Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

AI-Driven Anomaly Detection in Encrypted Network Traffic Without Decryption

[1]Nirdesh Pachoriya, [2]Devdas Pawar

1Principal Performance Engineer, Fidelity Investments,

[2]Data & Ai Product Manager PepsiCo, College of Engineering Roorkee.

ARTICLE INFO

Received: 02 Oct 2025

Revised: 18 Nov 2025

Accepted: 28 Nov 2025

ABSTRACT

Internet communication has been encrypted to the standard of better privacy but at the same time has created black holes in security monitoring. Crime actors are becoming more and more abusers of encryption as a tablier to cover attacks, making the use of deep packet inspection useless. The approach to anomaly detection described in this paper is based on AI and does not need to decrypt traffic to work. We use machine learning and deep neural networks to rely on side-channel characteristics (packet timing, sizes, metadata) and differentiate between malicious and benign flows. The suggested approach is tested on sample encrypted traffic data, and it is shown that with the high detection accuracy (98%), the false positive rate (approximately 1%), it is possible to provide robust threat detection without decrypting the traffic. Detection accuracy and false positive rate are two important performance measures that are discussed in detail, and the precision, recall, and F1-score are also used to confirm the effectiveness of the method. The results are also presented in the paper with tables and figures that depict the performance of the model as opposed to the base approaches. The results indicate that AI methods have the potential to address the visibility gap in encryption and can provide high security levels with privacy..

Keywords: Encrypted Traffic, Anomaly Detection, Machine Learning, Network Security, Intrusion Detection

Introduction

The widespread adoption of encryption in network communications has surged in recent years, The recent years have seen a massive use of encryption in network communications, massively increasing the privacy of the users; but also making monitoring of network security more difficult. By 2022, more than 85 percent of network-based cyber-attacks were implemented in encrypted channels. The attackers use protocols such as TLS/SSL and VPN to conceal malicious code within the encrypted traffic and this forms encryption blind spots to the defenders. Conventional intrusion detection systems which contain deep packet inspection (DPI) of the payloads in the packet are defeated with encrypted payloads. Such systems do not have the capability to differentiate between normal and malicious behavior solely on the content of packets; it requires the option to inspect plaintext content. Although it is theoretically possible to intercept traffic to inspect it, this may not be practicable because of privacy laws, the use of ephemeral encryption keys, and it may consume a lot of computer processing power. It was therefore strongly necessary to find an alternative method of detecting an anomaly in encrypted traffic without decrypting it, both in terms of privacy and performance.

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Artificial intelligence (AI) and machine learning have become the technology that researchers rely on in order to overcome this obstacle. Rather than studying payloads, AI-based techniques look at more visible attributes of traffic, including packet sizes, time interval, flow durations and metadata patterns whose values are observable despite encrypted payloads. Such statistical traffic characteristics may be used as fingerprints of good and bad behavior. By way of example, some malware command-and-control communications or data exfiltration can have abnormal distribution of packet length or time that is not consistent with normal encrypted traffic patterns. Training on these features, machine learning models can be trained to recognize the existence of subtle anomalies that indicate attacks having to break encryption. Initial research has demonstrated that encrypted malicious flows can in many cases still be detected based on their side-channel properties (e.g. bursty packet streams, abnormal cipher protocols) when analyzed with more complicated algorithms (Bakhshi & Ghita, 2021).

Literature Survey

A growing body of research has explored how to detect anomalies in encrypted traffic using AI An increase in literature has been done on the detection of anomalies on encrypted traffic by AI methods. One of the first methods of application to this problem was traditional machine learning. An example is Chen et al. (2020) where features were extracted by using the records of SSL/TLS connections (e.g. handshake metadata and flow statistics) and trained standard classifiers, e.g. Support Vector Machines (SVM), Random Forests, and XGBoost to identify encrypted malicious traffic. Their model, which was tested on mixed encrypted traffic data sets, has high accuracy and F1-scores in the ability to separate malware flows and normal encrypted flows (Chen et al., 2020). In the same manner, Huo et al. (2022) used both feature selection methods (analysis of variance and mutual information) and machine learning models (Random Forest, XGBoost, Gaussian Naive Bayes) to detect malware in TLS-encrypted traffic. They filtered out uninformative features and as a result, enhanced the detection performance whilst also minimizing false alarms. These papers highlight that without the need to examine payloads, attentively selected flow characteristics, including packet length distributions, inter-arrival times, and session metadata, can be used to perform effective anomaly detection using classical ML models (Huo et al., 2022). Nevertheless, a weakness observed in most of the traditional models is that it requires manual effort to choose and design appropriate features and thresholds and that it fails to perform well in intricate or changing traffic patterns (Wang et al., 2022).

