Theory and Practice of Science and Technology
2025, VOL. 6, NO. 8, 129-134
DOI: 10.47297/taposatWSP2633-456924.20250608

Encryption Traffic Anomaly Detection Technology Using Artificial Intelligence Methods

Qinglong Liang

International Exchange College, Guangzhou Maritime University, Guangzhou, Guangdong, 510725, China

ABSTRACT
To safeguard against numerous network attacks in environments where communication is not encrypted, an increasing number of online application services are adopting encrypted traffic to improve privacy and security. However, it was soon found that attackers could exploit secure communication channels to conduct cyber attacks. Traditional techniques, which rely on deep packet inspection to detect anomalous traffic, struggle to identify these types of attacks effectively. In recent years, the progress of artificial intelligence has prompted many researchers to apply AI-based approaches for detecting encrypted network traffic. This paper centers on the UNSW-NB15 dataset and provides a comprehensive retrospect of advancements in encrypted traffic anomaly detection techniques based on this dataset over the past few years. Initially, it categorizes and summarizes existing studies from three perspectives: traditional machine learning, deep learning, and unsupervised or semi-supervised learning methods, while comparing their performance in terms of classification accuracy, detection efficiency, and suitability for different scenarios. Furthermore, by evaluating and analyzing the experimental outcomes of these approaches on the UNSW-NB15 dataset, the paper highlights the strengths and limitations of various technical strategies. Finally, it outlines the current challenges and potential directions for future research in the field of encrypted traffic analysis.

KEYWORDS
Encrypted traffic; Anomaly detection; Artificial intelligence; UNSW-NB15; Deep learning

1 Introduction

Nowadays, with the rapid development of internet services, user data security and privacy protection requirements have continuously increased, leading to the broad application of encrypted communications protocols (such as HTTPS, TLS 1.3, etc.) [1]. However, while effectively ensuring privacy, the encryption of network traffic also brings new challenges to traditional cybersecurity protection means. Intrusion detection systems (IDS) relying on deep packet inspection (DPI) have found it difficult to directly analyze encrypted data payloads, thereby allowing malicious traffic to potentially hide within normal communications, increasing the difficulty of detection [2-3]. Particularly in complex attack scenarios such as zero-day attacks and encrypted tunnels, how to achieve effective anomaly detection without decrypting network traffic has become a current research hot spot in the field of cybersecurity [4].

To address this challenge, numerous studies have attempted to introduce artificial intelligence methods into encrypted traffic anomaly detection, including traditional machine learning, deep learning, and unsupervised/semi-supervised methods. These methods have shown some effectiveness in encrypted traffic anomaly detection, but still face limitations in terms of accuracy, computational overhead, interpretability, and practical deployment [5-8].

In recent years, researchers have commonly verified these methods using public datasets. Among these, the UNSW-NB15 dataset is extensively used in encrypted traffic anomaly detection research because it contains various real attack traffic and rich features [9]. However, existing literature mostly focuses on improving the performance of a single method, lacking a systematic comparative analysis of different methods on the same dataset.

Therefore, this paper focuses on the UNSW-NB15 dataset, reviewing and comparing the applications and performance of artificial intelligence methods that have utilized this dataset for encrypted traffic anomaly detection over the past five years. The primary contributions of this paper include: (1) Summarizing existing research progress across three areas: traditional machine learning, deep learning, and unsupervised learning; (2) Systematically comparing the detection accuracy, efficiency, and applicable scenarios of different methods based on UNSW-NB15 experimental results, and revealing their respective advantages and disadvantages; (3) Analyzing the challenges in current research and envisioning potential future development directions.

2 Background

Intrusion Detection System (IDS) is an active network security protection technology, typically categorized into signature-based and anomaly-based [10]. The former maintains a known attack signature database and can efficiently

* **Corresponding Author:** Qinglong Liang, luengkieron@gmail.com

detect known threats. The latter, by learning normal traffic network and identifying deviations in traffic behavior to detect potential intrusion, is more suitable for addressing the current environment of unknown attacks and encrypted traffic, and thus has gradually gained widespread attention in recent years [11].

