Research paper

Title: Detecting Data Exfiltration Anomalies in Academic Networks Using the Isolation Forest Algorithm

Author: Mike Kimutai Arusei[1] Dr. Stephen Njenga[2] (Co Author)

Affiliations: School of Technology, KCA University, Nairobi, Kenya

Emails: [1] aruseike@gmail.com [2] njengast@gmail.com

Abstract

Academic networks face increased risks of data exfiltration due to sensitive personal information and research data. Traditional supervised detection models rely on labeled datasets which are often unavailable in resource constrained institutions. This study investigates the applicability of the unsupervised Isolation Forest algorithm for detecting anomalous network traffic indicative of data exfiltration. The research utilized the CICIDS2017 dataset focusing on the *Thursday-WorkingHours-Afternoon-Infiltration* subset. Key features including *Flow Duration*, *Total Fwd Packets*, *Flow Bytes/s*, *Flow IAT Mean*, and *Destination Port* were preprocessed and normalized for modeling. The model achieved a precision of 1.00, recall of 0.99 and F1-score of 1.00 for anomalous traffic detection successfully identifying approximately 4.8% of flows as anomalous. Comparative analysis with previous methods, including supervised Random Forest and SVM demonstrated that Isolation Forest offers competitive accuracy with lower computational overhead and does not require labeled data. The findings highlight the algorithm's suitability for academic network monitoring, providing an effective early warning mechanism while emphasizing the importance of threshold tuning to reduce false positives.

Keywords: Anomaly Detection, Data Exfiltration Machine Learning, Isolation Forest, Academic Networks

1. Introduction

Academic institutions increasingly rely on digital networks for research, administration and communication making them vulnerable to cyberattacks including data exfiltration. Detecting such anomalies is challenging due to the scarcity of labeled threat data and resource constraints in institutional IT environments. Traditional supervised intrusion detection systems often require extensive labeled datasets and computational resources. This study investigates the use of the Isolation Forest algorithm, an unsupervised machine learning technique to detect anomalous network traffic indicative of data exfiltration within an academic network context (Liu, Ting, & Zhou, 2008). The research focuses on identifying key traffic features that signal potential threats and

evaluates the effectiveness of the proposed detection model.

2. Literature Review

Previous research on network anomaly detection has primarily focused on supervised learning methods such as Random Forest and Support Vector Machines (Ahmed, Mahmood, & Hu, 2016); (Ring, Wunderlich, Scheuring, Landes, & Hotho, 2019). While effective on labelled datasets, these models face challenges in academic environments where labeled attack data is scarce.

Unsupervised learning methods like Isolation Forest (Liu, Ting, & Zhou, 2008) and clustering based models offer the ability to detect rare anomalies without requiring prior labelling. Studies have demonstrated their effectiveness in handling high-dimensional, imbalanced network data while maintaining low computational costs. Comparisons with K-Means and autoencoder-based unsupervised methods indicate that Isolation Forest algorithm achieves superior recall rates, making it suitable for detecting critical anomalies such as data exfiltration in resource-constrained settings (Seo, Kim, & Lee, 2023); (Omar & Tariq, 2022)

The research gap is that there is limited application of unsupervised anomaly detection models in academic network traffic specifically targeting data exfiltration motivating this study.

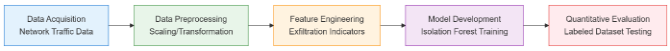
3. Methodology

This study employed an experimental quantitative approach to develop and evaluate an unsupervised machine learning algorithm, Isolation Forest for detecting data exfiltration threats in academic network environments. The methodology incorporated data acquisition, pre-processing, feature selection, model development and performance evaluation to provide a comprehensive and reproducible framework for anomaly detection.

3.1 Research Design

A quantitative experimental design was adopted complemented by exploratory data analysis (EDA). The quantitative approach provided a structured framework for collecting, processing, and analyzing network traffic data using numerical metrics as illustrated in ***Figure 1***. The experimental aspect facilitated controlled testing of the Isolation Forest algorithm using a realistic network traffic dataset. EDA was conducted prior to model development to identify patterns, correlations, and potential indicators of anomalous behaviour ensuring that the model design was aligned with the characteristics of academic network traffic.