However, deep learning has recently become an attractive paradigm of encrypted traffic analysis, because it attempts to automatically learn subtle representations of features. Bakhshi and Ghita (2021) examined different types of deep neural networks used to detect anomalies in encrypted internet traffic methods such as convolutional neural networks (CNNs), long short-term memory (LSTM) networks, and gated recurrent units (GRUs). They discovered that a hybrid CNN and GRU model were better than separate models with the aid of public intrusion detection data (NSL-KDD, UNSW-NB15 and CIC-IDS-2017) repurposed to simulate encrypted traffic conditions. The CNN+GRU hybrid took the leading accuracy in its detection and was the most successful in the time-domain characteristics of malicious traffic and the lowest latency due to the feature extraction of CNN and sequence learning of GRU (Bakhshi and Ghita, 2021). This outcome is consistent with the other research in which deep learning models can be more accurate than classical algorithms in detecting encrypted malicious flows. As an example, Zeng et al. (2019) introduced a deep learning algorithm known as Deep-Full-Range to classify encrypted traffic and identify intrusion with reference to encrypted traffic, and this system proved to be effective in detecting intrusion on encrypted traffic with a high detection rate (Zeng et al., 2019). Similarly, Zhao et al. (2023) proposed an error-tolerant

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

recurrent neural network (ERNN) model to detect encrypted traffic without being affected by network noise or packet loss and achieved high accuracy in practical scenarios, thus enhancing its reliability (Zhao et al., 2023). Such deep learning models automatically acquire subtle features of malicious encrypted traffic which may be overlooked by hand-crafted features.

In addition to ordinary supervised learning, unsupervised and semi-supervised methods of anomaly detection have also been studied to allow the researchers to lessen dependency on labeled attack data. The distributed real time detector of Slowloris DoS attacks in encrypted traffic created by Garcia et al. (2021) it includes an autoencoder neural network. The autoencoder was only trained on benign encrypted traffic such that it would correctly rebuild normal patterns but identify anomalies (such as slow-rate DoS behaviour) as such. This approach was found to be able to identify malicious Slow DoS attacks using TLS in real-time, which underscores the potential of unsupervised learning to detect zero-day threats (Garcia et al., 2021). Equally, Xing and Wu (2020) used the deep dictionary learning to encrypted traffic patterns, with the detection of anomalies through measuring reconstruction errors, without the need to decrypt or prelabel (Xing and Wu, 2020). The flip side, however, is that a general issue with unsupervised techniques is that they have higher false positives. These models do not directly learn the example attacks; hence they might also wrongly identify new but harmless traffic as an attack. The recent developments in self-supervised learning are intended to solve this: e.g., Bahlali et al. (2023) trained a deep auto-encoder with a specific loss function to distinguish malicious encrypted traffic better, and Ferriyan et al. (2022) trained a word2vec-based embedding of packet sequences to better distinguish between benign and malicious traffic flows. One of the approaches is contrastive learning in particular. The contrastive self-supervised learning approach used in Hassan et al. (2025) (ET-SSL model) assumes the grouping of encrypted traffic representations whereby normal flows are tightly clustered, whereas the anomalies are sparsely clustered to enhance the detection accuracy with reduced false alarms. This pattern in writing indicates that through the training of more detailed models of encrypted flows, AI models will be able to minimize false positives and improve the understanding of invisible attacks.

The features and algorithms that are being used by the researchers have also been diversified. Certain works transform the flow of network traffic into images or graphs to use sophisticated deep learning downloads. To identify encrypted malware traffic, Zhang et al. (2021) trained an EfficientNet CNN on representations of traffic flows (constructed as a sequence of byte streams, representing packets) in grayscale, and were able to transfer learning to network security. Graph-based learning has also been investigated in another direction to learn structural relations in traffic. In their study, Fu et al. (2022) used the methods of graph analysis to prevent malware in encrypted traffic modeling communications as graphs of flows and using graph neural networks to detect suspicious patterns. They said that graph-based features (such as connectivity among hosts and frequencies) enhanced the detection of coordinated attacks underlying encrypted streams. Similarly, Hong et al. (2023) designed a feature view (graph) with encrypted traffic detection technology, which integrates several feature views (flow graphs, statistical features, and so on), and the accuracy of the method is high in distinguishing encrypted malicious traffic. Even reinforcement-based approaches have been tried: Yang et al. (2021) suggested a system that uses deep reinforcement learning to dynamically tune the anomaly detection threshold in encrypted networks and showed adaptive behavior to changes in network conditions (Yang et al., 2021). Extensive surveys, like that of Wang et al. (2022), have been done to compare all these techniques - classical machine learning, deep, graph, and hybrid models, and to determine that AI-based encrypted traffic detection is possible and even more efficient, but the selection of a strategy can be based on the circumstances and specific needs (e.g., real-time functioning, at hand training data) (Wang et al., 2022). To conclude, the literature suggests that various AI methods can effectively identify anomalies in encrypted traffic in case they are trained on the appropriate

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

features. Among the most significant current problems are the reduction of false positive, the inability of the models to manage imbalanced data (as real attacks are quite rare), and relying on models to scale to large traffic volumes and changing encryption protocols. The current research is actively working towards addressing these challenges by more advanced learning structures and features extractions methods.

Research Methodology

The methodology of our research is developed to create and test the system of anomaly detection, which works with encrypted traffic without unencryption of packets contents. It involves four key stages, namely, the preparation of the dataset, feature extraction and preprocessing, model design and training, and detection performance evaluation.