In terms of experimental assessment, researchers usually rely on public datasets to conduct algorithm validation. Early KDD 99 and NSL-KDD datasets have been widely used, but there is a larger gap between their attack types and modern network environment [12-13]. In recent years, datasets such as UNSW-NB15 and CIC-IDS-2018 have been far used in encrypted traffic anomaly detection study due to their coverage of richer attack categories and network behavior [9,14]. Among these, the dataset of UNSW-NB15 includes 2,540,044 records, including 49 dimensions of traffic features, covering normal and multiclass attack traffic. Notably, this dataset has already been divided into training set and test set, with an overall of 175, 341 records and 82, 332 records respectively, making it one of the most commonly used benchmark datasets currently.

3 Review of existing methods

3.1 Traditional machine learning methods

Traditional machine learning relies on handcrafted feature extraction, combined with classifiers for detection. It relies on a lot of high-quality manually labeled information, which to some extent boosted detection accuracy, but its performance is limited when processing high-dimensional complex network traffic [5]. Its methods include k-nearest neighbors (k-NN), decision trees (DT), and random forests (RF), which are briefly introduced below.

K-nearest neighbors (KNN) is a simple distance-based classification method [15]. It selects the K points with the minimal distance to the target sample and classifies the target sample point into the type that has the highest proportion among these K points. KNN is suitable for classifying small-scale datasets, but its inferencing efficiency is low for high-dimensional complex network traffic datasets like UNSW-NB15.

Decision Tree (DT) consists of Root nodes, internal nodes, branches, and leaf nodes [16]. Internal nodes function as tests, branches illustrate the possible results of those tests, and leaf nodes provide the classification decisions. In practical research, decision trees are often combined with ensemble learning (such as PROAFTN, XGBoost) to improve performance and robustness [17].

Random Forest (RF) is an ensemble learning approach that utilizes many decision trees [18]. It addresses the problem of decision trees being prone to overfitting and their inability to handle high-dimensional data. The final answer is determined by voting on the judgments of each tree, with the majority rule, thereby improving the accuracy of the results.

3.2 Deep learning methods

Deep learning, grounded in deep neural networks, represents a specialized form of machine learning. It can automatically extract time series and spatial features, significantly improving detection capability, but it faces problems such as high computational overhead and insufficient interpretability [6]. Below, I will introduce four relatively mainstream deep learning methods.

Convolutional Neural Network (CNN) can automatically extract features from local to global, and is particularly suitable for processing image and time series data [19]. CNN is composed of pooling layer, convolutional layer, and fully connected layer. The primary function of the convolutional layer is to capture and extract localized features, the pooling layer for compressing information, and the fully connected layer for integrating features and outputting the final classification results.

Recurrent Neural Network (RNN) denotes a category of neural network architecture that is particularly developed to handle and analyze sequential information [20]. Structurally, RNN employs recurrent connection, which endows the network with memory function over time, thus providing the model with contextual temporal dependencies.

Long Short-Term Memory (LSTM) is an optimization of RNN [21]. It addresses the gradient vanishing and gradient explosion problems inherent in RNNs when dealing with long sequences, and is widely applied in Natural Language Processing (NLP) and other domains. The LSTM's core consists of the memory cell and gating mechanism, which enable it to selectively remember and update the memory cell.

Transformer model is a recent research hotspot in IDS/NIDS. It is built upon the self-attention mechanism, which makes it possible to capture long-distance relationships effectively without relying on recurrent structures [22]. Existing research indicates that Transformer has achieved superior detection accuracy compared to traditional deep learning models on UNSW-NB15 and similar datasets [23]. However, this model has a large computational scale, and still faces challenges for software deployment to the current network environment.

