AN ISOLATION FOREST MODEL FOR ANOMALY DETECTION OF DATA EXFILTRATION IN NETWORK TRAFFIC CASE STUDY – ACADEMIC INSTITUTIONS NETWORK ENVIRONMENTS

BY

MIKE K. M. ARUSEI

MASTER OF SCIENCE IN DATA COMMUNICATION

KCA UNIVERSITY

2025

AN ISOLATION FOREST MODEL FOR ANOMALY DETECTION OF DATA EXFILTRATION IN NETWORK TRAFFIC

CASE STUDY – ACADEMIC INSTITUTIONS NETWORK ENVIRONMENTS

BY

MIKE K. M. ARUSEI

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF SCIENCE IN DATA COMMUNICATION IN THE SCHOOL OF TECHNOLOGY AT KCA UNIVERSITY

OCTOBER, 2025

DECLARATION

I declare that this dissertation is my original work and has not been previously published or submitted elsewhere for award of a degree. I also declare that this contains no material written or published by other people except where due reference is made and author duly acknowledged.

Student Name: ***Mike Arusei*** Reg. No: ***14/04172***

Signature:  Date: ***20/10/2025***

I do hereby confirm that I have examined the master's dissertation of Mike Arusei and have certified that all revisions that the dissertation panel and examiners recommended have been adequately addressed.

Sign:  Date: …***21/10/2025***…….

Dr. Stephen Njenga

Dissertation Supervisor

ABSTRACT

Academic institutions are increasingly relying on interconnected networked and to manage critical data and services making them vulnerable to cybersecurity threats such as data exfiltration. Traditional security infrastructures often fail to detect these emerging threats, especially within resource constrained academic environments lacking sufficient expertise. To address this, an anomaly detection model was designed using unsupervised Isolation Forest algorithm, which analyzes key network features identify abnormal outbound traffic indicative of data breaches. The model was evaluated on the CICIDS2017 dataset, focusing on real-world infiltration scenarios, with a case study of academic institutions to ensure contextual relevance. Using recall, precision, and F1-score metrics, the model demonstrated effective detection capabilities. Its significance lies in providing a scalable and practical network security solution for academic institutions, supporting compliance with data protection regulations. However, limitations include reliance on the representativeness of the dataset and adaptability to emerging attack patterns. Future work should include exploration of continuous model refinement and integration with broader security frameworks.

ACKNOWLEDGMENT

First and foremost, I would like to thank God for giving me health, guidance and strength to get me through this Master of Science (MSc) journey especially when things got tough

I am truly grateful to my research supervisor Dr. Stephen Njenga from the School of Technology, KCA University, for his guidance and support throughout the whole process of my research work and writing this dissertation. Your encouragement was truly motivating and really kept me going.

I also thank all my lecturers at KCA University, School of Technology and my fellow course mates as your friendship and motivation made this journey easier and more enjoyable

TABLE OF CONTENTS

DEDICATION

I wish to dedicate this work to my wonderful parents David Arusei and Susan Arusei. Thank you as you have always been there for me and believing in me. To my dear wife Triza Kirwa, your endless encouragement and support have carried me throughout every challenge. To all my siblings, thank you for your presence and for encouraging me on every step of this work. And finally, to my beautiful daughters Tiffany and Talia, you are my greatest inspiration and joy and I thank you. I am truly grateful for your prayers and encouragement throughout this important academic journey as this accomplishment belongs to all of us.

LIST OF FIGURES

LIST OF TABLES

ACRONYMS AND ABBREVIATIONS

CICIDS2017 - Canadian Institute for Cybersecurity Intrusion Detection System 2017

KE-CIRT - Kenya Computer Incident Response Team

DoS - Denial of Service

DDoS - Distributed Denial of Service

AI - Artificial Intelligence

CSV - Comma-Separated Values

DNNs - Deep Neural Networks

SVM – Support Vector Machines

IQR – Interquartile Range

CIA Triad - Confidentiality-Integrity-Availability Triad

BYOD - Bring Your Own Device

C2 - Command and Control

IP - Internet Protocol

HTTPS - Hypertext Transfer Protocol Secure

Colab - Google Colaboratory

IAT - Inter-Arrival Time

DNS - Domain Name System

EDA - Exploratory Data Analysis

FTP - File Transfer Protocol

ICT - Information and Communication Technology

iForest - Isolation Forest

LMS - Learning Management System

NaN - Not a Number

PCA - Principal Component Analysis

NumPy - Numerical Python

ODPC - Office of the Data Protection Commissioner

Pandas - Python Data Analysis Library

ML – Machine Learning

OPERATIONAL DEFINITIONS OF TERMS

Anomaly Detection – This refer to technique used to identify patterns in data that do not conform to expected behaviour.

Data Exfiltration – This refers to unauthorized transfer or leakage of data from a computer or network to an external location.

Isolation Forest Algorithm – This is an unsupervised machine learning algorithm that detects anomalies by randomly isolating observations in a dataset.

Network Traffic – This refers to the flow of data packets over a network including both inbound and outbound communication between devices**.**

Outbound Traffic – This refers to data sent from an internal network to an external destination which can include normal use or suspicious activity like data exfiltration.

Contamination Rate – This is a parameter in anomaly detection algorithms that estimates the expected proportion of anomalies in the dataset.

Kenya Data Protection Act (2019) – This is a national law enacted to regulate the collection, processing, storage, and transfer of personal data in Kenya.

Unsupervised Learning – This is a type of machine learning where the model learns patterns from data without labelled outcomes or predefined categories**.**

Exploratory Data Analysis – This refers to the initial data examination in order to identify trends, outliers and patterns to inform model development.

Feature Selection – This refers to choose of the most relevant data attributes to be used in model training.

False Positives – This refers to the normal traffic that is incorrectly flagged as anomalous

False Negatives - This refers to instances where actual network anomalies are misclassified as normal traffic

Benign Traffic – This refers to normal non-malicious network traffic

Precision – This refers to the proportion of correctly identified anomalies

Recall – This refers to the proportion of actual anomalies correctly identified

Labelled Data – This refers to data that has predefined labels indicating whether an instance is malicious or normal

False Alarm - This refers to a situation where there are excessive false positives that potentially can cause missing of real threats.

CHAPTER ONE

INTRODUCTION

1.1 Background of the study

Academic institutions are increasingly embracing digital platforms and networks to support students teaching, research activities and administrative operations. These networks play a crucial role in the daily operations of universities and colleges from storing valuable research data to managing student records. However, as they digitise their operations, they expose themselves to cybersecurity threats that if not detected or prevented can lead to loss of intellectual property, identity theft, reputational damage and financial fraud.

One of the major network security concern and often overlooked threat is data exfiltration. Data exfiltration refers to unauthorized and silent transfer of sensitive information from institutional systems to external destinations. Data exfiltration can go unnoticed for extended periods allowing attackers to steal large volumes of institutions data silently unlike denial-of-service (DoS) attacks or ransomware that causes immediate disruptions. The information that is often targeted includes research findings, student records, examination materials and financial data (Cybersecurity Ventures, 2023)

Many academic institutions lack sufficient skilled personnel, advanced network security tools and robust systems that can detect silent breaches such as data exfiltration. In 2021, University of Nairobi experienced ransomware attack that led to the exposure of confidential personal and research data an incident that clearly highlighted the growing cybersecurity threats and vulnerabilities in the education network systems (Cybersecurity and Response Centre Kenya., 2022). Recently, the Kenya Computer Incident Response Team (KE-

CIRT) pointed out the rising trend of undetected data leaks in educational institutions often due to malware, poorly monitored networks and phishing attacks (KE-CIRT, 2024)

This study explores application of Isolation Forest algorithm in detecting data exfiltration in networks with the aim of developing intelligent anomaly detection model that will enhance threat detection and thus protection of sensitive institution data will be achieved

1.2 Problem Statement

Many academic institutions especially the ones with limited skilled personnel and resources remains vulnerable to data exfiltration attacks despite using the existing traditional network security tools such as Snort and Suricata. These tools often rely on known attack signatures and often miss sophisticated threats that disguise themselves as normal traffic such as reverse shell connections and command-and-control attacks.

Supervised machine learning methods on the other hand require expensive labelled data that often struggle to detect new unseen attacks while the existing unsupervised approaches like autoencoders and clustering demand heavy computational power and have difficulty spotting subtle anomalies. This leaves academic networks at risk of unauthorized data breaches thus putting sensitive data and information at risk of unauthorised transfers and increasing the chances of failing to comply with data protection laws.

To address this gap, the Isolation Forest algorithm, an unsupervised machine model was used to develop an anomaly detection model as it works by isolating rare and distinct data points rather than modelling normal behaviour thus making it more effective in the detection of hidden threats in high dimensional network traffic.

1.3 Objectives

1.3.1 Main Objective

To develop a model to detect data exfiltration anomalies in network traffic using the Isolation Forest algorithm with focus on its relevance to academic institutions networks

1.3.2 Specific Objectives

a) To investigate network traffic features that indicates possible data exfiltration.
b) To develop a model to detect data exfiltration anomalies in network traffic data using Isolation Forest algorithm
c) To evaluate the performance of the developed model using standard metrics

1.4 Research Questions

a) What network traffic features are most relevant for identifying potential data exfiltration activities?
b) How effective is the Isolation Forest algorithm in detecting anomalous patterns associated with data exfiltration in network traffic?
c) How does the performance of the Isolation Forest model compare using standard evaluation metrics and accuracy?

1.5 Significance of the Study

Academic institutions are increasingly relying on digital platforms especially in teaching, research and administration tasks. Data breaches risks have become a major concern especially in institution that have limited cybersecurity measures and infrastructure. The study addressed this concern by the development of a simple cost-effective anomaly detection model using Isolation Forest algorithm. Academic institutions will thus have a practical approach in

monitoring and securing network traffic without the need for high end expensive computational resources. This research presented security solution to enhance early detection of potential data exfiltration attempts of sensitive data and thus supporting compliance with data governance frameworks on safeguarding sensitive information such as Kenya Data Protection Act (ODPC, 2019)

1.6 Motivation of the Study

The motivation of this study was the need for an urgent, effective and practical solution in academic institutions networks to enhance detection of data exfiltration attacks without the need of extensive resources and labels. By utilizing Isolation Forest algorithm, this research project sought to develop a scalable and lightweight anomaly detection model that is capable of identifying unusual patterns. This will thus support academic institutions strengthen cybersecurity capabilities that are cost effective and technically feasible.

1.7 Scope of the Study

The study focused on the development of a model to detect data exfiltration anomalies in academic network traffic using the Isolation Forest algorithm. Although many other cyber threats exist, this study did not focus on other attacks such as ransomware, denial-of-service or phishing attacks. The model was built and evaluated using secondary data from the CICIDS2017 dataset (CIC, 2024). The findings were interpreted in the context of academic networks.

CHAPTER TWO

LITERATURE REVIEW

2.0 Introduction

This chapter reviews key literature on network traffic anomaly detection and focus is on the emerging data exfiltration challenges. Limitations of the traditional intrusion detection systems is clearly highlighted and the growing relevance of unsupervised machine learning algorithm especially the Isolation Forest (iForest) in identifying threats is analysed. Although Isolation Forest algorithm has significantly been applied in other sectors, its application in academic sector is limited. This is the gap that forms the foundation of this research study.

2.1 Anomaly detection Theoretical and Conceptual Foundations

Anomaly detection study in network traffic is built on the idea that unusual and anomalous patterns tend to be different from normal traffic. These anomaly differences are always measured and modelled using a range of various approaches as described below: -

2.1.1 Tree-based separation approach

Tree-based separation approach especially the Isolation Forest algorithm that has gained attention due to its ability to identify network anomalies by splitting data systematically. As illustrated in Figure 1 below, anomalous points are easier to isolate thus only requiring fewer partitions than typical observations. This makes tree-based separation approach both efficient and scalable (Liao, 2023)

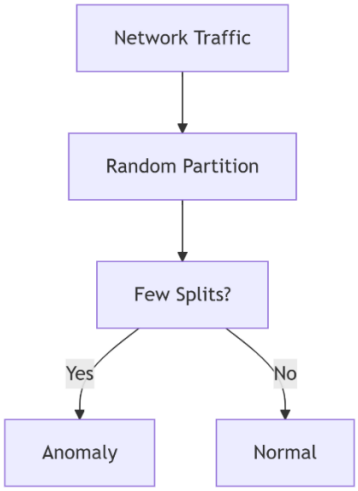


FIGURE 1

Tree-Based Separation (Isolation Forest) Approach

2.1.2 Information theory approach

Another perspective comes from information theory approach where measures such as unpredictability are used to capture irregularities in network traffic behaviour. This method detects unusual activity and also helps highlight features that carry the most diagnostic value as illustrated in Figure 2 below thus improving the precision of anomaly detection (Liu, Ting, & Zhou, 2008)

FIGURE 2

Information Theory Approach

2.1.3 Graph based methods

The most recent research is moving beyond looking at individual data points to looking at relationships. A good example is graph based methods where the connections between different network entities are considered as illustrated in Figure 3 below. This allows anomalies

to be spotted in the patterns of the graph-based connections rather than in the isolated metrics (Ma, et al., 2024).

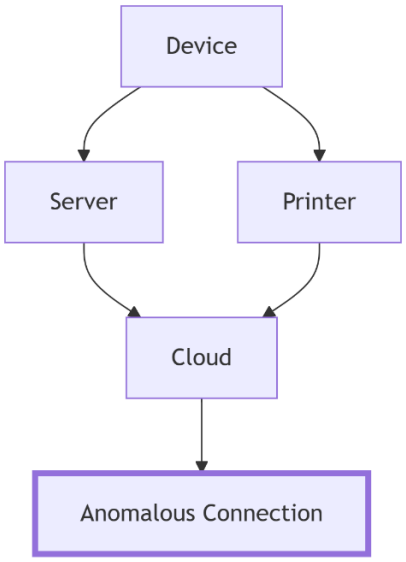


FIGURE 3

Graph-Based Methods

2.1.4 Ensemble learning frameworks

Ensemble learning frameworks as illustrated in Figure 4 below combines multiple detection models to balance the weaknesses of individual techniques thus creating systems that can easily adapt to new and emerging threats (Zhang, Chen, Wang, & Li, 2023)

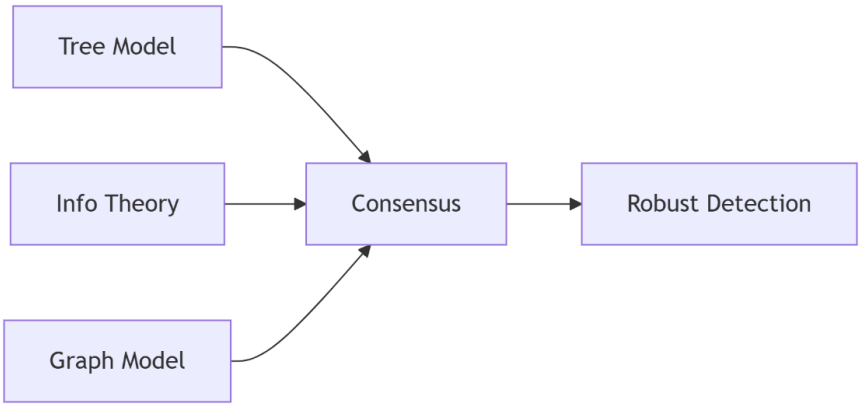


FIGURE 4

Ensemble Learning Framework

In summary, theoretical perspectives and conceptual foundations provide the foundation for this research study. While Isolation Forest algorithm serves as the primary anomaly detection method, its use and application is informed by the value of unpredictability based feature selection and therefore remains open to future integration with graph and ensemble techniques that are especially relevant in complex and dynamic environment of academic networks environments.