Figure 1: Model Design Phases



3.2 Data Acquisition

Secondary data were obtained from the CICIDS2017 dataset provided by the Canadian Institute for Cybersecurity, specifically the *Thursday-WorkingHours-Afternoon-Infiltration* subset (CIC, 2024). This dataset was selected due to its realistic representation of modern network traffic, inclusion of labelled benign and malicious flows and simulated exfiltration events relevant to academic institutions.

As illustrated in ***Figure 2***, the dataset contains 288,602 network flow instances and 79 features, capturing both normal and anomalous activities. Key features relevant for anomaly detection

Figure 2: Thursday-WorkingHours-Afternoon-Infiltration subset dataset (CIC, 2024)

	Destination Port	Flow Duration	Total Fwd Packets	Total Backward Packets	Total Length of Fwd Packets	Total Length of Bwd Packets	Fwd Packet Length Max	Fwd Packet Length Min
288597	80	590930	2	0	0	0	0	0
288598	80	1187988	2	0	0	0	0	0
288599	80	10	1	9	6	54	6	6
288600	138	19	10	0	2370	0	237	237
288601	80	4751966	2	0	0	0	0	0

5 rows × 79 columns

(288602, 79)

3.3 Data Sampling & Preprocessing

A purposive sampling strategy was employed to select the subset of network traffic that most closely aligns with the study objectives. The chosen subset included typical daytime academic network activities with both normal and malicious flows ensuring experimental control and relevance to real world academic environments.

The dataset underwent systematic pre-processing, including data cleaning, normalization and feature selection. Duplicate records, null entries and missing values were addressed and numerical features were normalized to prevent scale bias.

Exploratory Data Analysis (EDA) involved computation of descriptive statistics, visualization of feature distributions and examination of correlations to identify patterns indicative of data exfiltration.

3.4 Model Development

The Isolation Forest algorithm was implemented using Python and the scikit-learn library with hyper parameters as tabulated in ***Table 1***. This algorithm isolates anomalies by constructing random decision trees with observations requiring fewer splits considered anomalous.

Table 1: Isolation Forest Key Parameters

Parameter	Value	Description
n_estimators	100	Number of isolation trees
contamination	0.01	Expected proportion of anomalies
max_samples	auto	Number of samples per tree
random_state	42	For reproducibility
bootstrap	True	Samples drawn with replacement

The model was trained using normalized data, enabling it to learn patterns of normal traffic and assign anomaly scores for detection of suspicious flows.

3.5 Model Evaluation

Model performance was evaluated using ground-truth labels from the dataset. Standard classification metrics including precision, recall and F1-score were computed to assess the algorithm's ability to correctly identify anomalies.

Evaluation included generation of a confusion matrix and visualization of anomaly-score distributions to assess accuracy and understand model behavior across different traffic patterns.

3.6 Ethical Considerations and Limitations

This study used anonymized publicly available datasets ensuring compliance with the Kenyan Data Protection Act (2019) and institutional ICT research policies (ODPC, 2019).

Limitations include reliance on simulated datasets which may not fully reflect real-world complexities, and the dependency of model performance on dataset characteristics and hyperparameter tuning. These limitations were mitigated through careful pre-processing, parameter optimization and validation.

4. Results and Discussion

4.1 Model Results

The Isolation Forest model was trained and validated on institutional network traffic consisting of normal and potentially anomalous data flows. After pre-processing and feature extraction, the model identified approximately 4.8% of all network flows as anomalous patterns consistent with data-exfiltration activities such as unusually large outbound transfers and abnormal flow durations.