3.3 Unsupervised/semi-supervised learning methods

For unsupervised learning, its training data set has no labels, and it learns its inner structure and logical patterns by data mining on unlabeled samples. Its main types include clustering, dimensionality reduction, and anomaly detection, etc [24]. For semi-supervised learning, most of its training data set is unlabeled, with a small portion being labeled, aiming to reduce labeling cost while achieving high generalization performance [25-26]. Both these methods have potential in detecting unknown attacks, but they have a high false positive rate in experiments, making them difficult to deploy at scale in practical scenarios [8]. Below, I will introduce recent research on unsupervised/semi-supervised learning.

Xing et al. proposed an unsupervised and stable Encrypted traffic anomaly monitoring method called D2LAD [27]. D2LAD employs an idea similar to autoencoders. However, it performs feature representation and reconstruction through a deep dictionary learning structure. The model first learns a sparse dictionary representation of normal encrypted traffic, and then identifies anomalous traffic by computing the reconstruction error.

Kim et al. proposed an unsupervised learning method [28]. This method is a one-class learning method based on autoencoders. For detection, it uses packet/flow metadata stripped of node identification (e. g., IP, port) for feature extraction, and then determines anomaly by means of the reconstruction error.

Xu et al. designed a self-supervised method that combines graph encoders with graph contrastive learning [29]. By learning network traffic features through graph embedding, it improved the accuracy and generalization capability of encrypted traffic anomaly detection without the need for extensive annotations.

4 Comparative analysis on UNSW-NB15

4.1 Explanation of comparison dimensions

To ensure the fairness and consistency of the comparative analysis, this paper selects UNSW-NB15 as the unified benchmark dataset, and collates and conducts a comparison of relevant studies in the recent 5 years from four dimensions: method category, feature engineering, performance metrics, and applicability.

4.2 Comparison table

Table 1 compares the performance of different AI methods on UNSW-NB15. It can be seen that deep learning and self-supervised methods are generally superior to traditional methods in terms of performance metrics; they have higher accuracy but a large computational overhead. Semi-supervised methods show potential in unknown attacks detection, but lack stability. Ensemble and federated methods are closer to practical application, balancing efficiency and privacy, but are still constrained by interpretability and generalization.

4.3 Comparative analysis

4.3.1 Comparison of method categories

Deep learning methods performed best overall on UNSW-NB15, especially achieving high accuracy in the recognition of complex attack traffic; ensemble learning demonstrated more stable detection performance and can maintain balance across different attack types; Unsupervised and semi-supervised methods show some potential when addressing Unknown Attacks, but suffer from a higher false positive rate; Recently emerging self-supervised and federated learning frameworks, although the field of study is still under early development, have shown promise in reducing labeling requirements and protecting privacy.

4.3.2 Comparison of feature engineering

Different studies show significant differences in their processing methods on the UNSW-NB15 dataset. Some works still rely on traditional manual feature selection, such as supervised learning, traditional machine learning, and so on. These ways are simple and can reduce the dimension to a certain extent, but often have limited expressive power for high-dimensional complex traffic. Nowadays, many studies tend to favor automatic feature extraction, for example, by directly learning raw traffic features through deep neural networks, which reduces manual intervention and enhances the temporal and spatial expressive power of features. Some studies also combine multi-modal or hybrid feature engineering strategies, fusing statistical features with embedded features to balance detection accuracy and generalization ability.

4.3.3 Comparison of performance

Most studies achieved high detection accuracy and F1-score on the UNSW-NB15 dataset, generally exceeding 90%.