2.2 Network Traffic and Anomaly Detection

2.2.1 Understanding network Data Anomalies and Detection

An Anomaly or an outlier is typically defined as a data point or group of data points that deviate significantly from the majority of observations in a dataset. It can also be defined as a data object that significantly deviates from normal objects as if it was generated by different mechanism. Anomaly detection therefore is necessary to identify these deviated patterns in dataset as they do not conform to expected behaviour as illustrated in Figure 1 (Poonam, Deepika, & Gautam, 2021). From this illustration, normal data points are positively correlated in a diagonal shape while the anomalous data point (outlier) is far from them.

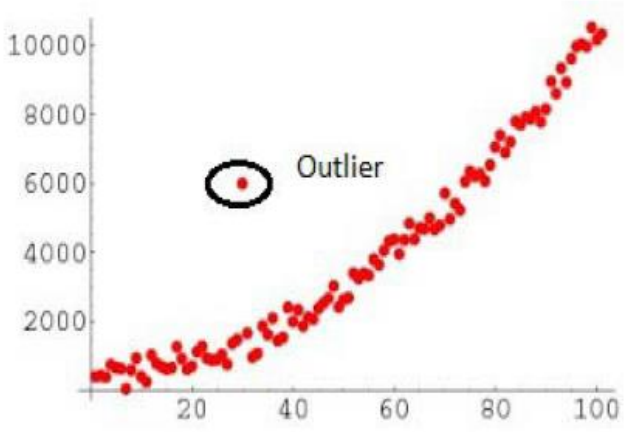


FIGURE 5

An illustration of outlier patterns in data

An irregular traffic pattern occurrence in a computer network could mean that a hacked computer is sending out sensitive data to an unauthorized destination. In anomaly detection, the normal behaviour is characterized by the model and any deviation that arises is an anomaly. The normal behaviour model represents the positive patterns that are allowed and the negative patterns that are considered as anomalies (Kumar V. , 2020)

In network systems, anomaly represents unusual or unexpected behaviour that may indicate issues such as cyberattacks, network misconfigurations or faults in network devices (Cheng, Wang, & Ma, 2022). These anomalies although some might not always harmful often require immediate attention as they can signal conditions that could affect network performance or compromise network security.

Anomaly detection therefore is crucial in network management as they may indicate security threats such as unauthorized access attempts, Distributed Denial of Service (DDoS) attacks or insider threats. Detecting them in real time is essential to prevent cyber criminals from escalating and causing significant damage to the network infrastructure.

Identifying anomalies early is significant as it ensures stable network operations and allows timely intervention to avoid network performance degradation. Prevention of extended network outages, reduced maintenance costs and consistency in service delivery will therefore be achieved (Chen, Zhang, & Li, 2023)

2.2.2 Taxonomy of Network Anomalies

Network traffic anomaly detection is guided by a taxonomy that classifies anomalies and irregularities into point, contextual and collective anomalies. This classification aids in the design of the model, determination of key features and also in the evaluation criteria of the model as described below: -

2.2.2.1 Point Anomalies

These anomalies represent isolated instances that deviate significantly from the normal behaviour. They are the easiest to detect as they stand out against background distribution of data. In cyber security, it includes the unusual large outbound network packets from the points and workstations that engages in low volume traffic and this signifies an early warning sign for possible data exfiltration attack (Chandola, Banerjee, & Kumar, 2009)

2.2.2.2 Contextual Anomalies

These anomalies arise when an event abnormality is dependent on its situational context. The same network behaviour may be benign and legitimate in one setting but also malicious and suspicious in another setting. A good example is high data throughput might be experienced during scheduled database synchronisations activities but identical activity that might occur during non-operational and off-peak hours could indicate malicious activities (Song, 2021). These anomalies however require spatial, user-behaviour or temporal context to be modelled alongside baseline traffic patterns.

2.2.2.3 Collective anomalies

These anomalies occur when a group of individually normal data points forms an unusual structure or sequence when they are analysed together. This is relevant in advanced persistent threat (APT) situations where intruders slowly exfiltrate data over some duration to prevent a trigger of volume-based alerts. Detection of these anomalies demands sequence-aware or aggregation-based models that are capable of recognising long term network deviations in behaviour.

Isolating these categories is critical in the configuration of Isolation Forest algorithm in the context of this study in order to detect both highly visible anomalies and stealthy deviations

that characterises sophisticated network intrusion and breaches in academic network environments.

2.2.3 Theoretical Framework of Anomaly Detection

This research project is guided by two main and key theoretical frameworks for network anomaly detection. Confidentiality-Integrity-Availability Triad (CIA Triad) and Anomaly Detection Theories. The two theories are important in explaining how data exfiltration attacks can be easily identified through unusual patterns in network traffic and how such incidents can pose serious risks to information privacy and security in an institution

2.2.3.1 Confidentiality-Integrity-Availability Triad (CIA Triad)

CIA Triad serves as the core theory and principle in cyber and network security. Confidentiality is directly impacted by data exfiltration as it involves unauthorised access and transfer of sensitive data (Alomari, Aldwairi, Alomari, & Al-Zoubi, 2022).

Breach in confidentiality always leads to significant reputational, financial and operational damage to an organization. This framework emphasises the need for organizations to enhance their network controls monitor capabilities to uphold confidentiality and prevent data leakage

2.2.3.2 Anomaly Detection framework

This provides the foundation for understanding how suspicious activities can be detected by carrying out analysis of any deviations from normal network traffic behaviour. In the context of data exfiltration, anomalies might include threats such as access to sensitive files outside normal working hours, unexpected spikes in outbound network traffic or connections to unusual and unfamiliar external servers.

Figure 6 below illustrates how machine learning and behavioural analysis of systems are increasingly becoming effective in identifying this threats that traditional signature-based tools have been unable to identify and detect (Zhang, Liu, & Chen, 2023)

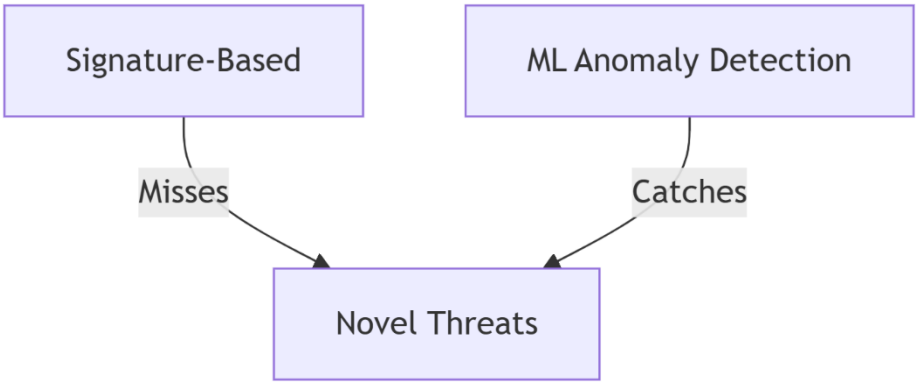


FIGURE 6

Signature Based vs Machine Learning Anomaly Identification

These two frameworks therefore offer a strong theoretical foundation for understanding data exfiltration problem especially through anomaly detection in network traffic.

2.2.4 Challenges in Detecting Data Exfiltration in Encrypted Traffic

Deployment of strong encryption protocols particularly Transport Layer Security version 1.3 (TLS 1.3) has transformed network security landscape by ensuring integrity and confidentiality of data in transit. TLS 1.3 is the most recent and secure version of the Transport Layer Security protocol that is designed to provide authenticated and encrypted communication over networks especially the internet

However, the same development presents some significant obstacles to network anomaly detection systems particularly in the detection of hidden data exfiltration. With payload contents masked, the traditional content-based inspection approaches especially deep packet inspection are rendered ineffective (Anderson & McGrew, 2019)

The recent research work highlights a growing dependence on flow level metadata including session duration, inter-packet timing and packet size distributions in identification of suspicious activity within encrypted network channels. Although these features provides indirect indicators of malicious activities, sophisticated challenges often creates network traffic patterns that closely resembles legitimate encrypted communications thus reducing anomaly detection efficieny (Papadogiannaki, Ioannidis, & Antonatos, 2021)

This challenge is further worsened in dynamic network environments where the behaviours significantly fluctuates. Anomaly detection approaches that are context-aware, adaptive and able to identify network deviations without reliance on static rule sets are required. Machine learning based approaches especially unsupervised model such as Isolation Forest algorithms, promises to be a viable alternative because normal encrypted network flow behaviour and outliers are able to be identified without the need of payload visibility. For these approaches to be effective, careful feature engineering and robust training on representative network traffic data is required

2.2.5 Unique Characteristics of Academic Networks

Academic computer networks represent distinct operational context for anomaly detection and is characterised by functional, policy-driven and structural factors that are different from government or corporate network infrastructures. These environments are normally heterogeneous and includes a wide range of connected components such as the traditional workstations, servers, IoT systems and laboratory systems (Gao, Sun, & Zhang, 2022). This results to the need for varied traffic profiles, communication protocols and access patterns.

In addition, academic network institutions often maintain open access policies to facilitate inter institutional data sharing, collaboration and remote access. Although this

facilitates research and innovation tasks, it creates vulnerabilities and complicates the establishment of stable mechanism for anomaly detection. Network traffic patterns are also influenced by the academic calendar with periods of intense data transfer being experienced during research data uploads, examinations and conference preparations. This variability of data transfers can mimic signatures of malicious network behaviour thus making anomaly detection to be more challenging.

Resource constraint faced by many academic cyber security and ICT teams is another critical factor. Staffing shortages and limited budgets often necessitates the need for automated low maintenance security measures that are capable of adapting to the evolving network usage patterns without the need for extensive manual network oversight. Therefore, unsupervised learning models such as Isolation Forest algorithm becomes a viable option due to its capacity to learn from diverse and changing network conditions thus making it relevant to this research study

2.3 Data Exfiltration Threats in Academic Institutions

The rapid digitization of academic institutions as a result of interconnections has brought immense benefits in terms of efficiency, communication and access to educational resources. However, this transformation has introduced significant cybersecurity threats such as data infiltration threat. Data exfiltration is a cyber threat that involves unauthorised transfer of sensitive data to external parties (Singh, 2023).

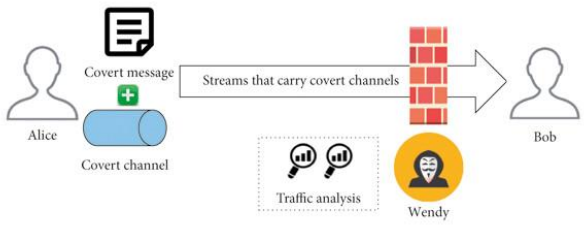
The most dangerous and hard to detect network security threat is data exfiltration. This threat involves an attacker stealing sensitive information and transferring it secretly to an external server. Unlike other disruptive attacks such as DDoS and ransomware, data exfiltration is stealthy and silent as the main goal of the attacker is unauthorised access while remaining undetected (Cybersecurity Ventures, 2023)

Academic institutions are becoming attractive targets to data exfiltration attacks due sensitivity of the data they hold such as student's personal information, academic records and research findings (Gregory, 2025). A report from Kenya Computer Incident Response Team highlights that there is a major increase in silent and stealthy data leaks in academic institutions that are often executed through backdoors, compromised user accounts or through phishing mechanisms. (KE-CIRT/CC, 2024)

Detection of data exfiltration threats with traditional defence mechanisms is proving to be a challenge because the behaviour of this attacks often mimics normal activity. The various data exfiltration threat vectors include: (Cybersecurity Ventures, 2023).

2.3.1 HTTPS Covert /Encrypted Traffic threat vector

This is when an attacker tunnel data through encrypted protocols such as DNS or HTTPS to avoid detection as illustrated in Figure 7. Many academic network systems use HTTPS protocol for normal operations such as accessing e-learning platforms. Attackers therefore exploits this trust to communicate with external Command and Control (C2) servers or disguise data leaks and thus normal traffic blends in



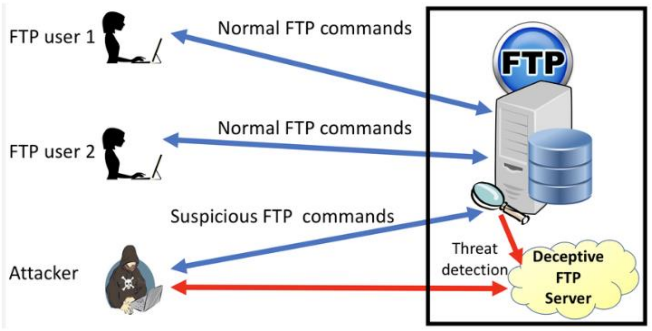
Source: https://onlinelibrary.wiley.com/doi/10.1155/2020/8892896

FIGURE 7

HTTPS Covert/ Encrypted Traffic vector attack

2.3.2 FTP Uploads threat Vector

FTP protocol is widely used in academic networks for legitimate file uploads. However, FTP often lacks encryption and unsecured making it a popular channel for data exfiltrating attacks. An attacker exploits compromised user credentials or weak passwords to upload institutional data such as exam materials or student records to remote servers as illustrated in Figure 8 below



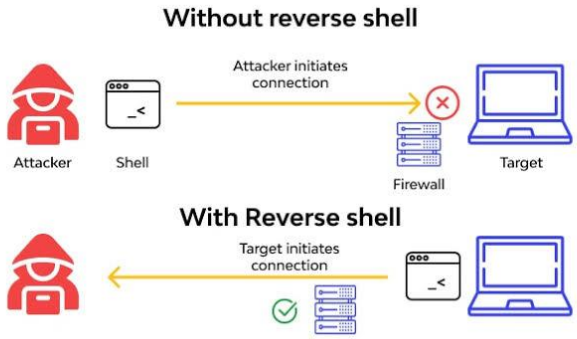
Source:https://www.researchgate.net/figure/FTP-services-hardened-against-APT-attacks_fig1_328326647

FIGURE 8

FTP-Based Data Exfiltration Attack Vector

2.3.3 Reverse Shells threat vector

This is a threat where an attacker gains remote command line access to victim's machine by establishing an outbound connection as illustrated in Figure 7 below. The attacker then leverages this to exfiltrate files silently across internal systems for examples lab computers that are always connected to the internet in an academic institution. These attacks are often difficult to detect using the existing traditional security tools as this threat vector often operates over trusted ports.