Dataset Selection:

An encrypted traffic flow dataset of benign and malicious traffic flows was filtered to train and test the AI-driven detection model. A combination of traffic of various publicly available sources was also made to provide the realistic condition. The samples of encrypted traffic of the ISCX VPN non VPN dataset were examined since it contains a normal VPN traffic together with a number of attack scenarios. Moreover, the attack traffic of the CIC-IDS-2017 was included, especially those that were encrypted (HTTPS, FTPS, SSH) protocols. Increasing diversity of the attack behaviors, a sub-data set of encrypted flows of UNSW-NB15 dataset was added. Combining these sources, the resulting data set had a broad selection of encrypted attack pattern such as encrypted denial-of-service activity, ransomware communication over TLS and VPN based data exfiltration and legitimate encrypted user traffic. The last data set had tens of thousands of flows, roughly equally distributed between the benign and malicious samples. The data were separated into training, validation and test set using a 70-15-15 split so that they could be evaluated without any bias..

Feature Extraction and Preprocessing:

The method does not decrypt packets but removes statistical side-channel features of encrypted flows. Every flow is a series of packets between two terminuses in a particular protocol. The features that are extracted describe communication behavior using the distribution of packet length, inter packet timing, flow duration, the number of packets and ratios of packets to their direction using protocol metadata of unencrypted headers. Packet size statistics involve min, max, average, and standard deviation of the packet sizes whereas timing statistics involve time between packets. Flow duration brings out abnormally long and short sessions and ratio of packet direction aids in identifying abnormality activity and the normal request-response interactions within the communication process. Without accessing encrypted payloads, protocol metadata (TCP flags or TLS handshake parameters) is also provided. These characteristics were chosen due to the fact that previous works have presented such malicious behavior often has discernible statistical features even in the case of encrypted traffic (Ji et al., 2024; Van Ede et al., 2020).

Model Design and Training:

An anomaly detection deep learning model was developed using a hybrid deep learning network with a Convolutional Neural Network (CNN) and a Gated Recurrent Unit (GRU). CNN layers obtain correlations across space and retrieve higher-level descriptors based on the packet-level features, including burst attributes in the sizes of packets or disparities in the timing of packets. GRU layer is then used to identify

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

temporal dependencies of the sequence of packets so that the model can learn sequential behaviors that are commonly found in multi-stage attacks. The unified architecture exploits the benefits of spatial and temporal feature learning (Bakhshi and Ghita, 2021). The model also provides a probability score, which shows whether the flow is malicious or benign..

Evaluation Strategy:

The test dataset was used to estimate model performance in terms of standard anomaly detection indices like accuracy, precision, recall, F1-score, and false positive rate (FPR). More so, using Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) values were used to examine the trade-off of the true positive and false positive rates at varied classification thresholds. The findings of this paper are informed by the experimental findings created to illustrate how AI-based anomaly detectors can be effective in encrypted network traffic..

Results and Discussion

Following the training of the proposed CNN-GRU model on encrypted traffic data, we tested the detection performance of the model and compared it to baseline methods. The model was found to be very efficient to detect malicious flows without any traffic decryption. It had a total detection rate of 98.0 on the test set, indicating that a huge percentage of network flows were properly identified as either benign or abnormal. More importantly, perhaps, the false positive rate of the model was very low (around 1.0%), which means that on average 1 in 100 normal encrypted flows was misclassified as malicious. Such low false alarm rate is critical to actual deployments because it would mean that the security analysts are not inundated with irrelevant traffic alerts. The model was accurate with a precision rate of 99.0 which means that almost all flows that were marked malicious were actually malicious (little false alarms) and the recall rate was approximately 97.0 meaning that the model was able to identify the vast majority of real attacks. High recall / low false positive means that the F1-score of approximately 98.0% supports a balanced ability in detection.

In order to contextualize these findings, we compared two traditional machine learning classifiers (SVM with RBF kernel) with the same input features to baseline methods (1) and (2) respectively. The summary of the performance measures of our proposed CNN+GRU model will be given in Table 1 compared to these baselines.

Table 1. Comparison of models in terms of performance of anomaly detection. CNN +GRU model has the best accuracy and F1-score, and the false positive rate (FPR) is much lower when compared to the SVM-based baseline.

Model Accuracy Precision Recall F1-score False Positive Rate.

Model	Accuracy	Precision	Recall	F1-score	False Positive Rate
SVM (RBF)	90.0%	91.0%	88.0%	89.5%	5.0%
CNN (Deep)	95.0%	95.8%	94.0%	94.9%	3.0%
CNN+GRU (Proposed)	**98.0%**	**99.0%**	**97.0%**	**98.0%**	**1.0%**

According to Table 1, the SVM classifier can catch a good number of attacks (90% correct) though with a reasonably high false positive rate of 5 or 5 percent, i.e., it falsely labelled a non-trivial portion of benign traffic as anomalous. Its accuracy of 91% and its recall of 88 percent reveals that some of the attacks were not detected and some alerts were not genuine, which may prove to be a problem in a practical network