Table 1 Comparative table on UNSW-NB15

Article	Method category	Method / Framework	Feature Engineering	Performance Metrics	Strengths and Weaknesses
Sattar et al. [30]	Self-supervised	ET-SSL	Automatic + flow-level	Acc: 95.1% Pre: 93.9% Rec: 94.8% F1: 94.3%	The detection rate of zero-day attacks is high, but the computational resource cost is large.
Bahlali et al. [31]	Full supervised	Auto encoder	Manual	Acc: 91.17%Pre: 81.32% Rec: 96.45% F1: 92.32%	Improve the accuracy of encrypted traffic detection, but the performance depends on the balance of the data set.
Liu et al. [32]	semi-supervised	GRU+GCN	Multi-modal (Sequence + Graph)	Acc: 99.51% Pre: 99.46% Rec: 99.36% F1: 99.41%	Efficiently handle sample imbalance and unknown attacks, but with relatively high computational complexity.
Bakhshi et al.[33]	DL	CNN + GRU	Automatic	Acc: 91.21% Pre: 91.19% Rec: 91.36% F1: 93.61%	The detection accuracy is high and the generalization ability is strong, but the training time is relatively long.
Bahlali et al.[34]	Semi-supervised	Custom-AE classifier	Manual + Automatic (SSL)	Supervised Acc: 90.95% F1: 92.20% Semi: 90.34%	Introduce an innovative masking strategy, but the generalization ability and practicality still need to be verified.
Tufan et al.[35]	ensemble	ensemble learning and CNN	flow-level + Automatic + Temporal	ensemble : Acc: 95.1% F1: 94.9% CNN: Acc: 95.9% F1: 95.8%	Modeling is based on real production environment data, but the attack types are limited and the model lacks interpretability.
Marfo et al.[36]	Federated	ANN + Federated Learning	Manual	Acc: 97.21% Pre: 98.09% Rec: 97.18% F1: 97.63%	Efficient detection was achieved while protecting privacy, but there was a data imbalance.

Deep learning-based CNN models achieved 95% accuracy on UNSW-NB15, and even GRU+GCN models achieved 99% accuracy on UNSW-NB15, while autoencoder-based unsupervised methods only achieved 90.34% accuracy. This leads to the conclusion that deep learning-based methods are generally superior to traditional and unsupervised methods, especially demonstrating stronger classification capability in complex attack scenarios [32, 34, 35]; Ensemble learning and multi-model fusion offer greater advantages in stability, capable of reducing single model fluctuations across different attack types. In contrast, unsupervised and semi-supervised methods, although slightly inferior in accuracy, show some potential in identifying unknown attacks, for example, the Area Under the Curve (AUC) of semi-supervised learning reaches 98% [32].

4.3.4 Comparison of applicability

The adaptation capabilities of different methods to actual network environment vary. Self-supervised and semi-supervised methods (e. g., ET-SSL by Sattar et al., Custom-AE by Bahlali et al.) can be trained under limited labeling conditions and possess the potential to identify zero-day attacks, but their generalization and practicality still require more verification [30, 34]. Deep learning methods (e. g., CNN+GRU by Bakshi et al.) excel in detection accuracy and generalization, but often have longer training times, which may be detrimental to deployment in real-time scenarios [33].

In specific scenarios, multi-modal feature fusion (e. g., GRU+GCN by Liu et al.) demonstrates strong capabilities in processing imbalanced samples and unknown attacks, making it more suitable for complex network traffic environments [32]. Ensemble learning (Tufan et al.) boosted the stability of detection results, but its applicability in large-scale environments is subject to certain limitations due to limited attack types and insufficient model interpretability [30]. The federated learning method (Marfo et al.) achieved a relatively high detection rate while guaranteeing privacy protection,

indicating its strong potential for application in distributed environments, but it still faces the problem of data imbalance [36].

4.4 Brief summary

Overall, deep learning and multi-modality methods are more suitable for high-precision, complex attack scenario, while self-supervision/semi-supervision and federated learning methods show potential in data scarcity, privacy protection, and zero-day attack detect, but the problems of generalization and deploy overhead still need to be solved.

5 Challenges and future directions

At the dataset level, despite its common use, UNSW-NB15 has limited and unevenly distributed attack types, which affects the model's generalization ability [37]. At the model level, while deep learning methods offer high accuracy, they suffer from black-box characteristics and high computational overhead issues; unsupervised and semi-supervised methods show potential for zero-day attacks, but have a high false positive rate. At the deployment level, how to balance real-time performance, scalability, and adversarial attack resistance remains a current pain point that needs to be urgently addressed. Future research can unfold in four areas: (1) Building larger-scale and more diverse encrypted traffic datasets and promoting cross-dataset verification; (2) Exploring self-supervised learning and generative models to enhance unknown attacks detection capability and reduce reliance on labeling; (3) Developing lightweight and interpretable models to meet the requirements of edge devices and practical deployment; (4) Combining federated learning with adversarial defense to enhance privacy protection and system robustness.