Source: https://www.wallarm.com/what/reverse-shell

FIGURE 9

Reverse Shells threat vector

2.4 Comparative Analysis of Existing Anomaly Detection Methods

Various anomaly detection techniques and algorithms have been developed to detect various anomalies in network environments, each bringing its own set of advantages and limitations. The complexity and scale of modern networks and the need for real time detection systems have driven the evolution of anomaly detection methods.

An overview of some of the commonly used approaches for anomaly detection in network systems, their advantages and limitations are summarised below: -

2.4.1 Statistical Methods

Statistical methods are often simple and computationally inexpensive. However, they often use predefined data distribution and may struggle to detect anomalies in high dimensional and nonlinear datasets. Modern network environments are characterized by these complexities and thus purely statistical methods might miss elusive patterns or overwhelmed by the volume of data. Examples of statistical detection methods with their merits and limitations are summarised in the Table 1 below (Wang, Chen, & Li, 2022).

Method	Description	Merits	Limitations
Z-scores Method	Measures how far a data point is from the mean in standard deviations.	Simple and easy to implement.	Assumes normal distribution. Ineffective in high dimensions.
Interquartile Range (IQR) Method	Identifies anomalies based on the range of the middle 50% of data.	Robust to anomalies in the data.	Limited to one variable (univariate) analysis. Struggles with high-dimensional data.
Hypothesis Testing Method	Tests whether a data point belongs to the expected distribution.	Can provide statistical validity.	Requires a predefined distribution. Computationally intensive.

TABLE 1

Comparison of Common Statistical based detection algorithms

2.4.2 Clustering-Based Methods

These methods are effective in detecting group anomalies. However, they struggle in detecting real-time and evolving threats in dynamic networks. Additionally, their sensitivity to parameter choices makes them unreliable if these parameters are not optimally selected. Summary of the major cluster-based methods are outlined in the Table 2 below (Xu, Shen, & Guo, 2023)

Method	Description	Merits	Limitations
K-Means Method	Partitions the dataset into *k* clusters; anomalies are points far from cluster centers.	Simple and effective for spherical clusters.	Sensitive to the choice of *k* (number of clusters). Struggles with noise in the data.
Density-Based Method	Groups points based on density; anomalies are located in low-density regions.	Effective for clusters of varying shapes and sizes.	Highly sensitive to parameter choices (e.g., neighborhood

Method	Description	Merits	Limitations
			radius, minimum points).
Hierarchical Clustering Method	Builds a tree-like structure of nested clusters; anomalies can appear as individual points or small branches.	Flexible in terms of cluster shapes; does not require pre-specifying the number of clusters.	Computationally expensive with large datasets. Difficult to scale.

TABLE 2
Comparison of common Cluster based detection algorithms

2.4.3 Machine Learning Methods

Machine learning techniques have emerged as the most powerful methods for anomaly detection especially in complex and high-dimensional datasets found in network traffic. Machine learning methods are broadly categorized into supervised and unsupervised learning approaches and each have distinct characteristics, advantages and limitations as reviewed below: -

2.4.3.1 Supervised Learning Algorithms

These are techniques that require labelled datasets to train models to differentiate normal and anomalous instances. The Table 3 below summarises the two main methods (Zhou, Wang, & Chen, 2022)

Method	Description	Merits	Limitations
Support Vector Machines (SVM)	Creates a hyperplane in a high-dimensional space to separate classes. Uses kernel	High accuracy with well-separated classes. Effective in high-dimensional spaces.	Requires large labelled datasets for training. Performance may degrade in dynamic

Method	Description	Merits	Limitations
	functions to handle nonlinear separations.		environments where patterns change over time.
Deep Neural Networks (DNNs)	Consists of multiple layers that learn complex patterns in large datasets, capable of modeling intricate feature relationships.	Highly effective for high-dimensional and complex datasets. Can learn feature representations automatically without extensive pre-processing.	Requires significant computational resources and time to train. Needs substantial labelled data, which may not be readily available. Risk of overfitting, especially with limited data.

TABLE 3

Comparison of common supervised machine learning Algorithms

2.4.3.2 Unsupervised Learning Algorithms

This are methods that does not require labelled data thus making them useful in scenarios where obtaining labels is difficult and expensive. Table 4 below summarises key features of the two most popular unsupervised techniques (Liu & Zhang, 2023)

Method	Description	Merits	Limitations
k-Nearest Neighbors (k-NN)	Identifies anomalies by calculating the local density around a data point. Points with fewer neighbors than a specified threshold are classified as anomalies.	Simple to understand and implement. Effective for small datasets with clear class boundaries.	Sensitive to the choice of distance metric. Computationally expensive for large datasets due to pairwise distance calculations.
Isolation Forest	Constructs an ensemble of random binary trees by recursively splitting	Efficient for high-dimensional data due to random partitioning.	May not capture nuanced anomalies in complex patterns (e.g., network traffic) unless

Method	Description	Merits	Limitations
	data on randomly chosen features and split values. Anomalies are isolated in fewer splits.	Requires minimal parameter tuning. User-friendly.	deviations are pronounced. Randomness can introduce variability; multiple runs may be needed for stability.

TABLE 4

Comparison of common unsupervised machine learning Algorithms

2.4.4 Deep Learning Approaches

Deep learning has become a dominant approach in network anomaly detection because it is able to automatically learn complex from a large-scale network traffic dataset. Deep learning can extract contextual patterns directly from pre-processed data unlike traditional machine learning approaches that heavily rely on manual feature engineering.

Features, merits and limitations of the major deep learning architectures are summarised in Table 5 below: -

Architecture	Description	Merits	Demerits
Convolutional Neural Networks (CNNs)	Widely used for detecting anomalies in network flows when traffic features are represented as structured data. Excel in identifying spatial correlations between traffic attributes and are effective in detecting malicious activity hidden within normal flow distributions.	Effective in capturing spatial patterns in network traffic features. Robust to noise.	May miss temporal network trends. Requires fixed input dimensions.

Architecture	Description	Merits	Demerits
Recurrent Neural Networks (RNNs)	RNNs—especially Long Short-Term Memory (LSTM) networks are suitable for modeling sequential dependencies. They capture temporal correlations in traffic sequences and are effective in detecting stealthy data exfiltration attacks that occur over extended durations.	Effective in modeling temporal and sequential dependencies.	Computationally intensive. Risk of vanishing gradients.
Transformer-Based Models	Leverage self-attention mechanisms to capture long-range dependencies efficiently. Offer high parallelization potential, enabling faster training for large-scale network datasets (Kim et al., 2022).	Effective in handling long-range dependencies. Highly parallelizable.	Require large datasets and substantial training resources.

TABLE 5

Key Deep Learning Architectures for Network Anomaly Detection

The major merit of deep learning approach is its ability to detect anomalies in encrypted network traffic by relying on flow metadata rather than payload content thus allowing effective anomaly detection in privacy sensitive environments where encryption is prevalent such as academic networks. However, the main disadvantage of deep learning approaches is that it requires substantial datasets for training, computationally intensive and lacks interpretability (Wang, Zhao, & Chen, 2023)

2.4.5 Comparative Strengths & Weaknesses of Supervised & Unsupervised approaches

Supervised and unsupervised machine learning algorithms have widely been applied to network anomaly detection and each of these approaches offers distinct merits and demerits

Supervised learning algorithms especially Random Forests, deep neural networks and Support Vector Machines (SVMs) heavily rely on labelled datasets in order to learn how to classify between normal and anomalous network traffic. Generally, they achieve high accuracy when they are trained on high quality labelled datasets. However, obtaining comprehensive and balanced labelled data in cyber security is challenging for new emerging network attack types and vectors. In addition, supervised models struggle with zero day attack an exploit that targets unknown vulnerability in the trained labelled data (Zhou, Wang, & Chen, 2022)

Unsupervised learning algorithms such as autoencoders, k-means clustering and Isolation Forest do not require labelled datasets but model normal network traffic behaviour and flags any deviations as potential anomalies traffic. This therefore makes unsupervised models suitable for the detection of previously unseen network threats. However, the main demerit is that they may produce high false positives especially in a dynamic environment such as an academic network where network traffic patterns and behaviour are different and evolves with time (Zhang, Chen, Wang, & Li, 2023)

A hybrid combination of supervised and unsupervised algorithms is widely being adopted to leverage the strengths of both where unsupervised methods are utilised to flag anomalies and supervised methods are for confirmation of malicious intention by an attacker

Table 6 below summarises these approaches highlighting the key merits, limitations and areas that are suitable to be applied.

Algorithm	**Merits**	**Limitations**	**Suitability**
Supervised	More accurate if a quality labeled dataset is available. Outputs are	Requires extensive labeled datasets. Poor at handling novel or unseen attacks.	Effective for detecting known attacks in a stable network environment.

Algorithm	Merits	Limitations	Suitability
	typically interpretable.		
Unsupervised	Effective in detecting zero-day or previously unknown attacks. Does not require labeled data.	May produce higher false-positive rates. Rare but normal events may be misclassified as anomalies.	Suitable for evolving networks and environments where threats are unknown or labels are unavailable.
Hybrid	Combines strengths of supervised and unsupervised approaches.	Very complex to implement, tune, and maintain.	Effective in mixed network environments with partial labelling or where both known and novel threats are present.

TABLE 6

Comparative analysis of supervised and unsupervised approaches.

2.5 Isolation Forest Algorithm

2.5.1 Theoretical Framework

Isolation Forest an algorithm introduced by Liu, Ting and Zhou in 2008 an unsupervised anomaly detection algorithm that identifies anomalies and outliers through recursive random partitioning as illustrated in Figure 10 below (Liu, Ting, & Zhou, 2008).

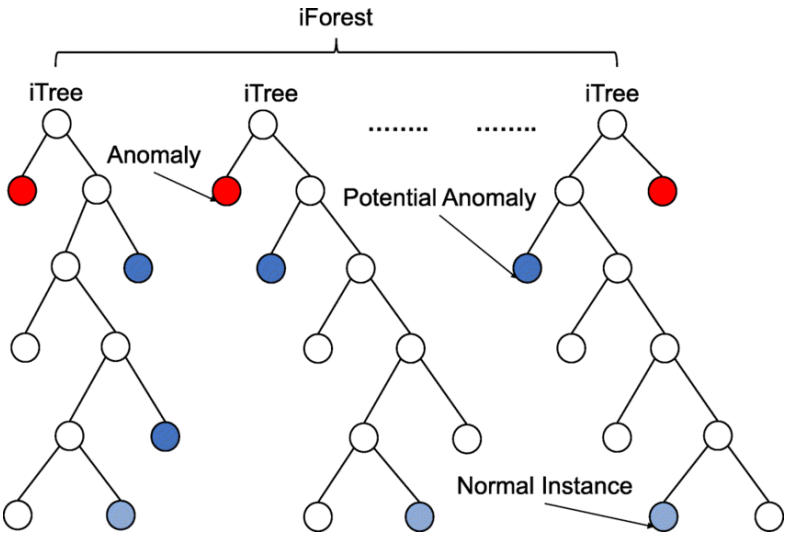


Image Source : https://www.researchgate.net/figure/Isolation-Forest-learned-iForest-construction-for-toy-dataset_fig1_352017898

FIGURE 10

Isolation Forest Algorithm Architecture

This algorithm operates on the following two main key principles: -

a) Anomalies are few and different and thus requires fewer random splits to isolate anomalies form normal data points

b) The algorithm uses and ensemble binary tree structure of isolation trees(iTrees) that are built with random split/feature selections as per *Figure 10* above.

In the training phase, multiple binary decision trees (iTrees) are trained with different subsamples drawn from original datasets. The iTrees follows the same procedure until it reaches the stopping condition as illustrated in Figure 11 below: -

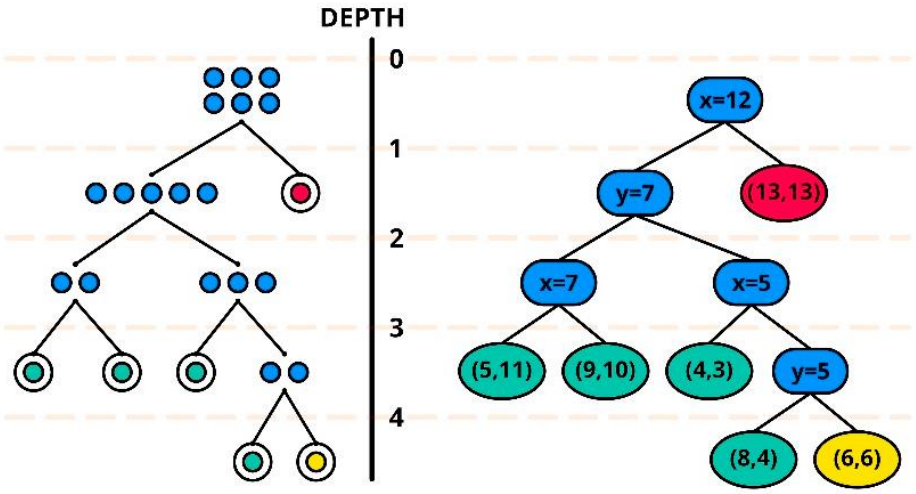


FIGURE 11

Conceptual illustration of binary tree with the isolated point.

This algorithm was developed to address the challenge of detecting anomaly in high-dimensional datasets common in many fields such as fraud detection, network traffic analysis and data mining.

