A HYBRID DEEP LEARNING AND GENETIC ALGORITHM MODEL FOR EXPLAINABLE NETWORK TRAFFIC CLASSIFICATION

Ibrohimov Azizbek Ravshonbek ugli
Head of Department, State Enterprise "Cybersecurity Center"

Haydarov Elshod Dilshod o‘g‘li
Head of the department, Tashkent University of Information
Technologies named after Muhammad al-Khwarizmi

Annotation

This thesis presents an explainable deep learning model for network traffic classification using a genetic algorithm. A ResNet-based classifier is combined with a GA-driven dominant feature selection method to enhance interpretability and optimize accuracy. Experiments on real-world encrypted traffic datasets achieved 97.24% accuracy while identifying critical statistical features. The model effectively balances accuracy, simplicity, and transparency, contributing to the advancement of explainable artificial intelligence in network security analysis.

Keywords: Deep Learning; Genetic Algorithm; Explainable Artificial Intelligence; Network Traffic Classification; Feature Selection; Residual Network (ResNet); Encrypted Traffic; Model Interpretability

Introduction

The rapid growth of Internet technologies and widespread use of encryption protocols such as SSL, TLS, and VPNs have significantly increased the complexity of network monitoring and security. As traditional payload-based inspection methods can no longer access encrypted content, traffic classification has shifted toward flow-based approaches, which rely on statistical attributes like packet size, inter-arrival time, and byte count to identify services and detect anomalies.

Deep learning (DL) has proven highly effective for this purpose, as it can automatically extract complex, non-linear patterns from traffic data. However, DL models often function as “black boxes,” providing little insight into how decisions are made—an issue that limits their trustworthiness in critical network security applications. To address this, Explainable Artificial Intelligence (XAI) techniques aim to make machine learning decisions more transparent and interpretable.

This research, based on the work of Ahn et al. [1], integrates deep learning with genetic algorithms (GAs) to combine high accuracy with interpretability. A Residual Neural Network (ResNet) is used to classify encrypted traffic flows, while a GA identifies and quantifies the most influential features contributing to these classifications. This hybrid framework not only achieves a 97.24% accuracy rate on public PCAP datasets but also reveals which network characteristics—such as packet timing, size, and byte count—play key roles in differentiating Internet services like instant messaging, VoIP, and web browsing.

Model design and methodology

This study presents a hybrid model that integrates deep learning and a genetic algorithm (GA) to classify network traffic with high accuracy and interpretability. The framework consists of two main parts: a ResNet-based deep-learning classifier and a GA-driven dominant feature selection module. The classifier identifies both encrypted and unencrypted traffic using flow-level statistical data, while the GA determines which features contribute most to classification accuracy. Together, these components address the "black-box" problem typical of deep neural networks.

The classifier employs the Residual Network (ResNet) architecture [2], which uses shortcut connections to pass information between layers. This design prevents vanishing gradients, allowing deeper models to learn efficiently. ResNet's layered structure helps the model capture complex dependencies in traffic behavior while maintaining stable training. Since it uses flow-level statistics rather than packet payloads, the classifier can effectively process encrypted traffic.

The dataset preparation begins with packet gathering, where network packets are collected and grouped into flows based on a five-tuple: source and destination IP addresses, source and destination ports, and transport protocol (TCP or UDP). Packets sharing the same 5-tuple within a certain time window are treated as a single flow, representing one communication session. Including bidirectional flows — both client-to-server and server-to-client directions — helps the model recognize symmetrical patterns in network communication.

After collection, data undergoes preprocessing and feature extraction. Each flow $F = \{p_1, p_2, \dots p_n,\}$ is transformed into a vector of statistical features, including the minimum, maximum, mean, and standard deviation of both packet size and inter-arrival time, as well as total packet count and total bytes. Considering both directions, each flow is represented by 20 features, describing its temporal and volumetric behavior. All features are normalized to [0, 1] and reshaped into a 2D structure suitable for convolutional processing in ResNet.

The model was implemented in TensorFlow, trained with 70% of the data, and tested on the remaining 30%. Optimal parameters were found experimentally: batch size = 300, epochs = 5000, filters = 128, and 4–16 residual blocks. Each block includes convolution, batch normalization, ReLU activation, and skip connections. The final ResNet achieved a 97.24% accuracy, confirming its strong generalization across different traffic types.