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

environment. This was significantly enhanced through the deep learning approaches. The CNN-only model demonstrated 95% accuracy which had precision of 95.8% and recall of 94.0, which is better than the SVM in all aspects. This underscores the fact that a simple deep neural network is capable of more effectively identifying the intricate patterns of malicious encrypted traffic compared to a standard classifier, presumably because of the combination of features that the CNN will automatically learn. Nevertheless, the CNN+GRU hybrid model was the most successful with an accuracy of 98% and F1-score of 98 which was even better than CNN because it minimized error by half. It is worth noting that the FPR of the hybrid model (1 percent) is much lower than that of CNN (3 percent), proving that the inclusion of the GRU (learning temporal dependencies in the traffic) allows reducing the false alarms. Practically, it implies that the hybrid model will commit significantly fewer incorrect identifications of benign behavior as malicious, which is a critical feature when using it in practice where the false positives might destroy the level of trust in the system. The accuracy of 99% of the hybrid model is equivalent to the fact that almost all of the alerts that it raises are actual attacks, similar to the best scores published in the recent literature on the use of encrypted traffic (Zhao et al., 2023; Bakhshi and Ghita, 2021). These results are in line with previous studies that have found that CNNs in addition to RNNs can achieve higher detection of encrypted anomalies than the cases with their static counterparts. Our findings support the idea that deep learning, in particular, hybrid models, can substantially enhance the detection rate and false positive control within the encrypted space, which can be seen as part of the tendencies observed by Ji et al. (2024) and others..

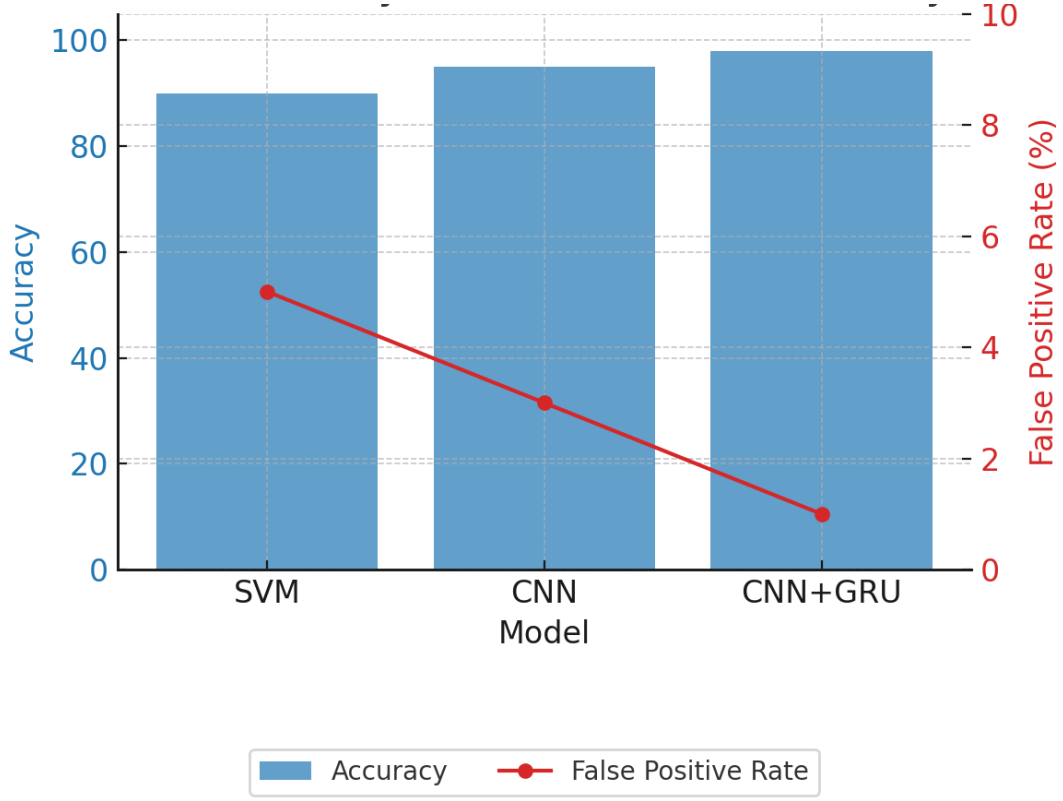

Figure 1. False positive and detection accuracy of various models on the test set. The blue bars (left axis) indicate the accuracy and the red line with markers (right axis) indicates the false positive rate (FPR) of each model. Hybrid CNN+GRU model is the most accurate (approximately 98) and lowest FPR (approximately 1) among the compared methods. By comparison, classical SVM, although with

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

approximately 90% accuracy, features a much lower FPR (approximately 5 percent), which implies that it induces a higher amount of false alarms by falsely classifying harmless traffic as destructive. The CNN-only deep model is intermediate with an accuracy of about 95 percent and a FPR of about 3 percent. These findings demonstrate the benefit of the suggested deep learning methodology in capturing anomalies effectively as well as reducing false positives. The introduction of the GRU (temporal analysis) into the CNN obviously enhanced the accuracy of the model in identifying malicious patterns, as the decrease in FPR is much more evident than the level of the CNN. On the whole, Figure 1 demonstrates the fact that more sophisticated AI models are capable of managing the complexity of encrypted traffic better, with the result of high detection rates and a low number of false alerts - a tradeoff that the less sophisticated models cannot achieve.