The algorithm constructs an ensemble of binary trees by randomly partitioning the feature space. The path length from the root of each iTree to an instance serves as the measure of isolation. Anomalies in Isolation Forest tend to have shorter paths compared to normal instances (Zhang et al., 2023).

As illustrated in Figure 12 below, binary search tree is randomly constructed to isolate each data point. As per the figure, one split is required to separate the anomalous point (13,13) from the rest of the data and four splits are required to separate the normal point (6,6). This suggests that abnormal data instance is more likely to be closer to the root node than the normal data instance. (Aggarwal, 2016)

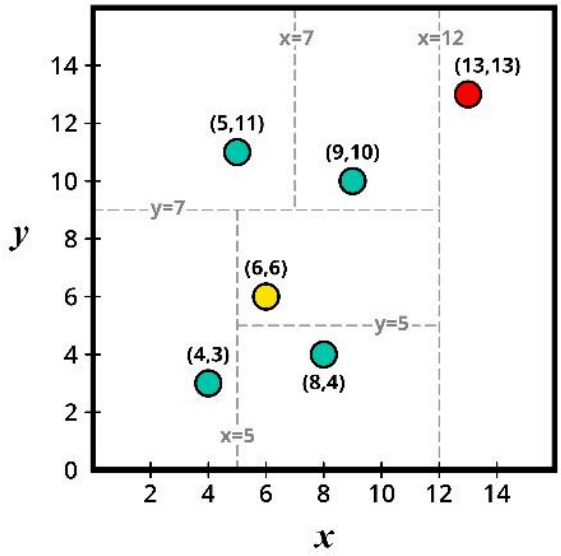


FIGURE 12

Conceptual illustration of isolation operations

2.5.2 Empirical Applications in Network Anomaly Detection

The Isolation Forest (iForest) algorithm has become a prominent tool in network anomaly detection due to its unique approach in isolating anomalies rather than modelling data distributions. This allows the algorithm to efficiently handle large and high-dimensional datasets such as in network traffic. The algorithm is scalability making it suitable for network environments where traffic data is voluminous and complex.

Isolation Forest success in detecting both known and unknown network anomalies has been demonstrated in several recent empirical studies thus justifying its crucial role in modern network security systems.

Recent work by Ahmed showed that the algorithm effectively identified Distributed Denial of Service (DDoS) attacks in large-scale networks (Ahmed, Khan, & Hashem, 2022). Their experiments utilized real-world datasets with varying levels of traffic thus demonstrating that iForest could differentiate between normal traffic and attack patterns with a high degree of

accuracy. The study showed that even during periods of high traffic load, the algorithm maintained low false-positive rates while providing near real time anomaly detection an important aspect in environments where speed and precision is critical in mitigating ongoing attacks.

Kumar and Singh examined the use of the algorithm in detecting insider threats, attacks that traditional signature-based methods often fail to detect within enterprise networks (Kumar & Singh, 2022). Insider threats are typically characterized by elusive deviations from normal user behaviour which is difficult to identify without advanced machine learning techniques. They applied iForest to internal network traffic and observed that the algorithm was able to isolate anomalous activities indicative of insider attacks with high accuracy. The algorithm outperformed conventional anomaly detection methods such as k-means clustering and principal component analysis. This makes Isolation Forest an excellent algorithm for proactive security measures in an academic institution setting where insider threats is a major concern.

Utilization of Isolation Forest algorithm in network anomaly detection is further supported by its performance across other sectors that are not network related illustrating its robustness and versatility. Li, Wang & Chang applied the algorithm in identification of fraudulent activities in e-commerce platforms (Li, Wang, & Chen, 2022). This study emphasized that iForest successfully isolated rare but impactful fraud cases from vast volumes of transactional data. Adaptability of the algorithm to detect anomalies in transactional data characterized by high dimensionality proved its effectiveness in network traffic

Zhang, Huang and Zhou demonstrated the application of the algorithm in healthcare specifically detection of anomalies in medical sensor data. Medical data shares similarities with network traffic as it is high-dimensional, continuous and vulnerable to sudden unexpected changes (Zhang, Huang, & Zhou, 2022). The researchers found out that the algorithm could

efficiently process this continuous stream of sensor data to detect anomalous events such as irregular patient activity and faulty readings. The capacity to work with real time data strengthened algorithm's suitability for network monitoring.

The empirical evidence from these diverse applications underscores the importance of Isolation Forest algorithm across different anomaly detection scenarios. The algorithm's success in sectors such as finance and healthcare where anomaly detection is critical further justifies its suitability for network anomaly detection. In cybersecurity, where rapid response to anomalies is critical, Isolation Forest algorithm stands out as an efficient and highly adaptable solution.

Algorithm's capacity for unsupervised learning make it more appealing for network anomaly detection as it does not require labelled training data that can be scarce or unavailable in many real-world scenarios. Unsupervised nature of Isolation Forest allows the algorithm to detect threats that signature-based systems might not detect thus providing a more comprehensive security framework for modern networks.

In summary, Isolation Forest algorithm recent applications in network anomaly detection and its application in other sectors demonstrates its adaptability and reliability in real time environments. As networks becomes more complex and the nature of cyber threats evolves, use of advanced anomaly detection algorithms like Isolation Forest algorithm is essential to maintain secure and robust network systems.

2.5.3 Hyperparameter Tuning for Network Traffic

Hyperparameter tuning is essential in optimising anomaly detection algorithms especially for network traffic datasets where feature distributions vary significantly with time.

Parameters such as subsampling size, contamination rate and number of trees in Isolation Forest directly influence false positive rates and detection accuracy

Hyperparameter tuning is often dataset specific in network anomaly detection. High volume network traffic environments especially in academic networks will mean the need for a larger number of estimators to improve robustness. Smaller subsampling sizes will accelerate detection in real time systems without compromising detection accuracy.

Hyperparameter optimisation however presents a unique challenge when compared to supervised learning methods as the ground truth for anomalies evaluation may be unavailable or incomplete and thus will often requires usage of proxy metrics for calibration (Rahman, Alam, & Ahmed, 2022)

2.5.4 Hybrid Isolation Forest Models

These models combine the core principle of Isolation Forest that involves partitioning feature space through recursive random splits but with complementary techniques in order to enhance detection adaptability, accuracy and robustness

The common one involves integration of feature selection before Isolation Forest processing in order to reduce noise and computational cost while retaining key features. The other involves incorporating supervised classifiers. A good example is using Isolation Forest for anomaly scoring and then passing flagged network instances to supervised or deep learning approach to refine classification (Zhao, Liu, & Sun, 2023)

This models also addresses high false positive rates possibilities a known demerit of unsupervised approaches by the combination of statistical approaches with Isolation Forest models in order to capture complex data distributions. Encrypted network environments with

a hybrid of temporal modelling and flow-based features can improve detection of silent threats such as data exfiltration over extended durations. (Ahmed, Khan, & Farooq, 2022)

These hybrid Isolation Forest Models are suitable for academic networks especially where network traffic is both heterogeneous and highly dynamic and also where the risk of new emerging unseen attack patterns is higher. However, the disadvantage of hybrid models is that they require more computational resources. In addition, careful integration of these models is required to avoid complexity that could hinder real time anomaly detection performance.

2.6 Datasets for Intrusion and Anomaly Detection

2.6.1 CICIDS2017 Dataset Overview

Several studies in network security projects have heavily relied on CICIDS2017 dataset developed by Canadian Institute for Cybersecurity due to its realistic representation of modern network traffic environments (CIC, 2024). The dataset combines normal traffic with a variety of documented malicious behaviours thus making it a popular standard for evaluating intrusion detection in network systems.

In this research, Thursday subset as illustrated in Figure 13 below was utilised as it contains more advanced and stealthy attack scenarios such as reverse shell activity, command and control communication and DNS tunnelling common methods used in data exfiltration attacks (CIC, 2024).

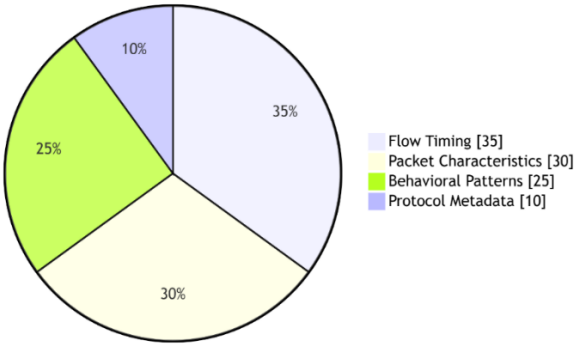


FIGURE 13

Key Feature Categories in CICIDS2017 Thursday Subset

These threats are relevant to academic institutions where sensitive data is targeted and traditional security tools may not handle. Thursday data subset reflects these real-world threats making it useful for this research. In addition, the dataset provided rich set of features such as flow duration, byte rates, packet rates and inter-arrival times features that supported the development of the model.

Recent literature by Wahome, Mbugua and Maina have highlighted the value of this datasets in supporting machine learning algorithms such as Isolation Forest which can detect silent and low volume data exfiltration attempts by cyber criminals that might otherwise be unnoticed. This thus reinforces this dataset relevance not only for intrusion detection systems in general but also for building scalable security solutions suitable for resource constrained settings. (Wahome, Mbugua, & Maina, 2023)

2.6.2 CICIDS2017 dataset Limitations

CICIDS2017 dataset presents some limitations that can influence research outcomes although it is regarded as the most comprehensive publicly available datasets for intrusion detection research work. Although it contains a broader range of attack scenarios such as brute force, infiltration, DoS and DDoS, the simulated nature of network traffic may not fully capture

the diversity and unpredictability of real-world network behaviour (Sharafaldin, Lashkari, & Ghorbani, 2023). This therefore limits its natural validity especially in the deployment in dynamic network environments

Class imbalance is another challenge as certain network attack type are overrepresented while others such as low and slow attacks or exfiltration or appears less often. This class imbalance leads to creation of biased models that performs well on only frequent attack types and fails to detect rare ones. In addition, dataset mostly uses unencrypted traffic that does not fully reflect the growing use of encrypted communication in modern networks. This poses limitations in the evaluating models that are intended for encrypted datasets (Lopez, Kaur, & Conti, 2022)

In addition, CICIDS2017 dataset was collected in a controlled lab environment. While the dataset was collected consistently, it may lack variability, concurrent traffic patterns and noise that are present in real networks. Therefore, models that are trained exclusively on CICIDS2017 dataset may overfit to synthetic characteristics and thus may underperform in real deployments.

CICIDS2017 dataset however remains highly relevant to this research despite these limitations. The dataset offers a rich and diverse set of labelled network traffic flows that have both normal and anomalous behaviours which enables robust unsupervised and supervised model evaluation. Also, this dataset details flow level features such as flow duration, byte counts and packet counts that are directly compatible with the Isolation Forest algorithm detection approach that is employed in this research study.

Given that the main aim of this research study is evaluation of anomaly detection capability in a reproducible manner, CICIDS2017 dataset provides an acceptable reference point in cybersecurity research studies. This is because the comparison of the results with

existing research work is possible. while real world datasets are more authentic, they lack labels and privacy restrictions making CICIDS2017 dataset ethically accessible to researchers and practical especially within the academic network research context.

2.6.3 Alternative Publicly available Datasets

Cyber security researchers have explored other existing publicly available datasets that have realistic attack scenarios and different traffic patterns to address the challenges of CICIDS2017 dataset. The major example is University of New South Wales – Network Based 2015 Dataset (UNSW-NB15) dataset that includes a mix of real and synthetic traffic with wide range of common attacks thus making it relevant for research studies involving encrypted threats (Moustafa & Slay, 2022.)

Another adopted dataset that focuses on Internet of Things network traffic is Telemetry, Operational and Network data for Internet of Things (TON_IoT) dataset. It constitutes application level data and packet captures. Although its not directly aligned with all academic network environments, this dataset provides important insights in the security of Internet of Things devices that are present in campus networks (Alshamrani, 2023)

Other existing datasets such as National Security Laboratory-Knowledge Discovery in Databases (NSL-KDD) and Measurement and Analysis on the WIDE Internet (MAWI) are useful for benchmarking purposes but their narrow scope and age limits their importance to the emerging network threats.

Merits and demerits of these alternative datasets are summarised in *Table 7* below

Dataset	Merits	Demerits
UNSW-NB15	Contains a mix of synthetic and real network traffic. Suitable for evaluating detection of modern attacks.	Limited volume compared to larger datasets like CICIDS2017.
NSL-KDD	Consistent and widely adopted for benchmarking. Extensively studied and documented in the literature.	Contains outdated attack patterns that may not reflect current threats.
TON_IoT	Specifically designed for IoT traffic studies. Includes multimodal data (network, telemetry, and operational).	Less relevant for non-IoT or traditional enterprise network environments.
MAWI	Comprises real backbone traffic with high variability. Captures realistic network behaviour and background traffic.	Lacks comprehensive labeled ground truth for all network flows.

TABLE 7

Common Public Datasets for Network Anomaly Detection

Choice of research dataset always depends on research objectives and in most cases, researchers always combines multiple datasets in their work in order to reduce overfitting to a single network traffic profile.

2.7 Feature Engineering and Data Pre-processing in Cybersecurity

Feature engineering and data pre-processing are important in anomaly detection for network data exfiltration in order to achieve accuracy in anomaly detection. This is because

raw network traffic data especially packet captures or flow logs are often noisy and of high dimension. Data pre-processing especially for anomaly detection models such as Isolation Forest algorithms is therefore necessary so that the model does not struggle with computational inefficiency and false positives

Features dimensionality reduction techniques have been widely used in network intrusion detection research works on order to reduce complexity at the same time preserving structure of network traffic patterns (Altin & Çakir, 2024). Principal Component Analysis (PCA) technique is important in the removal of redundant features especially in large scale datasets such as CICIDS2017. t-Distributed Stochastic Neighbor Embedding (t-SNE) technique on the other end supports exploratory visualization of anomalies that is useful in the detection of patterns of possible data exfiltration attempts.

