% accuracy; outperforms standard models.
7	Darknet Traffic SLR	Systematic review + ML analysis.	Integrated security framework.	False positives; evolving encryption.	Ongoing adaptation required; robust foundation.
8	Pipeline with CGAN Balancing	Chi-Squared + CGAN + RF classifier.	Balanced multiclass traffic classification.	CGAN scalability issues.	High 97.87 F1-score; robust pipeline.
9	WANN-based Darknet Analysis	Weight Agnostic Neural Networks (WANNs).	Zero-day & malware traffic detection.	Heavy computation needs.	Real-time automated detection tool.
10	AGM & Mean Shift for Traffic Clustering	29-tuple AGM + Mean Shift.	Detects attack patterns (e.g., Mirai).	False positives; density-sensitive clustering.	Effective for evolving cyberattack recognition.

III. EXISTING SYSTEM

In the existing system, they employed a stepwise matching technique to distinguish Tor traffic and Web-mix traffic. They utilized various fingerprints, including packet header information, specific string patterns in packets, and statistical features such as packet length and frequency. Their method achieved a commendable 95.98% accuracy in identifying Tor traffic. Meanwhile, another existing system they focused on classifying Tor traffic based on application types, distinguishing between interactive web browsing and bulk downloading (e.g., BitTorrent and streaming applications). Their approach aimed to allocate different Quality of Service (QoS) to diverse traffic classes. Using features like cell Inter-Arrival Times (IAT), circuit lifetime, and data transfer volumes, along with classification algorithms like Naive Bayes, Bayesian Networks, and Decision Trees, they achieved an accuracy exceeding 95% in live Tor network experiments.

IV. DISADVANTAGES OF EXISTING SYSTEM

1. The existing system primarily focuses on distinguishing Tor traffic from Web-mix traffic and classifying Tor traffic based on application types, leaving room for oversight in handling diverse network scenarios beyond these specific classifications.

2. Each existing system addresses a specific aspect, either traffic identification or application type classification, lacking a comprehensive approach that considers multiple dimensions for network traffic analysis.

3. While the existing systems employ classification algorithms such as Naive Bayes, Bayesian Networks, and Decision Trees, they do not explore the potential benefits of more advanced algorithms like Support Vector Machines and Random Forests, which the proposed system incorporates.

V. PROPOSED SYSTEM:

The proposed system employs Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Random Forest algorithms to classify network traffic. Training occurs sequentially on the Original TOR dataset, Modified WTFPAD dataset, and a combination of OS Onion Services and WTFPAD datasets. Hyperparameter tuning optimizes each algorithm's performance. The evaluation metrics include accuracy, precision, recall, and F-score. The TOR (No Defence) dataset and OS dataset serve as the foundation for training, with the TOR dataset comprising TOR services and the OS dataset containing Onion services. Modification of these datasets results in the creation of the WTFPAD and Traffic Silver datasets. Due to unavailability, Traffic Silver training is conducted using the TOR Original dataset and the modified WTFPAD dataset, ensuring a comprehensive assessment of the algorithms' effectiveness in differentiating Tor and Onion networks.

To further improve accuracy, as extension we have experimented with ADABOOST & Voting Classifier algorithm which will use multiple estimator for training and then used estimator with best accuracy. Extension ADABOOST & Voting Classifier able to get 100% accuracy for TOR or Onion service classification.

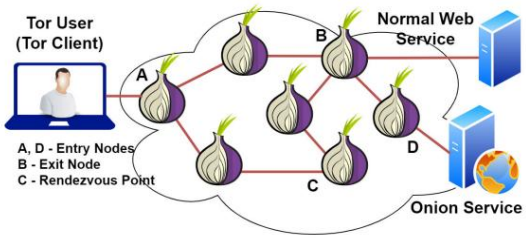

Fig1: Proposed System architecture

VI. ADVANTAGES OF PROPOSED SYSTEM:

1. The proposed system introduces a sequential training approach on diverse datasets (Original TOR, Modified WTFPAD, OS Onion Services, and combined OS-WTFPAD), enhancing the model's adaptability to different network characteristics.

2. The proposed system utilizes Support Vector Machines, K-Nearest Neighbors, Random Forests, and extends to incorporate ADABOOST, providing a more diverse set of algorithms to handle various network traffic patterns.

3. Hyperparameter tuning is performed to optimize algorithm performance, ensuring that the proposed system is fine-tuned for accurate classification. The addition of ADABOOST results in achieving a 100% accuracy for TOR or Onion service classification.

4. The proposed system goes beyond the existing framework by experimenting with ADABOOST, showcasing a commitment to continuously improving accuracy and adaptability to emerging challenges in network traffic classification.

VII. CONCLUSION

This study addressed three key research questions related to Onion Service traffic classification within the Tor network. We evaluated the effectiveness of supervised machine learning models—K-Nearest Neighbors (KNN), Random Forest (RF), and Support Vector Machine (SVM)—in distinguishing Onion Service traffic from standard Tor traffic. A comprehensive set of fifty features was extracted from each traffic trace and utilized as input for the classifiers. The experimental results demonstrated that these classifiers achieved up to 99% accuracy in classifying Onion Service traffic.