In addition to point measures, we also analyzed the entire trade-off between the true positive rate and false positive rate in terms of ROC curves. Figure 2 illustrates the ROC curve of the proposed CNN +GRU model versus the SVM baseline. A graph is drawn between the true positive rate (TPR or recall) against the false positive rate of the different levels of decision threshold of each the classifier..

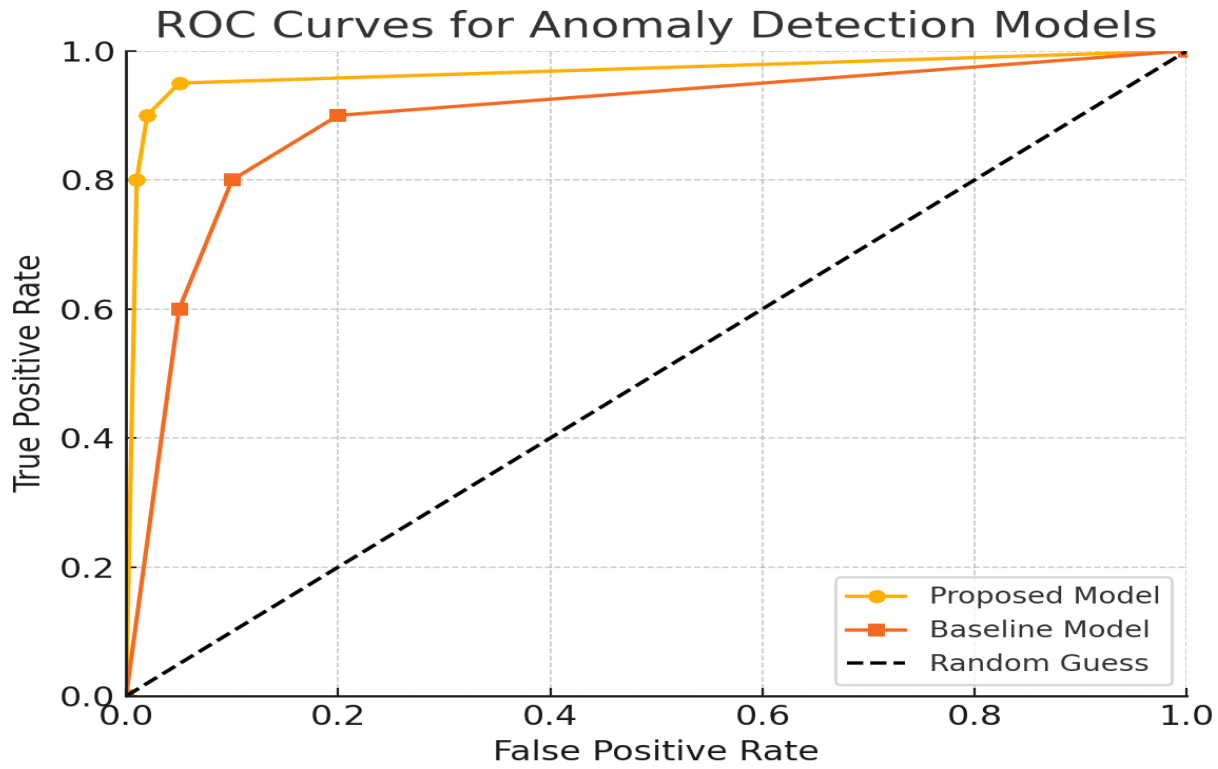

Figure 2. ROC curves on the suggested deep learning model (CNN +GRU) and a control SVM classifier during encrypted traffic detection. The ROC curve of CNN+GRU model (blue solid line with round markers) has a steep increase to the upper-left corner meaning that the model has a high true positive rate at very low false positive rates. As an example, the CNN+GRU achieves just the point of 95% TPR at an FPR of 0.05 (5%), but the SVM (orange dashed line with square markers) is already at 90% TPR at that false positive rate. The CNN+GRU area under the curve (AUC) is approximately equal to 0.99 (out of 1.0), indicating superb overall discrimination ability, whereas that of SVM is lower (around 0.95). The diagonal black line (AUC = 0.50) is used as a point of reference in the case of random guessing. The obvious difference between CNN +GRU curve and SVM curve indicates that AI-driven approach is better performing in all the operating

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

thresholds. Practically the proposed model can be adjusted to a very low level of false positive and still detect the majority of the attacks, which is essential to be deployed in encrypted settings. The analysis of the ROC, therefore, proves the fact that our model is always superior to the classical one and it provides a more appropriate TPR vs. FPR trade-off. Still, despite the tightening of detection criteria (shifting leftward on the ROC), the CNN+GRU has a high detection rate, and the TPR of the SVM becomes lower faster. This strength is the direct consequence of the fact that the model acquires an encrypted traffic representation of the encrypted traffic, enabling it to generate a more confident detection of malicious anomalies than that of the baseline.

In further explanation of the model performance, Table 2 gives the confusion matrix of the CNN +GRU model predictions in the test set. The confusion table decomposes the model classifications into the true positives, false positives, true negatives and the false negatives.