SHapley Additive exPlanations (SHAP) method has emerged as a key method in the identification of features that contributes most to anomaly scores. Integrating SHAP method with Isolation Forest algorithm can reduce feature sets thus improvement of real time performance will be achieved without compromising accuracy. This is very critical in environments with limited cyber security staffing especially in university IT departments (Journal of Big Data, 2025)

2.8 Regulatory and Ethical Considerations

2.8.1 Kenyan Data Protection Act (2019)

The Kenya Data Protection Act (DPA) of 2019 has operationalised Article 31 of 2010 Kenyan Constitution of Kenya which recognises data privacy as one of the fundamental human rights (Mankone, 2023). The act establishes a comprehensive legal and regulatory framework that governs lawful collection, processing, storage and distribution of personal data across

public and private sectors. Its guidelines are fairness, confidentiality, transparency and integrity of personal data and are aligned with international standards including European Union General Data Protection Regulation (GDPR). This ensures Kenya's approach on data privacy is globally interoperable (Africa Law Partners, 2023)

Office of the Data Protection Commissioner serves as the primary oversight authority that have statutory powers to mandate corrective actions, impose administrative fines and issue compliance notices in cases of data breach (Mankone, 2023). This oversight is very crucial for maintaining lawful data processing in institutions such as colleges and universities that operates with high volume and complex ICT environments while enabling operational efficiency

DPA regulation of systematic large-scale processing of monitoring of personal data is highly relevance to network traffic anomaly detection. Section 31 of the DPA requires that organisations that are engaged in such activities especially continuous surveillance and monitoring of network activities are required to appoint a Data Protection Officer and also conduct Data Protection Impact Assessment. This should be done before the deployment of network monitoring systems in order to assess and mitigate any potential privacy risks to data

These obligations are very important in the context of academic network security. Network logs that are used for intrusion detection may contain personally identifiable information including MAC addresses, IP addresses, potentially user account details and device identifiers and. Such information is categorised as personal data under the DPA and are therefore fully subjected to statutory protection.

Therefore, anomaly detection framework including Isolation Forest based intrusion detection model must adopt privacy by design and privacy by default principles. This constitutes reduction of unnecessary data collection and ensuring encryption of storage and data transmission (Mankone, 2023)

Institutions can achieve a balance between statutory privacy requirements and network security objectives if these safeguards are adopted. This will ensure that advanced network monitoring systems operates within the legal boundaries that are established by the Kenya DPA while at the same time maintaining public trust in academic network environments.

2.8.2 Institutional Policies in Academic Environments

Academic institutions in Kenya such as University of Nairobi and Kenya Methodist University have established their own internal cybersecurity data protection and privacy policies that ensure full compliance with the Kenya DPA (2019) while at the same time safeguarding institutional operational and educational objectives (Mankone, 2023). These frameworks serve as compliance tools and guiding documents that balance security measures while preserving ethical responsibility and academic freedom while handling data (Africa Law Partners, 2023)

These policies are aligned with the principles that are in the international best practices such as General Data Protection Regulation (GDPR) and national legislation frameworks. Kenya Methodist University Privacy Policy (2025) is embedded within Data Protection Policy framework. It specifies the categories of technical and personal data that are collected through online platforms such as network access logs and academic records. It outlines the safety measures audit trails, regular compliance reviews and encryption to prevent misuse. User rights and institutional responsibilities are also defined in the policy to ensure that monitoring activities remain lawful (KEMU, 2025).

Institutional policies in summary provides a structured legal and ethical context that facilitates the deploying of advanced anomaly detection systems as per the proposed Isolation Forest-based intrusion detection model. This is to ensure systems maintains user trust, conforms to both privacy by design and security by design principles and operates with a

clearly defined authorisations (Mankone, 2023), (Africa Law Partners, 2023). Therefore, integrating technical safeguards and policy is critical in academic network environments where open nature of networks heightens vulnerability. The diversity of network users makes it necessary for the of a robust data privacy protection.

2.9 Operationalization of Variables

Operationalisation of variables ensures that both the process of data collection and analysis are conducted in a systematic and an enhanced measurable manner. Operational of the variables were aligned with network traffic anomaly detection research to ensure research outcomes were measured precisely and reproducible

Independent variables are input features that are extracted from raw network traffic data. They serve as predictors in network anomaly detection model. These features were selected based on their relevance in previous network intrusion detection works especially in the representation of network traffic characteristics that are able to distinguish anomalous activity from normal behaviour

Dependent variable represented final classification output of the model that indicated whether a given network traffic flow is normal or anomalous. An intermediate variable, the anomaly score was included as part of model decision-making process. It provided a continuous measure of deviation from normal traffic patterns before binary classification was applied.

The operational definitions for the variables used are summarised in Table 5 below.

Variable	Type	Operational Definition
Flow Duration	Independent	Total time (in milliseconds) for a network flow from start to finish. Short or abnormally long durations can indicate data exfiltration or reconnaissance.
Total Bytes Sent	Independent	Total amount of data (in bytes) transmitted from source to destination. Excessively high values may suggest large-file transfers or exfiltration; low values could indicate probing or scanning.
Destination IP	Independent	IP address receiving network traffic. Rarely contacted, blacklisted, or unusual destination IPs may indicate unauthorized or suspicious communication.
Time of Activity	Independent	Timestamp when the network flow occurred. Traffic outside normal operating hours (e.g., after business hours) may suggest unauthorized access attempts.
Flow Bytes/s	Independent	Average data transfer rate (bytes per second) for a flow. Unexpected drops or sudden spikes may reflect malicious activity, flooding, or abnormal usage patterns.
Anomaly Label	Dependent	Binary classification output of the model, where 1 represents normal (benign) traffic and -1 represents anomalous traffic. Derived by applying a threshold to the model's anomaly score.

TABLE 8

Operationalisation of Variables

The above variables as tabulated in Table 8 above collectively enabled effective transformation of raw network traffic data into a structured analytical input for the Isolation Forest based anomaly detection model. Clear definition of each of the above metrics ensured consistency in feature extraction. This allowed accurate labelling and supported rigorous evaluation of the isolation forest model detection capability.

Anomaly Score an Intermediate variable quantified how much a network traffic flow deviated from the normal behaviour with values that will range from 0 to 1. It was calculated based on the average path length that was needed to isolate a data point in a randomly constructed tree. The formula for the anomaly score is as follows (Liu, Ting, & Zhou, 2008): -

$$s(x, n) = 2^{-\frac{E(h(x))}{c(n)}}$$

where:

- $s(x, n)$ is the anomaly score for data point x,
- $E(h(x))$ is the average path length to isolate x,
- $c(n)$ is a normalization factor based on the average path length in a binary search tree Zhou, 2008).

Any score closer to 1 mean the flow is most likely anomalous while a score near 0 suggests normal network traffic.

The Anomaly Label assigns each network flow a class based on a threshold usually 0.5 as illustrated below: -

$$\text{Anomaly Label} = \begin{cases} -1, & \text{if } s(x, n) \geq 0.5 \quad \text{(Anomalous)} \\ 1, & \text{if } s(x, n) < 0.5 \quad \text{(Normal)} \end{cases}$$

From the above, any network flow with a score above 0.5 was labelled as anomalous (-1) thus indicating unusual network behaviour signifying possible security threats while a score below 0.5 was considered normal (1).

This method allowed the model to learn how typical network behaviour looks like and flag deviations effectively thus supporting proactive detection of potential security incidents (Liu, Ting & Zhou, 2008).

2.10 Conceptual Framework

Based on literature review insights done and operationalization of variables, the conceptual framework of the model was visually depicted as illustrated in the Figure 14 below.

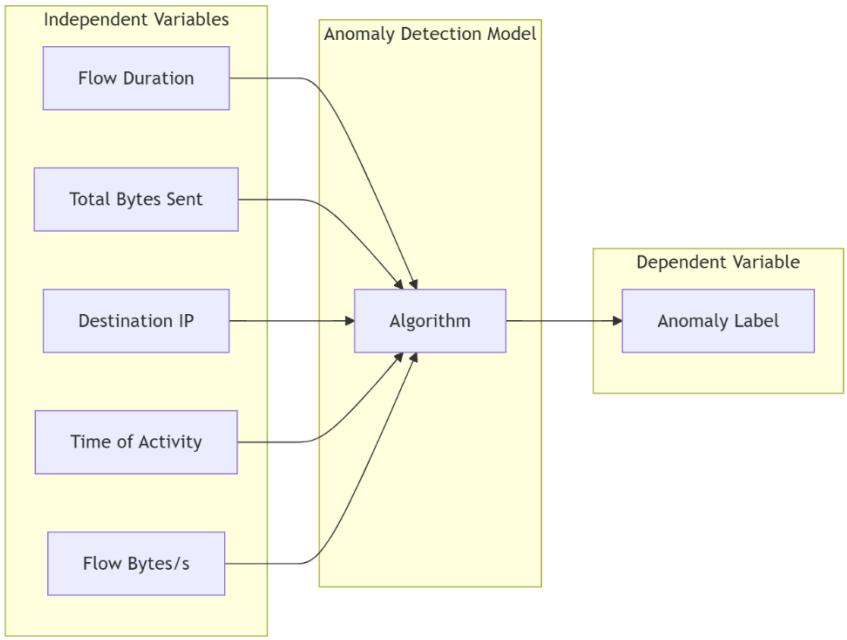


FIGURE 14

Conceptual Framework of Isolation Forest anomaly detection

The above conceptual framework is based on a structured transformation process. Input features/independent variables from network traffic data are analysed using the isolation forest algorithm that produces an intermediate/anomaly Score that informs binary classification output of Anomaly Label indicating whether network flow is normal or anomalous.

As illustrated in figure 14 above, the input features are extracted from raw network traffic logs and included into flow characteristics and contextual markers categories. Flow characteristics features such as Flow Duration, Bytes Sent, Flow Bytes quantified aspects of how data moved through the network. Contextual marker features such as Destination IP and Time of Activity provided additional context around network traffic session.

Intermediate variable anomaly score generated internally by the model that reflects degree of deviation of network traffic instance from normal pattern will be vital in deriving final classification

Dependent variable anomaly label is the final output/ binary classification result of the model where 1 signifies normal traffic while -1 signifies anomalous traffic

2.11 Research Gap and Synthesis

The reviewed literature demonstrates extensive progress in network anomaly detection using statistical, clustering and machine-learning techniques. However, most of the existing approaches focuses on general intrusion detection and performs poorly in identifying data exfiltration intrusions particularly within encrypted academic network environments.

Traditional statistical and clustering methods relies on predefined patterns and assumptions that fails to capture complex and nonlinear patterns of modern network traffic. Supervised models although accurate requires large, labelled datasets that are rarely available in academic settings. Unsupervised techniques such as Isolation Forest algorithm have shown better adaptability and efficiency. From the review, most studies applying Isolation Forest algorithm using benchmark datasets such as CICIDS2017 dataset that do not represent the dynamic and heterogeneous nature of academic networks (CIC, 2024).

In addition, there is limited exploration of Isolation Forest algorithm or hybrid models in detecting encrypted or covert data transfers where traffic inspection is hindered by HTTPS and VPN use. Feature engineering in this context thus remains underdeveloped with few studies targeting flow level attributes that may signal data exfiltration behaviour.

Therefore, a clear research gap exists in adapting and fine-tuning unsupervised models such as Isolation Forest algorithm to academic network environments incorporating context

specific features and aligning detection approaches with regulatory frameworks such as Kenya's Data Protection Act (2019). This study sought to bridge these gaps by developing and evaluating an optimized Isolation Forest algorithm based model for detecting data exfiltration anomalies in academic network traffic.

2.12 Conclusion

Based on this literature, it can be concluded that data exfiltration is a growing and serious cyber threat especially in academic institutions setting as it often operates with limited cybersecurity solutions.

Isolation Forest algorithm thus stood out as a viable solution for detecting stealth and silent threats without the need of having labelled data or large computational resources. While Isolation Forest success in other sectors is well documented, its application to academic networks setup for detecting data exfiltration attacks is not well explored a gap that this research addressed

CHAPTER THREE

RESEARCH METHODOLOGY

3.0 Introduction

This chapter presents research methodology that was adopted in the design, development and evaluation of the unsupervised machine learning algorithm the Isolation Forest to detect data exfiltration threats. Structured approach that included research design, data acquisition, data pre-processing, feature selection, exploratory data analysis, model training and evaluation is provided. The methodology reflects practical application in building and evaluating an anomaly detection solution that is capable of identifying unauthorized data transfers in network environments especially for academic institutions that have limited labelled data.

3.1 Research Design

The study employed quantitative experimental research design that was complemented by exploratory data analysis that guided the development of the model using Isolation Forest algorithm. Quantitative method provided a structured approach for data collecting, processing and analysis through numerical metrics while experimental aspect allowed manipulation of inputs and testing of the algorithm with simulated network traffic dataset that reflects real conditions.

The model designed can be applied to an existing network traffic dataset which contains both benign network activities and simulated infiltration attempts. Although the model can operate in an unsupervised environment, the output generated was compared with the available ground truth labels to assess its effectiveness in identification of abnormal or suspicious behaviour.

Exploratory data analysis was conducted as a foundational step to ensure the model was built on a solid understanding of the dataset. The analysis assisted in revealing trends, relationships and patterns within the data before model development.

The combination of structured research design and data exploration ensured the development of a model that can perform well technically and also reflects a clear understanding of how network attacks and threats occurs in academic network environment.

Choice of experimental design was motivated by the need to train and evaluate the model in a controlled and data driven anomaly detection environment as illustrated in Figure 15 below

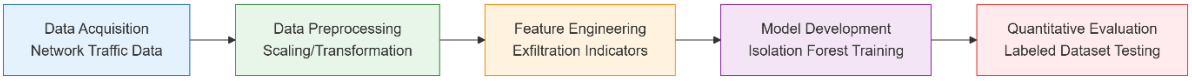


FIGURE 15

Anomaly Detection Model Design Phases

3.2 Data Acquisition

3.2.1 Source of Data

To simulate realistic network environment that includes benign traffic and malicious attacks, secondary data sourced from publicly available network traffic datasets commonly used in cybersecurity anomaly detection research was used. The primary dataset that was selected and used was CICIDS2017 dataset provided by the Canadian Institute for Cybersecurity specifically Thursday Working Hours Afternoon subset (CIC, 2024).

3.2.2 Justification for Dataset Selection

This dataset was chosen because it is realistic in representation of modern network traffic behaviour especially in cyber security research. The data has a comprehensive coverage

of modern network traffic scenarios and includes labelled attacks and thus make it more relevant to anomaly-based intrusion detection research works

Prior to generation of this dataset, the other publicly available network datasets such as NSL-KDD, KDD99 have been criticized for being outdated and lacks the complexity of real-world traffic. The Canadian Institute for Cybersecurity based at the University of New Brunswick in Canada developed a new suite of datasets under controlled lab conditions to reflect real world network environments in response to these limitations

CICIDS2017 dataset was generated over a period of five consecutive weekdays; Monday to Friday where typical user behaviour such as emailing, video streaming, web browsing, chatting, file transfers and various malicious activities such as data exfiltration attempts over HTTP and FTP protocols were performed.