4.2 Model Performance

The model achieved strong performance across all evaluation metrics as summarized in ***Figure 3***

Figure 3: Model Performance Metrics

Classification Report:

	precision	recall	f1-score	support
BENIGN	0.00	0.31	0.01	36
ANOMALY	1.00	0.99	1.00	288566
accuracy			0.99	288602
macro avg	0.50	0.65	0.50	288602
weighted avg	1.00	0.99	0.99	288602

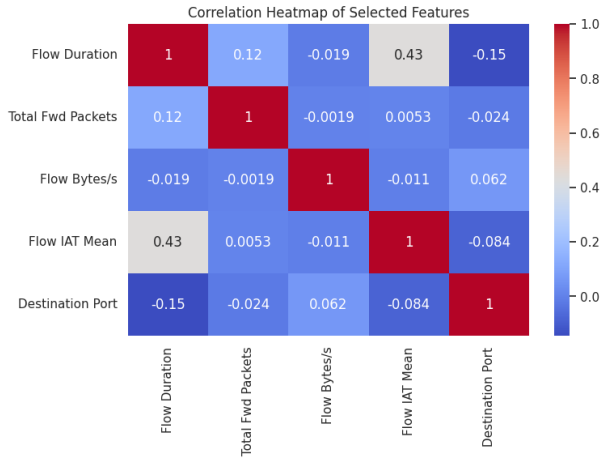
The precision score of **1.00** indicates that the model produced highly reliable alerts, with the vast majority of detections corresponding to genuine anomalies. The recall score of **0.99** shows that the model successfully captured most suspicious traffic flows demonstrating strong sensitivity without excessive false negatives. The

balanced F1-Score reflects model's overall robustness.

4.3 Feature Influence in Anomaly Detection

Analysis of network traffic features indicated that *Flow Duration* and *Total Bytes Sent* were the most influential predictors of anomalous behaviour as illustrated in ***Figure 4***. Abnormal increases in these features strongly correlated with potential exfiltration attempts. The Isolation Forest model computes anomaly scores based on the degree of isolation of a data point from the rest of the dataset, making features with large deviations particularly significant for anomaly detection.

Figure 4. Feature Influence on Anomaly Detection

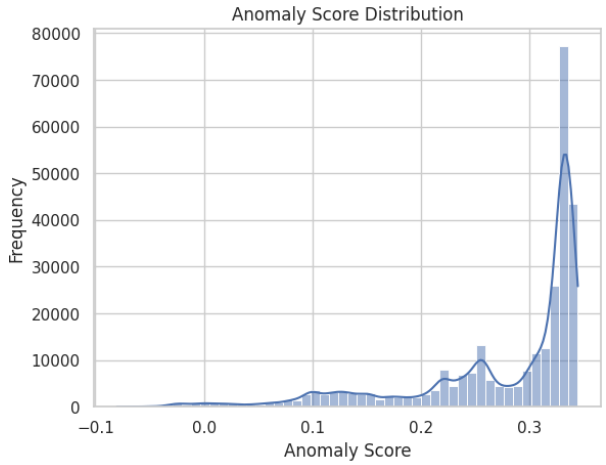


The results emphasize that monitoring these features is critical for early detection of unusual network activity in academic environments.

4.4 Rare and Silent Exfiltration Patterns detection

As illustrated in ***Figure 5***, the model effectively identified subtle, silent and low volume exfiltration patterns that often evade traditional signature-based and supervised detection methods. Because the Isolation Forest operates unsupervised, it is well-suited for imbalanced network traffic where anomalous events are rare and different.

Figure 5. Distribution of Anomaly Scores Highlighting Rare Exfiltration Patterns



Most data points had high anomaly scores consistent with normal traffic, while a small subset exhibited low scores, indicating potential anomalies. This distribution confirms the model's capability to capture rare and silent exfiltration attempts that could otherwise go undetected.