6 Conclusion

This paper reviews artificial intelligence encrypted traffic anomaly detection research based on the UNSW-NB15 dataset, comparing the feature engineering, performance, and applicability of different methods. The results indicate that deep learning demonstrates superiority in accuracy, while emerging methods such as self-supervised learning and federated learning show potential in unknown attacks and privacy protection. Future work needs to strike a balance between enhancing generalizability and deployment feasibility.

References

[1] Keshvadi, S., Sharma, Y. Exploring HTTPS Certificate Ecosystem: Analyzing the Entire IPv4 Address Space. In: 2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Regina, SK, Canada, 24-27 September 2023. pp. 547--552

[2] Wang, W., Zhu, M., Zeng, X., Ye, X., Sheng, Y. Malware traffic classification using convolutional neural network for representation learning. In Proceedings of the 2017 International Conference on Information Networking (ICOIN), Da Nang, Vietnam, 11–13 January 2017; IEEE: New York, NY, USA, 2017; pp. 712–717.

[3] Anderson B, Paul S, McGrew D. Deciphering malware's use of TLS (without decryption) [J]. Journal of Computer Virology and Hacking Techniques, 2018, 14(3): 195-211.

[4] Sommer R, Paxson V. Outside the closed world: On using machine learning for network intrusion detection[C]. Proceeding of 2010 IEEE Symposium on Security and Privacy. Los Alamitos, CA: IEEE Computer Society, 2010: 305-316.

[5] Shone, N., Ngoc, T. N., Phai, V. D., Shi, Q. (2018) A deep learning approach to network intrusion detection. IEEE Transactions on Emerging Topics in Computational Intelligence, 2(1): 41–50.

[6] Thirimanne, S., Jayawardana, L., Yasakethu, L. (2022) Deep neural network based real-time intrusion detection system. SN Computer Science, 3: 145.

[7] Tama, B. A., Lim, S. (2021) Ensemble learning for intrusion detection systems: A systematic mapping study and cross-benchmark evaluation. Computer Science Review, 40: 100388.

[8] Kim D, Im H, Lee S. Adaptive autoencoder-based intrusion detection system with single threshold for CAN networks. *Sensors*. 2025;25(13): 4174.

[9] Moustafa N, Slay J. UNSW-NB15: A comprehensive data set for network intrusion detection systems[J]. Proceedings of the 4th International Conference on Cyber Security and Digital Forensics (CyberSec), 2015; pp. 1–6.

[10] Liao H J, Lin C H R, Lin Y C, et al. Intrusion detection system: A comprehensive review[J]. *Journal of Network and Computer Applications*, 2013, 36(1): 16-24.

[11] Sommer, R., Paxson, V. (2010) Outside the closed world: On using machine learning for network intrusion detection. 2010 IEEE Symposium on Security and Privacy, pp. 305–316.

[12] Tavallaee, M., Bagheri, E., Lu, W., Ghorbani, A. A. (2009) A detailed analysis of the KDD CUP 99 data set. IEEE Symposium on Computational Intelligence for Security and Defense Applications, pp. 1–6.