The dataset also contains labelled exfiltration activities and thus allowed supervised evaluation of the model against known exfiltration behaviours. Dataset realistic simulation of external and insider attacks make it ideal for training and evaluating the model to detect anomalies in network traffic

Therefore, this dataset not only meet technical requirements of this study but also supports specific objectives of to investigation, modelling and evaluation of the model in detecting data exfiltration anomalies in network traffic.

3.2.3 Overview of the Dataset

This dataset contained approximately 288,602 network flow instances and more than 79 features capturing both benign and malicious activities that are relevant to academic network environments. Some of the key dataset network traffic features are tabulated in Table 9 below:

Feature	Description
Flow duration	▪ This refers to duration of how long each communication session lasts
Total bytes sent and received	▪ This refers to amount of exchanged data
Flow Bytes per second	▪ This refers to the speed of data transmission
Source and destination IP addresses and ports	▪ This refers to source and destination where traffic is coming from and going to
Time of activity	▪ This refers to time the traffic occurred
Packet sizes and counts	▪ This refers to how data is broken up and sent
Protocols used	▪ This refers to network communication rules used in sending data

TABLE 9

Key dataset network traffic features

Some of the above features played an important in the development of the model. Attackers intending to steal data in a network environment often leave behind unusual patterns such as large volume of data being sent at odd hours to some unknown IP addresses and destinations

3.3 Data Sampling Procedure

3.3.1 Target Population

The research study mainly focused on academic institutions including colleges, universities and technical training institutions as they form relevant and practical target population as they are increasingly exposure to network and cybersecurity threats and they have limited defence capabilities.

This institution often operates under limited and constrained ICT environments that have minimal cybersecurity investments and yet they support a diverse user population that

generates huge network traffic patterns. Therefore, they are ideal for testing anomaly-based detection using unsupervised machine learning models.

Key characteristics used in defining the target population included: -

- Limited cybersecurity infrastructure as many academic institutions lack modern intrusion detection systems
- Lack of skilled personnel as technicians might lack specialized skills in threat detection.
- Diverse network traffic sources as network traffic originates from faculty, students, administration and guests on shared infrastructure.
- Risks of insider threats risks as BYOD (Bring Your Own Device) usage contributes to internal risks exposure
- Usage & reliance on email platforms, third party cloud services and LMS (Learning Management Systems) increases network vulnerability.

3.3.2 Sampling Technique

This research study adopted purposive sampling strategy a method where samples are selected based specific criteria that is highly relevant to the research objectives. This enabled deliberate selection of data that met predefined characteristics which were necessary for the study. In this work, 'Thursday-WorkingHours-Afternoon-Infiltration' subset of CICIDS2017 dataset was selected because it closely aligns with the research study main objective of detection of data exfiltration and infiltration anomalies within an academic network environment.

The subset simulated typical daytime network traffic activities in an academic institution network setting and contained both normal user behaviour and realistic anomalous scenarios. Justification for the choice of this dataset subset is summarised below: -

- Relevance to academic network settings - The data subset reflected working hours traffic patterns in academic institution networks
- Existence of relevant threats in the data subset - The data subset contained simulated data infiltration attacks such as FTP-based uploads, reverse shell and covert infiltration
- Existence of a balanced network traffic composition - The data subset included both malicious and benign traffic thus supported training of the model.
- Focus on specifics and experimental control achieved - The data subset is narrow and targeted and thus allowed controlled experiments and precision in model evaluation

In summary, using purposive sampling technique ensured that the selected data was relevant and representative of the network conditions faced in academic institutions networks thus enhancing model's applicability and effectiveness.

3.4 Data Preparation and Pre-processing

The acquired data underwent a systematic pre-processing and analysis to prepare the data for anomaly detection modelling and to ensure the dataset was suitable able to show meaningful differences between normal and anomalous network activities as illustrated in Figure 16 below:

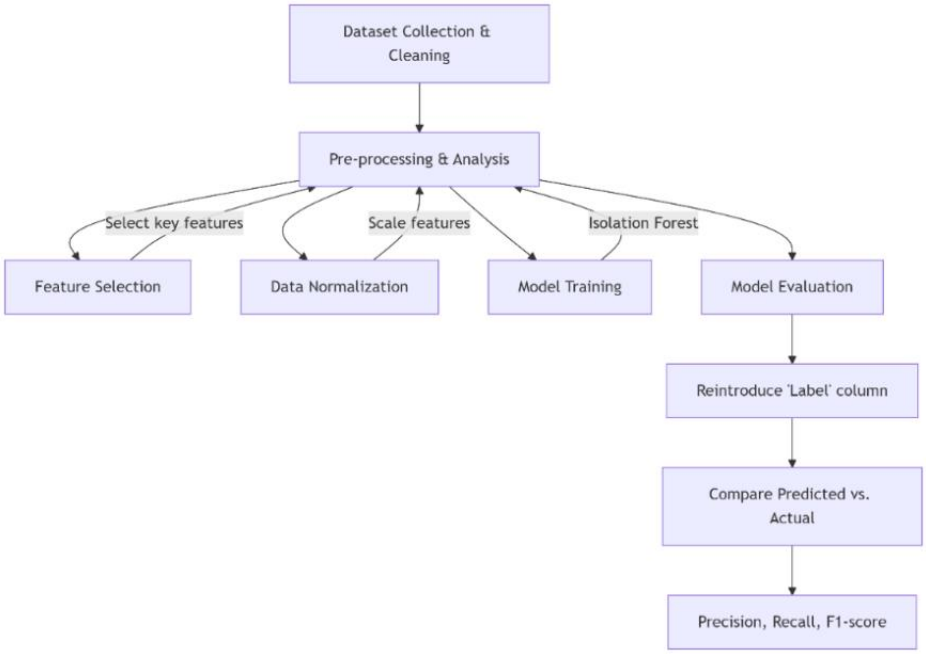


FIGURE 16

Data Preparation and Pre-processing

3.4.1 Exploratory Data Analysis (EDA)

Before feature selection process, Exploratory Data Analysis (EDA) was carried out in order to understand the structure of the dataset used, identify key patterns and guide informed decisions on the features to be included in the model.

Exploratory data analysis focused on data exfiltration and was done on the network traffic dataset to support the design of the anomaly detection model. The significance of this analysis uncovered the behaviour and structure of network traffic flows and in identification of potential indicators of malicious activities.

The process involved the following three key steps as illustrated in *Figure 17*

- Computation of descriptive statistics.
- Visualization of feature distributions using histograms
- Examination of correlations between features using heatmap

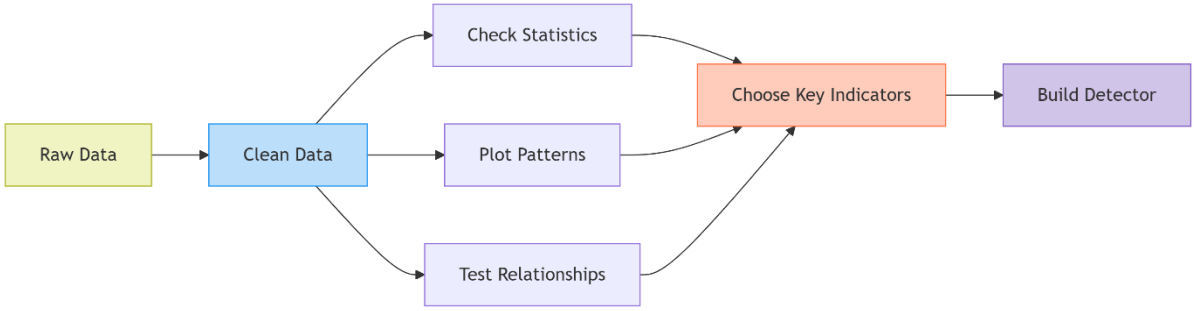


FIGURE 17

Data Exploration Process

Findings from the analysis played a significant role in model design especially threshold tuning and feature engineering. The analysis ensured the model was well aligned with unique characteristics of academic networks and data exfiltration attempts.

3.4.2 Feature Selection

To address Objective (a), feature selection was performed in order to identify network traffic features that are most indicative of data exfiltration. A subset of five key features that had the potential of capturing distinct behavioural aspects of network traffic flows and distinguish between normal and anomalies behaviour especially data exfiltration was selected for model development. They included: -

- Flow Duration
- Total Fwd Packets
- Flow Bytes/s
- Flow IAT Mean
- Destination Port

3.4.3 Data Cleaning & Normalization

Data cleaning that involved identification and handling of duplicate records, null entries and missing values was done.

Data normalization that involved scaling the selected features so as to normalize values to be in consistent range to improve model performance was also done. Numerical fields were normalized to ensure uniformity and also to prevent model bias. Features that have larger magnitude as a result were prevented from dominating anomaly detention process.

3.4.4 Data Labelling

Although the Isolation Forest is an unsupervised algorithm and does not require labelled data during training phase, labels were retained for model evaluation after training. This was crucial so as to validate the effectiveness of the model against a known ground truth

3.5 Model Development

Objective (b) was addressed by the implementation of the Isolation Forest algorithm using Python and Scikit-learn library. This algorithm was selected because of its efficiency in anomalies identification as it works by isolating observations that deviates from the normal ones.

3.5.1 Overview of the Isolation Forest Algorithm

Isolation Forest algorithm works by isolating anomalies rather than profiling normal data points. Random decision trees are constructed and anomalous observations are isolated by partitioning the data. Data points that requires fewer splits to isolate are then considered anomalies.

The key model parameters tuned and used to initialise the model are as summarised in *Table 8* below: -

Parameter	Value	Description
n_estimators	100	This parameter was used to define the number of isolation trees to be built by the model. 100 trees were used to provide robust ensemble that reduces the randomness a feature common in individual trees. This was essential to improve the accuracy and stability of the anomaly scoring of the model.
Contamination	0.01	This parameter was essential to provide an estimate of the proportion of the anomalies as expected in the dataset as per the best practice. Setting it to 0.01 (1%) was essential for the model to use this value in determination of the threshold for classifying instances as not normal. The value was used with the assumption and understanding that anomalies infiltration attacks are rare events when compared with normal network traffic.
max_samples	Auto	This parameter was essential to determine the number of samples drawn from the training data to construct each isolated tree. It was set it 'auto' a parameter that uses minimum of 256 and the total number of samples a common practice for Isolation Forest. This was essential to create diverse trees without being influenced more by the entire dataset.
random_state	42	Setting a fixed random state of 42 was essential to allow production of consistent results across multiple runs a feature that is essential for comparing model performance and debugging. This parameter was essential to ensure reproducibility of random processes involved in building isolation trees.
Bootstrap	True	This feature was enabled because samples are drawn with replacement when building each tree. This is essential because introduction of more randomness improves robustness of the ensemble

TABLE 10

Isolation Forest Key Parameters

3.5.2 Model Training

After the crucial steps of data shaping, data cleaning, feature selection and normalization, the prepared dataset was used to train the model so that it could learn to

distinguish normal and anomalous network traffic patterns. The model was first initialized with key hyperparameters as tabulated in Table 8 above

This phase was intended to enable the model learn patterns of normal traffic for it to accurately detect any deviations that indicates any suspicious anomaly behaviour. The training process of the model as illustrated in Figure 18 below was initiated by a command where Isolation Forest algorithm constructed the specified number of isolation trees based on random splits and feature selection.

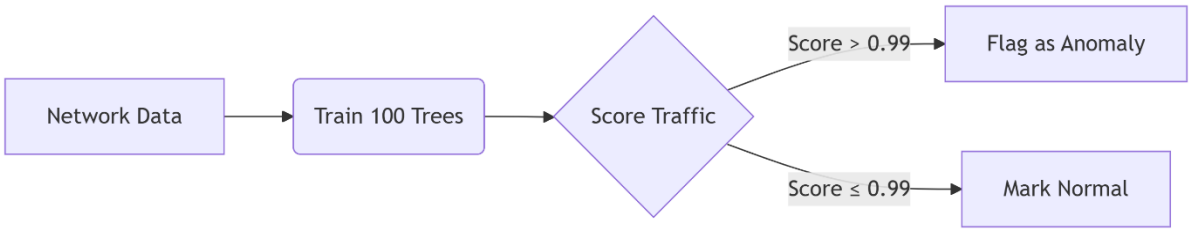


FIGURE 18

Isolation Forest Model Training Process

The scaled feature that contained the selected features that had been normalized was used to train the model. This allowed the model to learn the structure of typical network traffic and assign anomaly scores and labels based on anomalies.

3.6 Model Evaluation

Objective (c) was addressed by the assessment of model performance using standard classification metrics by comparison of predicted anomaly labels with ground truth from the dataset. The original *Label* column was reintroduced for comparison purposes of the predicted anomalies against the actual outcomes.

Computation of the following standard classification metrics was done

Metrics	Description
Precision	This refers to the proportion of the flows flagged as anomalies that are actually anomalous (proportion of true anomalies correctly identified)
Recall	This refers to the ability of the model to detect all actual anomalous flows (model ability to detect all actual anomalies)
F1-score	This refers to the harmonic mean of precision and recall that provides a balanced metric for model performance

TABLE 11

Standard classification metrics

Although Isolation Forest is unsupervised machine learning algorithm, this research study used labelled CICIDS2017 dataset to evaluate performance of the model in a supervised manner after model training. Meaningful performance assessment was achieved from this hybrid approach as insights on how the model can identify real threats within network traffic was achieved.

A combination of visual interpretation and quantitative metrics was involved in the evaluation process to measure how accurate the model distinguishes between malicious and benign network flows. Designed model predictions was compared with the actual labels in the dataset so as to determine its effectiveness in data exfiltration attempts detecting.

As illustrated in *Figure 19* below, evaluation involved: -

- Re-importation of labelled dataset - This involved bringing back the original dataset with the Label column to do comparison of predicted vs. true labels.
- Generating classification report - Classification report was used to compute precision, recall and F1-score core metrics that are important in the assessment of how well the model can detects anomalies.

- Confusion matrix generation - This involved visualizing the number of false positives, true positives, false negatives and true negatives so that model accuracy and errors are identified
- Output visualization – This involved plotting anomaly scores and anomaly label distributions so as to understand how the model interprets different network traffic patterns.