To improve transparency, the GA-based feature selection identifies which features most affect the classifier's decisions. Each possible feature subset is represented by a binary vector θ, where "1" means the feature is included and "0" means it is excluded. The optimization objective balances simplicity and accuracy:

$$k = \lambda_1 p_1 + \lambda_2 p_2,$$

where p_1 is the proportion of dropped features and p_2 is classification accuracy, with $\lambda_1 + \lambda_2 = 1$. By adjusting λ_1 and λ_2, the model can prioritize interpretability or performance. Through selection, crossover, mutation, and elitism, the GA evolves optimal feature masks that remove redundant variables without reducing accuracy.

The dominance rate (I) quantifies feature importance:

$$I = \frac{K}{N} * 100$$

where K is the number of times a feature is selected across Nexperiments. High dominance rates identify the most influential statistical features, such as mean inter-arrival time, average packet size, and total byte count.

Experimental results show that lowering λ_2, increases feature reduction but slightly decreases accuracy (from 97% to about 94%). The model performs consistently across eight service types—instant messaging, email, file transfer, P2P, remote access, streaming, VoIP, and web browsing—achieving balanced precision, recall, and F1-scores.

Overall, the proposed ResNet–GA hybrid achieves both high performance and explainability. ResNet ensures robust classification, while GA reveals which statistical features drive decisions, turning a black-box DL model into an interpretable, adaptable tool for intelligent network monitoring and cybersecurity applications.

Conclusion

This thesis presented an explainable deep learning-based traffic classification model that integrates a genetic algorithm for dominant feature selection. The hybrid system bridges the gap between model accuracy and interpretability, addressing a critical issue in modern network analysis. Using flow-based statistics and a ResNet architecture, the classifier achieved 97.24% accuracy across diverse encrypted and non-encrypted datasets. The incorporation of a GA allowed the model to identify key features contributing to classification performance while minimizing redundancy. The feature importance analysis, quantified through the dominance rate, provided clear interpretive insights into how specific traffic characteristics influence classification.

The experiments revealed a clear trade-off between performance and interpretability, controlled by the parameters λ_1 and λ_2. A balanced configuration maintained near-optimal accuracy with substantial feature reduction. This adaptability demonstrates the model's practical potential in real-world network management, where computational efficiency and transparency are equally vital.

Overall, the proposed ResNet + GA framework contributes a significant methodological advancement to the field of Explainable Artificial Intelligence (XAI) in network security. By transforming the opaque deep learning process into an interpretable and quantifiable system, it lays the groundwork for more trustworthy and transparent traffic classification and intrusion detection systems. Future research can extend this work to real-time adaptive systems, blockchain-based security frameworks, and online GA optimization to enable dynamic, explainable, and autonomous network defense mechanisms.

References

1. S. Ahn, J. Kim, S. Y. Park, and S. Cho, “Explaining Deep Learning-Based Traffic Classification Using a Genetic Algorithm,” IEEE Access, vol. 9, pp. 4738–4750, 2021.
2. A. Adadi and M. Berrada, “Peeking inside the black-box: A survey on explainable artificial intelligence (XAI),” IEEE Access, vol. 6, pp. 52138–52160, 2018.
3. G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescape, “Mobile encrypted traffic classification using deep learning,” IEEE Trans. Network and Service Management, vol. 16, no. 2, pp. 445–458, 2019.
4. W. Wang, M. Zhu, X. Zeng, X. Ye, and Y. Sheng, “Malware traffic classification using convolutional neural networks,” Proc. Int. Conf. Inf. Netw. (ICOIN), pp. 712–717, 2017.
5. D. Gunning, “Explainable Artificial Intelligence (XAI),” DARPA Report, 2017.
6. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” Proc. IEEE CVPR, pp. 770–778, 2016.
7. H. Kim and N. Feamster, “Improving network management with software-defined networking,” IEEE Commun. Mag., vol. 51, no. 2, pp. 114–119, 2013.
8. M. Karakus and A. Durresi, “Quality of service in software-defined networking: A survey,” J. Netw. Comput. Appl., vol. 80, pp. 200–218, 2017.