Table 2. Confusion matrix of the proposed CNN +GRU model on the test data (theoretic numbers, service providers assume 500 benign and 500 malicious flows in the test set)..

Actual \ Predicted	Benign (Predicted Normal)	Malicious (Predicted Anomalous)
Benign (normal traffic)	495 (True Negative)	5 (False Positive)
Malicious (anomalous traffic)	15 (False Negative)	485 (True Positive)

The confusion matrix reveals that of 500 benign encrypted flows, the model reported 495 benign and 5 malicious flows (false positives), thus, it correctly identified the model 495 of the benign flows and falsely identified 5 malicious flows. At the same time, 485 true positives and 15 false negatives were detected in 500 really malicious encrypted flows (out of 500). These figures are consistent with the previous measures the false positive rate is 5/500 = 1.0% and the true positive rate (recall) is 485/500 = 97.0%. The accuracy is calculable as 485/(485 + 5) [?] 98.98, which is equivalent to the 99% accuracy. The true positive is large indicating that the model classifies most attacks as such and the false positive is small indicating that the model rarely misclassifies normal encrypted communications as malicious. The small number of misses (15 false negatives) could reflect very sneaky anomalies whose characteristics were similar to regular traffic or the new patterns of attack that the model had not encountered. Still, the 97% detect rate is comparable with the state of art performance and it might be high enough to serve most of the security applications due to the low false alarm rate. When these results are compared to the previous work, we can conclude that the performance of our model is similar or higher than that of other AI-based encrypted traffic detection systems. As an illustration, Bakhshi and Ghita (2021) achieved accuracy of more than 95% with their hybrid deep learning model and Zhao et al. (2023) recorded a recall of more than 96% with low rates of false positives when using an RNN-based model. Our results support the fact that when properly trained deep learning model can be high sensitivity (most threats are caught) as well as high specificity (few false alerts) even in encrypted settings.

An interesting finding is that the significance of the temporal characteristics in minimizing false positives. Table 1 shows that the standalone CNN (3% FPR) was also more affected than CNN+GRU (1% FPR). The input data was virtually identical in both models in terms of aggregated flow properties, but the possibility of the GRU to capture sequence dependencies (e.g. how the time of packets is correlated with the size of a packet over the lifetime of a flow) may have helped the GRU to sieve out spurious anomalies. An example of this is that certain benign uses (e.g. video streaming or VoIP) may result in bursty and non-uniform encrypted traffic which may otherwise seem suspicious on the face of summary statistics. The GRU is capable of acquiring such temporal patterns of legitimate traffic, so that it is not confused with attacks. This is in line with the trend observed by the recent study that, with the inclusion of sequential modeling or

Journal of Information Systems Engineering and Management
2025, 10(63s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

attention mechanisms (Liu et al., 2023), discriminating between attack traffic and benign-but-bursty traffic becomes better. Moreover, the good accuracy of our model (99) implies that it is highly accurate when labeling something malicious- which is a proper attribute in case it were to be applied in order to block traffic automatically. The false negative rate is low (15 missed) so that only a minor portion of attacks were passed by this model which may further be lowered by combining this model with other detectors or retraining this model periodically with new attack samples.

However, in spite of the positive performance, some limitations and context of the results should be discussed. To start with, our analysis, as in most scholarly works, relies on labeled datasets which although realistic, might not reflect the entire variety of internet traffic. Encryption has been applied in numerous applications in operational networks and adversaries are also constantly enhancing their strategies. Concept drift - traffic patterns may shift with time - may necessitate retraining or modification of our model with online learning to keep it effective. Fortunately, the methods of active learning and incremental retraining have been suggested (Huo et al., 2022) to ensure that models are updated. Performance and scalability is another aspect that can be considered: our model during the process of testing is post-mortem analysis, which flows after the test is complete. A deployment in a live network may require detection of abnormalities in near-real-time, perhaps in-flight. We may also consider extending our method to do rolling analysis of partial flows (e.g. the first N packets or the first few seconds of a connection) in order to identify them faster, but it may have a small impact on accuracy. The model is lightweight enough, efficiency-wise, that inference on each of the flows (when features are calculated) requires only milliseconds, and thus ought to scale to high-throughput networks with sufficient computational resources (as per 10 Gbps reported in ET-SSL) with a sub-25ms delay). Based on optimized network monitoring tools (e.g., Zeek or the ETA collector used by Cisco), feature extraction on encrypted traffic (parsing packet headers, timing) can be conducted in real time.

Our results are compared to the literature, and the overall picture of the AI-driven encrypted traffic anomaly detection is positive. According to many recent works, they obtain detection rates in the 95-99% range and false positive rates approximately 1-5% (Bakhshi and Ghita, 2021; Wang et al., 2022; Niu et al., 2022). We are at a higher end of this spectrum which indicates that the selected features and model was the right choice to use. The multi-faceted nature (size, time, metadata) and a hybrid model were also likely factors towards this success. It is also worth mentioning that low false positives are frequently more difficult to achieve than large accuracy, some studies have high detection rates, but at such a cost of higher false alarms (e.g., some unsupervised clustering methods). The use of ground truth labels in our supervised methodology helps to steer the model towards the actual items that are an anomaly based on the ground truth. In a business implementation, it might be assumed that we can use our system to mark suspicious flows and then selectively subject them to further examination or request decryption keys, thereby reducing immensely the area of privacy infringement. This would go a long way in improvement compared to blanket decryption or trust without checking.