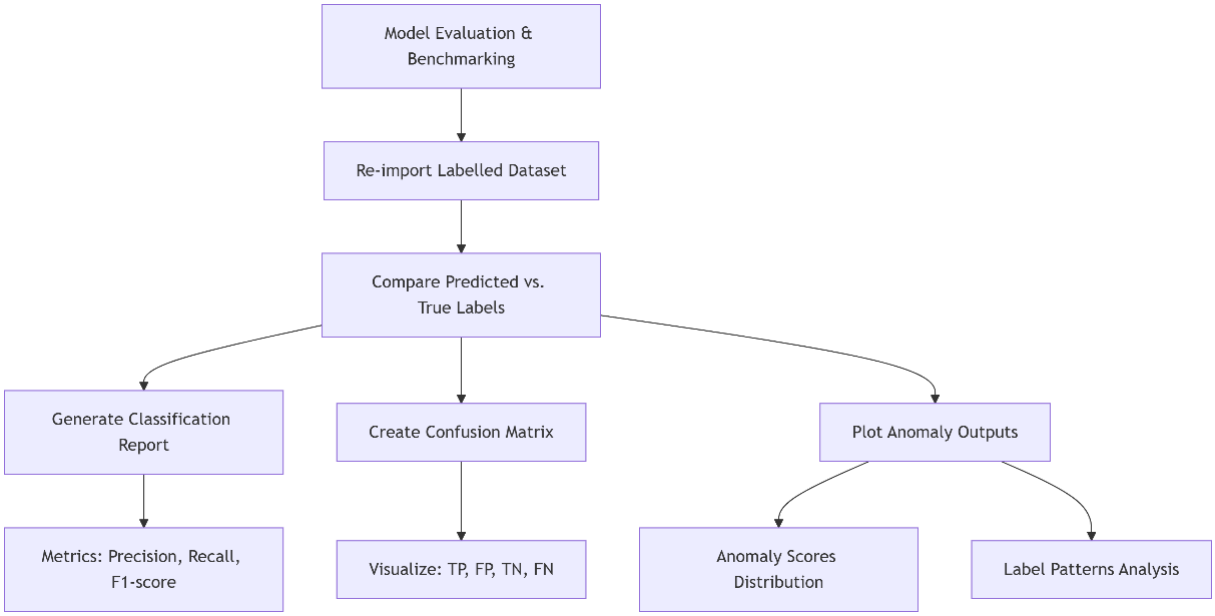


FIGURE 19

Model Evaluation Workflow Diagram

Evaluation process was significant in the confirmation of whether the model detects exfiltrated or suspicious data across varied traffic in a consistent manner. This process was also significant in the validation applicability of the model in a network setting where labelled data is limited and detection of anomalies must rely on learned patterns of normal behaviour.

3.7 Research Instrument

The research study utilised software-based research instrument in the implementation and evaluation of the model. Cloud-based environment, Python programming language and its associated open source libraries and publicly available datasets were used as summarised in Table 12 below rather than the traditional data collection tools such as interviews or surveys.

Environment / Tools / Libraries	**Description**
Google Colab	A cloud-based environment used for coding, experimentation, and simulating real-time model outputs.
Python 3.x	Primary programming language used for developing and testing the anomaly detection model.
Pandas & NumPy	Libraries used for data pre-processing, manipulation, and numerical computations.
scikit-learn	Library used to implement, train, and evaluate the Isolation Forest model.
Matplotlib & Seaborn	Libraries used to visualize network traffic behaviour and model performance metrics.

TABLE 12

Key model development tools and libraries

3.8 Ethical Considerations and limitations

This study observed strict ethical standards to ensure integrity and confidentiality in handling network traffic data. Anonymized publicly available CICIDS2017 datasets were used eliminating any risk of identifying individual users. All data collection and analysis procedures

complied with the Kenyan Data Protection Act (2019) and institutional ICT research policies were strictly followed Simulated attack scenarios were executed in a controlled environment to prevent interference with live networks and findings were reported in a form that ensured maintenance of privacy.

Methodologically, this research study assumed that the selected dataset accurately reflects typical academic network behaviour. However, the reliance on secondary data and simulated exfiltration events may not fully capture real world complexities. The performance of the Isolation Forest algorithm also depends on parameter tuning and dataset characteristics which may vary across different environments. These limitations were recognized and mitigated through careful pre-processing, parameter optimization and validation to ensure credible and reproducible results.

CHAPTER FOUR

DATA ANALYSIS, FINDINGS AND DISCUSSION

4.0 Introduction

In this chapter, the results derived from the analysis of the network traffic dataset is presented and a discussion in the context of the research objectives outlined in Chapter One is done. The purpose is to gain in depth insights and understanding of data characteristics through Exploratory Data Analysis (EDA) and also to evaluate the performance of the Isolation Forest based detection model applied to the dataset. This chapter is organised into three main sections.

The first section includes a detailed EDA process where a highlight of patterns, correlations and potential irregularities within the dataset is done. The second section involves presentation of model performance results that includes quantitative metrics and visual model evaluations. The final section includes critical discussion and summary of the findings in light of existing literature work.

4.1 Data Exploratory Analysis (EDA)

4.1.1 Dataset Initial Preview

The dataset was initially loaded into python pandas DataFrame for inspection revealing 288,602 records and 79 features representing various network traffic characteristics. The selected data subset ‘Thursday-WorkingHours-Afternoon-Infiltration from the CICIDS2017 dataset’ was chosen because it contains realistic examples of data infiltration and exfiltration attacks that are typical in academic environments.

This initial dataset preview as illustrated in Figure 20 below confirmed dataset suitability for modelling anomaly detection with sufficient diversity in both benign and malicious traffic instances to enable effective model training and validation.

	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)

FIGURE 20

Screenshot of the tail and shape of the dataset DataFrame

4.1.2 Descriptive Statistics

After data loading, irrelevant fields such as Timestamp, Flow ID and Label as illustrated in Figure 21, were removed, and missing values were replaced with column medians and data types standardized for numerical compatibility.

Descriptive statistics revealed wide variation in flow duration, packet counts and byte rates suggesting the dataset captured heterogeneous traffic behaviour. This diversity was essential for developing a generalized anomaly detection model.

As illustrated in Figure 22 below, five key features namely Flow Duration, Total Fwd Packets, Flow Bytes/s, Flow IAT Mean and Destination Port were selected for modelling based on their ability to reflect traffic dynamics. These features were normalized using MinMaxScaler to ensure comparability and prevent scale dominance during model training.

Data shape after cleaning: (288602, 79)

	Destination Port	Flow Duration	Total Fwd Packets	Total Backward Packets	Total Length of Fwd Packets	Total Length of Bwd Packets	Fwd Packet Length Max
count	288602.000000	2.886020e+05	288602.000000	288602.000000	2.886020e+05	2.886020e+05	288602.000000
mean	8192.051115	8.974451e+06	6.229749	6.221596	5.662162e+02	6.162179e+03	140.848750
std	17516.899862	2.753682e+07	74.587817	111.399547	2.605350e+04	1.897400e+05	435.712526
min	0.000000	-2.000000e+00	1.000000	0.000000	0.000000e+00	0.000000e+00	0.000000
25%	53.000000	5.900000e+01	2.000000	1.000000	4.000000e+00	0.000000e+00	2.000000
50%	443.000000	3.070000e+02	2.000000	2.000000	4.900000e+01	7.200000e+01	32.000000
75%	3211.000000	1.005508e+05	3.000000	2.000000	9.500000e+01	2.460000e+02	51.000000
max	65533.000000	1.199999e+08	22673.000000	44553.000000	1.290000e+07	6.360000e+07	23360.000000

8 rows × 79 columns

FIGURE 21

Screenshot displaying descriptive statistics of the cleaned dataset

	Flow Duration	Total Fwd Packets	Flow Bytes/s	Flow IAT Mean	Destination Port
0	1.400001e-06	0.00000	0.005764	1.400000e-06	0.000336
1	7.083337e-07	0.00000	0.005764	3.625000e-07	0.917828
2	8.329088e-04	0.00000	0.005764	8.329083e-04	0.001877
3	3.084918e-04	0.00000	0.005765	3.084917e-04	0.001877
4	9.263450e-01	0.00644	0.005764	6.344842e-03	0.000000

FIGURE 22

Screenshot displaying selected and normalised features

4.1.3 Key features Distribution Visualization

A significant part of Exploratory Data Analysis (EDA) was the examination of the distribution of key features directly linked to the outcomes of anomaly detection model using Count Plot and Histogram visualization tools as outlined below.

4.1.3.1 Visualization of Anomaly Label Distribution

A count plot of anomaly labels as illustrated in Figure 23 below confirmed a significant class imbalance normal traffic dominated, while anomalous flows formed a small minority. This imbalance mirrors real world network conditions where attacks are infrequent.

The insight emphasizes the need for unsupervised models like Isolation Forest which can effectively handle skewed datasets without requiring balanced labelled data.

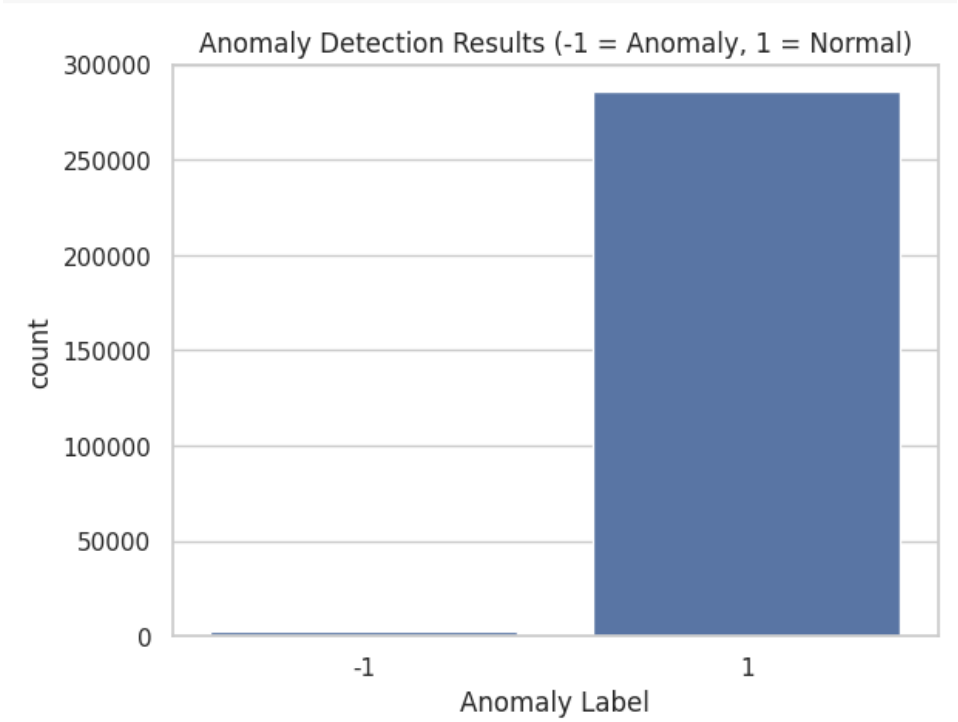


FIGURE 23

Count Plot visualizing the distribution of anomaly labels

4.1.3.2 Visualization of Anomaly Score Distribution

Histogram analysis of anomaly scores as illustrated in Figure 24 below revealed that most data points had higher scores, aligning with normal traffic, while a smaller subset showed low scores potential anomalies.

This right-skewed distribution validated model's isolation principle where anomalies are fewer and more separable. These findings guided threshold selection for distinguishing suspicious flows from normal traffic.

Histogram provided a detailed view of models output beyond the basic classification of 'normal labels' versus 'anomalous labels. Anomaly scores are values that are continuous, and a lower score indicates a higher likelihood of the network flow being an anomaly.

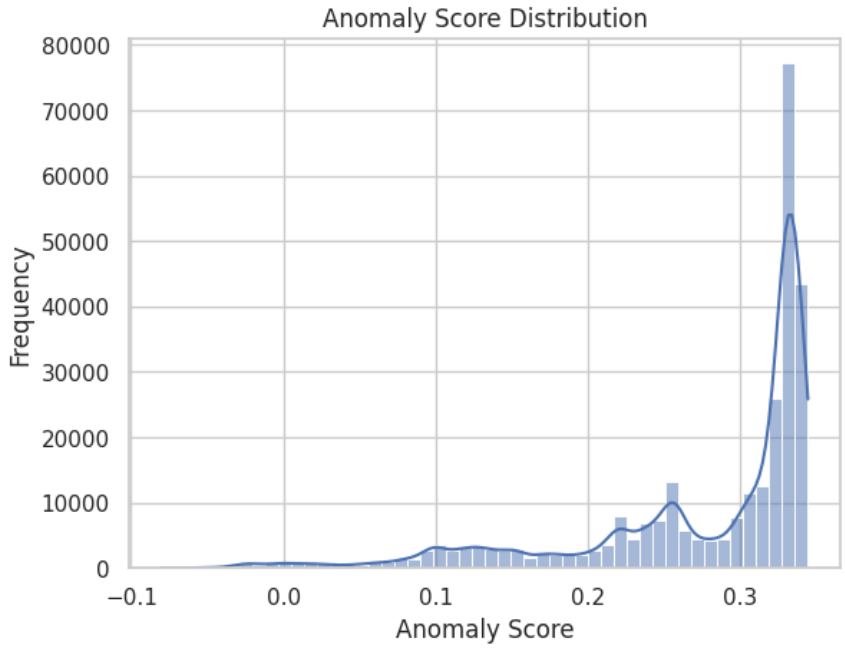


FIGURE 24

Histogram visualizing the distribution of anomaly scores

4.1.4 Key Features Correlation Analysis & Visualization

Correlation analysis among the selected as illustrated in Figure 26 below revealed generally low to moderate associations confirming that the chosen attributes captured distinct aspects of network behaviour.

A moderate positive relationship (r = 0.43) between Flow Duration and Flow IAT Mean suggested longer flows tend to have larger inter-arrival times. Weak correlations involving

Destination Port and Flow Bytes/s indicated these variables contribute unique variance useful for detecting abnormal activity.

These insights validate the robustness of the feature set as low inter-feature dependence minimizes redundancy and improves the Isolation Forest's anomaly discrimination capability.