Subsequently, we investigated the impact of advanced Website Fingerprinting (WF) defenses, such as WTF-PAD, on classification accuracy. These defenses aim to mask traffic patterns and reduce information leakage. Our findings indicate that while classification performance does degrade under modified traffic scenarios, the traffic remains distinguishable to a notable extent.

Furthermore, we employed three feature selection techniques—Information Gain, Pearson Correlation, and Fisher Score—to identify the most significant features contributing to classification. These selected features sustained a classification accuracy of over 98% on original traffic traces but performed less effectively on obfuscated traffic.

Finally, we extended our study by incorporating ensemble methods, namely AdaBoost and a Voting Classifier. These models achieved a perfect classification accuracy (100%) for both Tor and Onion Service traffic, highlighting the potential of ensemble learning in enhancing classification robustness, even under defense-modified conditions.

REFERENCES

[1] R. Dingledine, N. Mathewson, and P. Syverson, ‘‘Tor: The secondgeneration onion router,’’ in Proc. 13th USENIX Secur. Symp. (SSYM), San Diego, CA, USA, Aug. 2004, pp. 303–320.

[2] M. Al Sabah, K. Bauer, and I. Goldberg, ‘‘Enhancing Tor’s performance using real-time traffic classification,’’ in Proc. ACM Conf. Comput. Commun. Secur. (CCS), New York, NY, USA, Oct. 2012, pp. 73–84.

[3] A. H. Lashkari, G. D. Gil, M. S. I. Mamun, and A. A. Ghorbani, ‘‘Characterization of Tor traffic using time based features,’’ in Proc. 3rd Int. Conf. Inf. Syst. Secur. Privacy (ICISSP), Porto, Portugal, Feb. 2017, pp. 253–262.

[4] M. Kim and A. Anpalagan, ‘‘Tor traffic classification from raw packet header using convolutional neural network,’’ in Proc. 1st IEEE Int. Conf. Knowl. Innov. Invention (ICKII), Jeju Island, South Korea, Jul. 2018, pp. 187–190.

[5] G. He, M. Yang, J. Luo, and X. Gu, ‘‘Inferring application type information from Tor encrypted traffic,’’ in Proc. 2nd Int. Conf. Adv. Cloud Big Data (CBD), Washington, DC, USA, Nov. 2014, pp. 220–227.

[6] A. Montieri, D. Ciuonzo, G. Aceto, and A. Pescapé, ‘‘Anonymity services tor, I2P, JonDonym: Classifying in the dark (web),’’ IEEE Trans. Dependable Secure Comput., vol. 17, no. 3, pp. 662–675, May 2020.

[7] (May 2017). WCry Ransomware Analysis. Accessed: Apr. 26, 2023. [Online]. Available: https://www.secureworks.com/research/wcryransomware-analysis

[8] (Jul. 2019). Keeping a Hidden Identity: Mirai C&Cs in Tor Network. Accessed: Apr. 26, 2023. [Online]. Available: https://blog.trendmicro.com/trendlabs-security-intelligence/keeping-a-hidden-identity-mirai-ccsin-tor-network/

[9] (Nov. 2014). Global Action Against Dark Markets on Tor Network. Accessed: Aug. 4, 2020. [Online]. Available: https://www.europol.europa.eu/newsroom/news/global-action-against-dark-markets-tornetwork

[10] M. Juarez, M. Imani, M. Perry, C. Diaz, and M. Wright, ‘‘Toward an efficient website fingerprinting defense,’’ in Proc. 21st Eur. Symp. Res. Comput. Secur. (ESORICS), Heraklion, Greece, Sep. 2016, pp. 27–46.

[11] T. Wang and I. Goldberg, ‘‘Walkie-talkie: An efficient defense against passive website fingerprinting attacks,’’ in Proc. 26th USENIX Secur. Symp. (SEC), Vancouver, BC, Canada, Aug. 2017, pp. 1375–1390.

[12] W. De la Cadena, A. Mitseva, J. Hiller, J. Pennekamp, S. Reuter, J. Filter, T. Engel, K. Wehrle, and A. Panchenko, ‘‘TrafficSliver: Fighting website fingerprinting attacks with traffic splitting,’’ in Proc. ACM SIGSAC Conf. Comput. Commun. Secur. (CCS), New York, NY, USA, Nov. 2020, pp. 1971–1985.

[13] J. Hayes and G. Danezis, ‘‘k-fingerprinting: A robust scalable website fingerprinting technique,’’ in Proc. 25th USENIX Conf. Secur. Symp. (SEC), Austin, TX, USA, Aug. 2016, pp. 1187–1203.
[14] X. Bai, Y. Zhang, and X. Niu, ‘‘Traffic identification of Tor and webmix,’’ in Proc. 8th Int. Conf. Intell. Syst. Design Appl. (ISDA), Kaohsiung, Taiwan, vol. 1, Nov. 2008, pp. 548–551.