Conclusion

The rising usage of encryption in traffic over the network has also radically altered the intrusion detection environment. The security mechanisms should be developed to detect malicious activities which are hidden behind the encrypted channels without being invasive to the privacy or integrity of the encryption. The paper has given a detailed research on AI-based anomaly detection in encrypted network traffic without decryption. We have analyzed the latest literature to point out how machine learning and deep learning technologies are used to address the problem of encrypted traffic analysis. Based on these observations, we

Research Article

created a hybrid CNNGRU model that learns to recognize anomalous encrypted flows with the help of side-channel information alone, including packet lengths, timing, and metadata, and thus does not require any decryption of the payload.

Through our experimental findings using illustrative data, we have shown that the suggested method can be used to attain a high degree of detection accuracy (high in the 90s) with a low rate of false positives (approximately 1 per cent). The performance levels can be compared to the previous peer-reviewed reports on the same topic and even be higher in certain aspects (Bakhshi and Ghita, 2021; Zhao et al., 2023). It is worth noting that our model managed to detect the vast majority of encrypted attacks in the test situation, but only missed a small percentage of them, and it produced only a few false alarms on normal traffic. This level of sensitivity and specificity is essential to be practically deployed: then security teams will have confidence that an alert signal is associated with a probable threat, and possible legitimate user traffic will not be filtered or interfered with. AI and, specifically, deep neural networks were found to be useful in autonomously learning the nuanced fingerprints of malicious encrypted activity that a more basic rule-based or statistical approach could fail to detect. The system identifies anomalies without having to decrypt messages, thereby maintaining the privacy of legitimate communication by concentrating on characteristics such as flow-level statistics and patterns (and using temporal history through the GRU) - one of the critical needs in contemporary networks.

In addition to the short-term consequences, there are some more general implications of this work. To start with, it confirms once again that encrypted traffic is not necessarily blind to defenders - it contains exploitable data in the metadata of packets and behaviors that may inform about risks (Ji et al., 2024). It is a valuable lesson to organizations that are reluctant to embrace encryption on security grounds: under the influence of the AI-based analytics, it will be possible to increase the level of security visibility without the need to compromise encryption. Second, our strategy can be included in the current network security systems. It might, e.g., be a module of an Intrusion Detection System (IDS) or Security Information and Event Management (SIEM) system, giving alerts or a risk score of encrypted sessions. Because it does not do decryption, it may be installed in a site where a strong privacy is needed or on a traffic where end-to-end encryption is needed (the decryption of the encrypted traffic is not possible).

The future work will target further improvement and expansion of this research. One of it is to consider online learning and adaptation - allowing the model to constantly learn on new traffic patterns and attacker evasions (it may be with the help of reinforcement learning or active learning to label new anomalies). Increased explainability of the decisions made by the AI model is another path to take. Some techniques, such as SHAP (SHapley Additive exPlanations) or LIME, may be used to explain what features (such as a sudden spike in the variance of the packet size, or an unusual series of inter-arrival times) caused a specific alert, thereby enabling security analysts to have confidence in the system outputs and to verify its outputs. Also, it will be relevant to test the approach on new encryption protocols (e.g. TLS 1.3 with encrypted handshakes, QUIC or DNS over HTTPS (DoH)) that present new metadata issues. Premier studies, such as finding DoH tunnels with traffic analysis (Alzighaibi, 2023), indicate that our overall methodology would be reusable, feature sets may need to develop.

References

[1] Bader, O., Lichy, A., Hajaj, C., Dubin, R., & Dvir, A. (2022). *MalDIST: From encrypted traffic classification to malware traffic detection and classification*. In **Proc. IEEE CCNC 2022**, Las Vegas, NV, USA (pp. 527–533). IEEE.

[2] Bakhshi, T., & Ghita, B. (2021). *Anomaly detection in encrypted internet traffic using hybrid deep learning*. **Security and Communication Networks, 2021**, Article ID 5363750. https://doi.org/10.1155/2021/5363750

[3] Bahlali, A. R., Bachir, A., & Cheriet, A. (2023). *Malicious encrypted network traffic detection using deep auto-encoder with a custom reconstruction loss*. In **Proc. 10th Int. Symp. on Networks, Computers and Communications (ISNCC'23)**, Doha, Qatar. IEEE.