[13] NSL-KDD dataset. University of New Brunswick. Available: http://www.unb.ca/cic/datasets/nsl.html
[14] Sharafaldin, I., Lashkari, A. H., Ghorbani, A. A. (2018) Toward generating a new intrusion detection dataset and intrusion traffic characterization. ICISSP, pp. 108–116.
[15] Cover T, Hart P. Nearest neighbor pattern classification[J]. IEEE Transactions on Information Theory, 1967, 13(1): 21-27.
[16] Safavian S R, Landgrebe D. A survey of decision tree classifier methodology[J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1991, 21(3): 660-674.
[17] Al-Obeidat F, El-Alfy E S M. Hybrid multicriteria fuzzy classification of network traffic patterns, anomalies, and protocols[J]. *Personal and Ubiquitous Computing*, 2019, 23(5-6): 777-791.
[18] Biau G, Scornet E. A random forest guided tour[J]. *Test*, 2016, 25(2): 197-227.
[19] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
[20] Elman J L. Finding structure in time[J]. Cognitive Science, 1990, 14(2): 179-211.
[21] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation, 1997, 9(8): 1735-1780.
[22] Zhang Y, Wang X, Li J, et al. *An effective intrusion detection system based on wrapper feature selection and Multi-Head Attention Transformer*. Scientific Reports, 2025, 15: 11348.
[23] Alzahrani A, Alsaeedi A, Alharthi M, et al. *IDS-MTran: A multi-scale Transformer model for intrusion detection systems*. Scientific Reports, 2024, 14: 74214.
[24] Jain A K, Murty M N, Flynn P J. Data clustering: a review[J]. ACM Computing Surveys, 1999, 31(3): 264-323.
[25] Chapelle O, Schölkopf B, Zien A. (2006) Semi-supervised learning. Cambridge: MIT Press.
[26] Ongsulee P. Artificial intelligence, machine learning and deep learning. Proceedings of the IEEE 15th International Conference on ICT and Knowledge Engineering (ICT&KE). IEEE, 2017: 1-6.
[27] Xing J C, Wu C M. Detecting anomalies in encrypted traffic via deep dictionary learning[C]. Proceeding of IEEE. *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2020: 734-739.
[28] Kim M-G, Kim H. Anomaly detection in imbalanced encrypted traffic with few packet metadata-based feature extraction. Computer Modeling in Engineering & Sciences, 2024, 141(1): 585–607.
[29] Xu R J, Wu G W, Wang W P, Gao X, He A, Zhang Z P. Applying self-supervised learning to network intrusion detection for network flows with graph neural network[J]. Computer Networks, 2024, 248: 110495.
[30] SATTAR S, KHAN S, KHAN M I, et al. Anomaly detection in encrypted network traffic using self-supervised learning[J]. Scientific Reports, 2025, 15: 1-21
[31] BAHLALI A R, BACHIR A, CHERIET A. Malicious encrypted network traffic detection using deep auto-encoder with a custom reconstruction loss[C]. Proceedings of the 2025 IEEE International Conference on Cyber Security and Privacy*, 2025: 1-10.
[32] LIU M, YANG QC, LIU SL. Semi-Supervised Encrypted Malicious Traffic Detection Based on Multimodal Traffic Characteristics[J]. SENSORS, 2024, 24(20): 6507.
[33] BAKHSHI T, GHITA B. Anomaly Detection in Encrypted Internet Traffic Using Hybrid Deep Learning[J]. SECURITY AND COMMUNICATION NETWORKS, 2021: 5363750.
[34] BAHLALI A R, BACHIR A, LABED A. Self-Supervised Learning Meets Custom Autoencoder Classifier: A Semi-Supervised Approach for Encrypted Traffic Anomaly Detection[J]. IEEE ACCESS, 2025, 13: 139141-139154.
[35] TUFAN E, TEZCAN C, ACARTÜRK C. Anomaly-Based Intrusion Detection by Machine Learning: A Case Study on Probing Attacks to an Institutional Network[J]. IEEE ACCESS, 2021, 9: 50078-50092.
[36] MARFO W, TOSH D K, MOORE S V. Network Anomaly Detection Using Federated Learning[C]. Proceedings of 2022 IEEE Military Communications Conference (MILCOM). Rockville, MD, USA: IEEE, 2022: 1-6.
[37] Moustafa N, Slay J. UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). Proceedings of Military Communications and Information Systems Conference (MilCIS). IEEE, 2015: 1-6.