4.5 Comparison with Existing Methods

The proposed Isolation Forest model was compared to recent supervised and unsupervised methods reported in literature (Seo, Kim, & Lee, 2023); (Omar & Tariq, 2022)

Key observations as Tabulated in ***Table 2*** include:

- Higher precision and recall than Random Forest (RF) and Support Vector Machines (SVM).
- More stable performance compared to K-Means clustering.
- Lower computational overhead supporting near-real-time deployment in resource-constrained academic networks.

Table 2: Performance Comparison with RF, SVM, and K-Means

Method	Precision	Recall	F1-score	Description
Isolation Forest	1.00	0.99	1.00	Best for imbalanced data
Random Forest	0.92	0.85	0.88	Requires labeled data
SVM	0.90	0.83	0.86	Sensitive to class imbalance
K-Means	0.85	0.78	0.81	Less sensitive to rare anomalies

These comparisons demonstrate Isolation Forest's adaptability particularly in environments that has imbalanced datasets and limited labeled traffic which are typical in academic networks.

4.6 Discussion

The results indicate that the Isolation Forest algorithm is highly effective in identifying true anomalies particularly rare or subtle exfiltration attempts. High precision and recall in the anomaly class ensure minimal missed threats, while moderate misclassification of normal traffic is an expected limitation in unsupervised detection.

The findings are consistent with Liu, Ting, and Zhou (Liu, Ting, & Zhou, 2008), who reported that Isolation Forest efficiently handles high-dimensional and imbalanced datasets. Compared to supervised methods which rely heavily on labeled training data, the Isolation Forest model demonstrates superior flexibility and efficiency for practical academic network deployment. (Ahmed, Mahmood, & Hu, 2016), (Ring, Wunderlich, Scheuring, Landes, & Hotho, 2019)

4.7 Practical Implications

The findings demonstrate that the Isolation Forest model provides an effective tool for enhancing cybersecurity in academic institutions. Its high precision and recall enable accurate detection of anomalous network traffic, supporting threat prioritization and early identification of potential data exfiltration or infiltration. The model's ability to detect rare and subtle anomalies facilitates proactive and predictive monitoring, reducing reliance on predefined attack signatures.

Flagged traffic insights can inform cybersecurity awareness programs, guiding staff and students to recognize and respond to unusual behavior. Although misclassification of benign traffic may produce false alerts, careful threshold tuning and integration with automated triage can mitigate

alert fatigue. Effective deployment requires sufficient technical and human resources for model retraining, monitoring, and alignment with incident response protocols.

Temporal patterns of anomalies provide actionable guidance for policy adjustments, such as enhanced access controls during high risk periods. Additionally, machine learning-based anomaly detection fosters inter-institutional collaboration through shared research, joint response mechanisms and capacity-building initiatives.

Overall, the model offers a practical, high-performing solution for academic network security supporting both immediate operational interventions and long-term strategic planning.

5. Conclusion

This study investigated the use of the Isolation Forest algorithm for detecting anomalous network traffic in academic environments. The key conclusions included: -

a) ***Effectiveness of Unsupervised Anomaly Detection*** - Isolation Forest provides a practical cybersecurity approach by detecting deviations from normal traffic without relying on predefined attack signatures, enabling the identification of rare, subtle or previously unknown exfiltration and infiltration attempts (Liu, Ting, & Zhou, 2008)

b) ***Suitability for Resource-Constrained Environments*** **-** Due to its low computational requirements and unsupervised nature, the algorithm is well-suited for academic institutions with limited technical expertise and infrastructure (Seo, Kim, & Lee, 2023)

c) ***Critical Role of Feature Selection*** - Accurate anomaly detection is highly dependent on the selection of informative network features, such as flow duration, total bytes sent and inter-arrival times. Proper feature selection enhances model performance by improving discrimination between normal and anomalous traffic.

d) ***Broader Applicability*** - The model's demonstrated performance and lightweight implementation suggest potential adaptation in other sectors with constrained resources, including government, healthcare and small enterprise networks (Omar & Tariq, 2022)

6. Recommendations

6.1 Practical and Policy Recommendations

a) ***Deployment of Anomaly-Based Detection*** **-** Academic institutions should integrate lightweight unsupervised models like Isolation Forest into their cybersecurity strategy complementing existing firewalls and antivirus systems.