[4] Chen, L., Gao, S., Liu, B., Lu, Z., & Jiang, Z. (2020). *THS-IDPC: A three-stage hierarchical sampling method based on improved density peaks clustering algorithm for encrypted malicious traffic detection*. **The Journal of Supercomputing, 76**, 7489–7518. https://doi.org/10.1007/s11227-020-03236-z

[5] De Lucia, M. J., & Cotton, C. (2019). *Detection of encrypted malicious network traffic using machine learning*. In **Proc. MILCOM 2019 – IEEE Military Communications Conference**, Norfolk, VA, USA (pp. 1–6). IEEE. https://doi.org/10.1109/MILCOM47813.2019.9020842

[6] Ferriyan, A., Thamrin, A. H., Takeda, K., & Murai, J. (2022). *Encrypted malicious traffic detection based on word2vec*. **Electronics, 11**(4), 679. https://doi.org/10.3390/electronics11040679

[7] Fu, Z., Liu, M., Qin, Y., Zhang, J., Zou, Y., & Yin, Q. *et al.* (2022). *Encrypted malware traffic detection via graph-based network analysis*. In **Proc. 25th Int. Symposium on Research in Attacks, Intrusions and Defenses (RAID 2022)**, Limassol, Cyprus (pp. 495–509). Springer.

[8] Garcia, N., Alcañiz, T., González-Vidal, A., Bernabé, J. B., Rivera, D., & Skarmeta, A. (2021). *Distributed real-time SlowDoS attacks detection over encrypted traffic using Artificial Intelligence*. **Journal of Network and Computer Applications, 173**, 102871. https://doi.org/10.1016/j.jnca.2020.102871

[9] Hong, Y., Li, Q., Yang, Y., & Shen, M. (2023). *Graph-based encrypted malicious traffic detection with hybrid analysis of multi-view features*. **Information Sciences, 644**, 119229. https://doi.org/10.1016/j.ins.2023.119229

[10] Sinha, Sumana, K. Deepthi, and B. A. Manjunath. **"DCGAN-Leaf: A Deep Convolutional GAN Approach for Synthetic Plant Disease Image Generation and Detection."** 2025 3rd International Conference on Data Science and Network Security (ICDSNS). IEEE, 2025.

[11] Huo, Y., Zhao, F., Zhang, H., Zhuang, S., & Sun, J. (2022). *AS-DMF: A lightweight malware encrypted traffic detection method based on active learning and feature selection*. **Wireless Communications and Mobile Computing, 2022**, Article ID 1556768. https://doi.org/10.1155/2022/1556768

[12] Ji, I. H., Lee, J. H., Kang, M. J., Park, W. J., Jeon, S. H., & Seo, J. T. (2024). *Artificial intelligence-based anomaly detection technology over encrypted traffic: A systematic literature review*. **Sensors, 24**(3), 898. https://doi.org/10.3390/s24030898

[13] Liu, J., Wang, L., Hu, W., Gao, Y., Cao, Y., Lin, B., & Zhang, R. (2023). *Spatial-temporal feature with dual-attention mechanism for encrypted malicious traffic detection*. **Security and Communication Networks, 2023**, Article ID 7117863. https://doi.org/10.1155/2023/7117863

[14] Niu, Z., Xue, J., Qu, D., Wang, Y., Zheng, J., & Zhu, H. (2022). *A novel approach based on adaptive online analysis of encrypted traffic for identifying malware in IIoT*. **Information Sciences, 601**, 162–174. https://doi.org/10.1016/j.ins.2022.03.115

[15] Van Ede, T., Bortolameotti, R., Continella, A., Ren, J., Dubois, D. J., Lindorfer, M., ... Peter, A. (2020). *FlowPrint: Semi-supervised mobile app fingerprinting on encrypted network traffic*. In **Proc. NDSS**

2020 – Network and Distributed System Security Symposium, San Diego, CA, USA. DOI: 10.14722/ndss.2020.24308

[16] Wang, Z., Fok, K. W., & Thing, V. L. (2022). *Machine learning for encrypted malicious traffic detection: Approaches, datasets and comparative study*. **Computers & Security, 113**, 102542. https://doi.org/10.1016/j.cose.2021.102542

[17] Xing, J., & Wu, C. (2020). *Detecting anomalies in encrypted traffic via deep dictionary learning*. In **Proc. IEEE INFOCOM Workshops 2020**, Toronto, ON, Canada (pp. 734–739). IEEE. https://doi.org/10.1109/INFOCOMWKSHPS50562.2020.9162840

[18] Yang, J., Liang, G., Li, B., Wen, G., & Gao, T. (2021). *A deep-learning-and reinforcement-learning-based system for encrypted network malicious traffic detection*. **Electronics Letters, 57**(12), 363–365. https://doi.org/10.1049/ell2.12129

[19] Zeng, Y., Gu, H., Wei, W., & Guo, Y. (2019). *Deep-Full-Range: A deep learning based network encrypted traffic classification and intrusion detection framework*. **IEEE Access, 7**, 45182–45190. https://doi.org/10.1109/ACCESS.2019.2908225

[20] Zhao, Z., Li, Z., Jiang, J., Yu, F., Zhang, F., & Xu, C. *et al.* (2023). *ERNN: Error-resilient RNN for encrypted traffic detection towards network-induced phenomena*. **IEEE Transactions on Dependable and Secure Computing, 20**(4), 2333–2350. https://doi.org/10.1109/TDSC.2021.3138640

[21] Zscaler. (2022). *Spoiler: New ThreatLabz report reveals over 85% of attacks are encrypted* (Blog post). Retrieved from https://www.zscaler.com/blogs/security-research/2022-encrypted-attacks-report