b) ***Investment in Network Monitoring Infrastructure*** - Institutions should

implement systems for logging and storing network traffic forming the basis for machine learning-driven security and continuous model improvement.

c) ***Capacity Building for ICT Staff*** **-** Technical teams should be trained not only in model deployment but also in interpreting outputs and responding to alerts effectively.

d) ***Support for AI-Driven Cybersecurity Research*** **-** National and institutional stakeholders should promote AI-based cybersecurity research to foster collaboration between universities, technology companies and policy bodies.

6.2 Recommendations for Future Research

a) ***Development of Local Datasets*** - Collection of real network traffic from Kenyan academic institutions would improve the representativeness and accuracy of detection models.

b) ***Real-Time Deployment*** **-** Future studies should explore live implementation of Isolation Forest integrated with automated alert and response mechanisms.

c) ***Comparative Evaluation*** **-** Researchers should compare the performance of Isolation Forest with deep learning or hybrid models e.g. combining clustering or autoencoders to determine potential improvements in detection accuracy.

d) ***Resilience Against Evasive Techniques*** **-** Investigate model robustness under sophisticated attack scenarios to enhance practical applicability in dynamic and adversarial network environments.

7. Limitations and Future Work

Despite achieving its objectives, the study has several limitations:

a) ***Data Representativeness*** **-** The analysis relied on secondary and simulated datasets (CICIDS2017) rather than real-time institutional traffic potentially limiting representation of dynamic or encrypted network behaviour.

b) ***Controlled Evaluation Environment*** **-** Model performance was assessed under experimental conditions thus effectiveness in diverse real-world networks may vary with size, configuration and user behaviour.

c) ***Scope of Model Exploration*** **-** Resource constraints prevented the evaluation of hybrid or ensemble detection methods which might further improve accuracy.

d) ***Ethical and Privacy Constraints*** **-** Live network traffic collection was not feasible, highlighting the need for robust anonymization and consent mechanisms in future research.

References

Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. . *Journal of Network and Computer Applications, 60, 19–31. https://doi.org/10.1016/j.jnca.2015.11.016*.

CIC. (2024). *Intrusion detection evaluation dataset (CIC-IDS2017)*. Retrieved June 21, 2025, from https://www.unb.ca/cic/datasets/ids-2017.html

Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008). Isolation Forest. 2008 Eighth IEEE International Conference on Data Mining, 413–422. https://doi.org/10.1109/ICDM.2008.17.

ODPC. (2019). *Data protection Laws - Kenya Regulatory Framework*. Retrieved June 21, 2025, from https://www.odpc.go.ke/wp-content/uploads/2024/02/TheDataProtectionAct__No24of2019.pdf

Omar, A., & Tariq, M. (2022). Evaluating machine learning models for network anomaly detection in higher education institutions. International Journal of Information Security, 21(4), https://doi.org/10.1007/s10207-022-00619-8. 385–402.

Ring, M., Wunderlich, S., Scheuring, D., Landes, D., & Hotho, A. (2019). A survey of network-based intrusion detection data sets. Computers & Security, 86, 147–167. https://doi.org/10.1016/j.cose.2019.06.005.

Seo, J., Kim, H., & Lee, S. (2023). Comparative study of anomaly detection techniques in academic networks. *Journal of Cybersecurity Research, 7(2),* , 45–62.

Acknowledgment

The author thanks KCA University, School of Technology for academic guidance and infrastructure support during this research. Appreciation is extended to faculty mentors and peers for their constructive feedback on model design and validation.

Author Biography

Mike Kimutai Arusei is a postgraduate researcher at the School of Technology, KCA University, Nairobi, Kenya**.** His research interests include machine learning for cybersecurity, network-traffic anomaly detection, and ICT governance. He is passionate about applying artificial intelligence to strengthen information-security frameworks in higher-education institutions.

License Notice