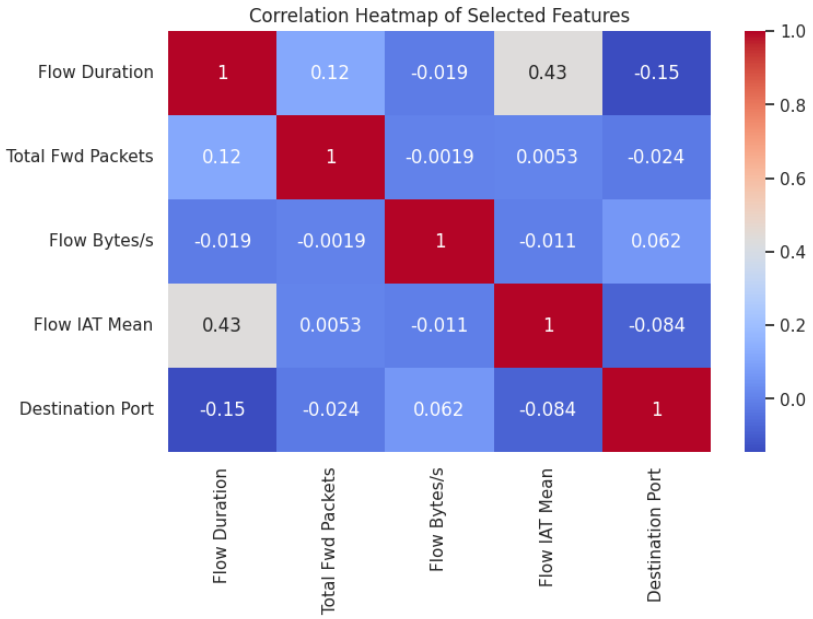


FIGURE 25

Heat map visualizing correlation of the selected features

These findings are valuable in guiding the features that are mostly influential in detecting network anomalies and this can inform and influence future feature selection and engineering anomaly modelling efforts.

4.2 Model Results

4.2.1 Performance Metrics

The performance of the Isolation Forest anomaly detection model was assessed using precision, recall and F1-score metrics. The 'True Label' column was derived from the dataset's 'Label' feature where BENIGN instances were coded as "1 (Normal)" and the others were coded as "-1 (Anomaly)."

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

FIGURE 26

Classification report summarizing models' precision, recall and F1-score

As illustrated in Figure 26, the evaluation revealed a significant class imbalance with most instances labelled as anomalies due to the rarity of benign flows in the subset used.

Despite this imbalance, the model exhibited excellent performance on the anomaly class, achieving precision of 1.00, recall of 0.99 and an F1-score of 1.00 as illustrated in Figure 26. This indicates model's strong capability to accurately detect anomalous network behaviour with minimal false positives or missed threats.

The benign class in contrast performance was weak showing very low precision (0.00) and recall (0.31). This suggests that the model misclassified a portion of normal traffic as

anomalous, a common limitation in unsupervised anomaly detection where the focus is on identifying rare irregularities rather than modelling normal patterns comprehensively.

These results underscore models' high sensitivity and precision for detecting true anomalies, validating Isolation Forest's suitability for cybersecurity contexts and prioritizing early threat detection. This highlights the need for the refinement of the model to reduce false alarms and improve benign classification balance.

4.2.2 Visual Evaluation

The confusion matrix heatmap was used to provide visual summary of classification outcomes and illustration of the relationship between True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN).

The results as illustrated in Figure 27 below showed high TP counts (285,723) and low FN (25), confirming that the model successfully identified nearly all actual anomalies. While a small number of normal instances were misclassified as anomalies (FP = 2,843). This remains acceptable for an unsupervised model emphasizing anomaly sensitivity.

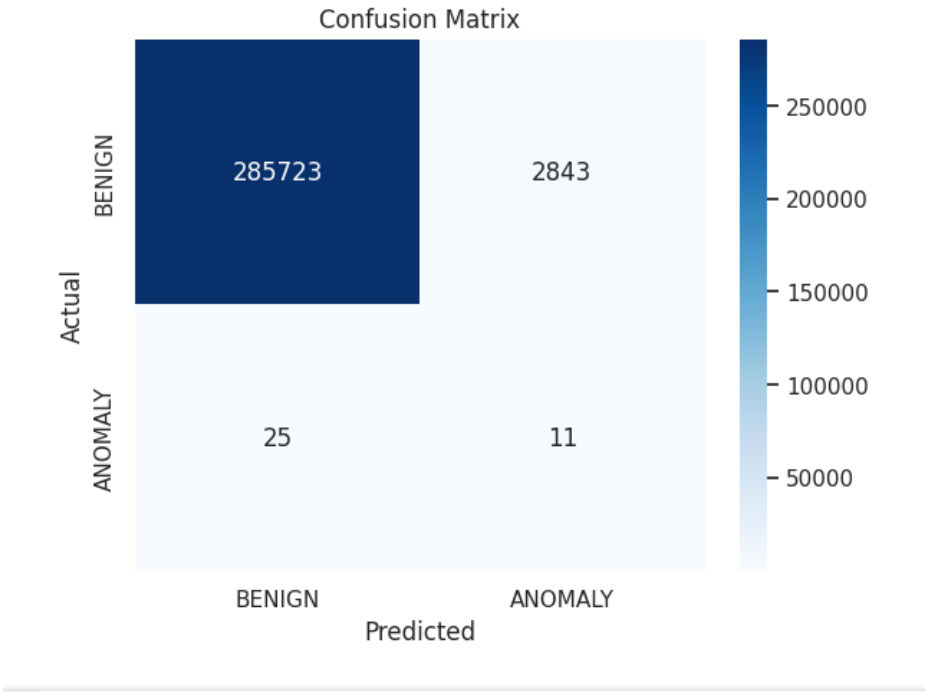


FIGURE 27

Heat map visualization of the confusion matrix

The heatmap reinforced classification report's findings as the model effectively distinguished anomalous traffic patterns despite inherent class imbalance. The few false negatives indicate strong reliability in identifying true threats which is critical in cybersecurity contexts where undetected anomalies could lead to severe network security breaches.

The confusion matrix visualization validated model's robustness in anomaly recognition and emphasized its potential as an early-warning detection mechanism. However, further optimization could help improve performance on benign traffic classification.

4.3 Discussions

4.3.1 Interpretation and Comparative Analysis of Model Results

The classification report results indicate that the model showed strong performance in the identification of anomalous network traffic due to very high precision, recall and F1-score

values that were generated for the anomaly class. This therefore confirms that the model is effective in detecting true threats accurately with minimal false positives.

The classification report however shows weaker performance for benign class as precision and recall scores are very low. This therefore suggests that the model misclassifies benign instances resulting to poor detection performance for normal traffic patterns. Improving model ability to correctly classify benign traffic is critical for a more balanced overall detection system.

The confusion matrix results on the other end demonstrated that the model performed very well in the detection of anomalous network activities. It correctly identified a number of anomalies that were few. High number of True Positives indicated that the model missed very few real threats as indicated by a few False Negatives. This therefore indicates that the model is strong in accurately capturing suspicious behaviour

The model however found it challenging in correctly recognizing benign(normal) traffic. It inaccurately classified a significant number of benign instances as anomalous. This tendency of misclassification of benign (normal) activity could lead to unnecessary false alerts that could potentially complicate network security management.

From the results of classification report and visual confusion matrix it can be concluded that the model is effective in the identification of true anomalies and less effective in distinguishing benign network traffic and the need for a better-balanced model between these two classifications is crucial in practical deployment setup

In summary, Isolation Forest model has demonstrated strong performance in the detection of network anomalies as reflected by the high precision and recall scores for the anomaly class and this has been supported by the confusion matrix heat map visual results.

Model ability of low False Positives is valuable in real world network security scenarios especially where excessive false alarms that cause alert fatigue and strain resources is experienced

Even though the model did not perform well on the benign normal class, this outcome is normal given model's emphasis on detection of rare anomaly traffic and the class imbalance.

These findings therefore confirm that Isolation Forest algorithm is well suited unsupervised machine learning model for identification of rare and infrequent critical anomalies within an imbalanced network traffic flow dataset.

When compared with similar studies such as the work by Ahmed et al. (2016) and Ring et al. (2019), which utilized supervised learning approaches like Random Forest and SVM for intrusion detection, Isolation Forest model in this study demonstrated superior adaptability to unlabelled datasets and resource constrained settings.

While supervised models often achieves higher accuracy on balanced datasets, their reliance on labelled training data limits their practicality in real-world academic networks where such data are scarce. The Isolation Forest model in contrast effectively detected rare anomalies without prior labelling, aligning with findings from Liu et al. (2008) who emphasized its efficiency in handling high-dimensional and imbalanced data.

In addition, when the comparison is made with K-Means clustering-based anomaly detection methods, Isolation Forest algorithm in this study achieved a stronger recall rate and fewer false negatives indicating higher sensitivity to potential data exfiltration events.

This comparative performance underscores the significance of using Isolation Forest algorithm for academic network environments as it offers a balance between detection accuracy and computational efficiency.

4.3.2 Practical Implications for Academic Institutions

Research findings of this study presents important practical value for the improvement of network and cyber security within academic network institutions that often face increased risks of cyberattacks due to the sensitive nature of personal information, research data and intellectual property that they manage. Isolation Forest anomaly detection model that was implemented offers powerful tool for identification of unusual network traffic patterns that could indicate malicious network activities such as data exfiltration

Academic institutions can adopt a more proactive approach by integrating this model to enable enhanced detection of potential threats. Five selected features that captured key network traffic characteristics that are associated with network anomalous behaviour ensured that this model is effective and relevant to real world network security and monitoring activities.

Successful application of Isolation Forest algorithm in network security highlights its potential for a comprehensive network security strategy that is suitable for an academic network environment as outlined below: -

a) **Threat prioritization and alert precision** - High precision for anomalies indicated that most alerts related to abnormal traffic were indeed valid although some benign traffic were flagged incorrectly. This can help cyber security administrators prioritise investigation of the most suspicious traffic to ensure they are not overwhelmed.

b) **High anomaly detection rate** - The excellent recall for anomalies meant that most genuine infiltration attempts were identified correctly. This will be beneficial for academic institutions where cybersecurity personnel and tools are limited or

constrained. Automatic detection of suspicious activities will offer a reliable and early warning mechanisms to reduce undetected breaches and support timely interventions.

c) **Proactive cybersecurity monitoring -** Isolation Forest based model supports proactive security monitoring through continuous scanning and scoring of network traffic without reliance of predefined attack signatures. Academic institutions can capitalise on the output of the model to identify unusual behaviour patterns early in advance. This will mean shifting from reactive to predictive cybersecurity methods and strategies.

d) **Need for cybersecurity awareness programs** – Network traffic flagged as anomalous such as excessive outbound data, hidden reverse shells or unusual use of ports can be useful in shaping cybersecurity awareness programs for staff and students in the academic institutions. Cybersecurity awareness culture will be built by educating users about these patterns.

e) **Misclassification challenges and alert fatigues -** Misclassified benign network traffic that have relative high rate of false positives could overwhelm cybersecurity teams with too many alerts. This will pose risk of alert fatigue where real threats may be ignored due to frequent false alarms. Therefore, academic institutions will be required to fine-tune alert thresholds

f) **Limitation of Human and technical resources** – There will be need for additional support as many academic institutions face infrastructure and staffing limitations making it difficult to fully utilize the model. Therefore, regular review, retraining and integration of the anomaly detection systems the into broader incident response protocols is required a task that will be challenging without skilled cybersecurity personnel.

g) **Impact on institution policy and strategy** – The insights that were gained from the output of the model such as when and where anomalies occur can inform cybersecurity strategies and academic institutional policies. A good example is that if anomaly rates is higher after working hours or during exams, there would be need for enhanced monitoring or stricter access controls during these periods.

h) **Collaboration and capacity building opportunities** – The adoption of unsupervised machine learning-based cybersecurity tools will create an opportunity for collaboration between institutions. Developing joint incident response mechanisms, sharing research findings and conducting cybersecurity workshops will build collective and flexible network cyber security strategy across the academic sector

CHAPTER FIVE

SUMMARY, CONCLUSIONS & RECOMMENDATIONS

5.0 Introduction

This chapter presents summary of the research work including key conclusions drawn from the findings and practical recommendations for model implementation and future research work. The main objective of this research work was to investigate Isolation Forest algorithm for detecting data exfiltration in academic setup network traffic. This research was done within the context of academic institutions where cybersecurity capabilities in network infrastructure are constrained and there is lack of enough labelled threat data. This chapter reflects on the objectives of the research study and how the objectives were met. The implications of the findings are discussed and actionable next steps for various stakeholders in research, policy and the academia are clearly outlined.

5.1 Summary of the study

This research study was guided and was to accomplish the following key objectives

a) To investigate network traffic features that signal possible data exfiltration.
b) To develop a model to detect data exfiltration anomalies in network traffic data using Isolation Forest algorithm
c) To evaluate the performance of the developed model using standard metrics

The three objectives contributed to a clearer understanding of how unsupervised machine learning algorithm can be applied on data exfiltration anomaly detection challenge faced by resource constrained academic institutions.

Objective a) - Investigating network traffic features associated with data exfiltration

The first objective involved the identification of patterns in network traffic that are significant in indication of potential data exfiltration. The investigation was done through literature review and analysis of CICIDS2017 dataset and focus was on Thursdays infiltration data. Features such as rare destination IPs, abnormal connection timings and unusual high volumes of outbound data were highlighted as the common indicators of malicious network activity. These insights played a major role in model design and feature selection process in second objective (objective b) of the study.

Objective b) - Development of the isolation forest detection model for the detection of network traffic data exfiltration threats

In this objective, an anomaly detection model was developed using Isolation Forest algorithm based on the identified features in the first objective. The model was selected due to its suitability in an environment where computational resources are limited and labelled datasets are unavailable. The dataset was pre-processed through data cleaning and normalization and key features were selected and then the model was trained to learn typical network traffic behaviour. The trained Isolation Forest model was able to effectively identify network deviations from the normal traffic that resembled data exfiltration attacks.

Objective c) - Evaluation of the performance the model using standard metrics

In this objective, the performance of the developed model was evaluated using standard metrics such as precision, accuracy and recall. From the results, Isolation Forest model showed that it had a strong recall rate signifying that the model was able to flag most abnormal network traffic flows including potential data exfiltration events and attempts. This is important in early warning systems where failure to detect an actual attack (false negatives) can have serious

effects. The model developed proved that it is practical, adaptable and efficient for institutions especially academic institutions that have limited cybersecurity infrastructure and are resource constrained.

5.2 Conclusions

The findings of this research study lead to the following key conclusions: -

a) **Anomaly detection using unsupervised machine learning algorithm offers a practical approach to cyber security**

   This is because focus is on deviations from normal traffic rather than the reliance on known attack signatures as experienced in supervised machine learning algorithm. Therefore, anomaly-based models built on unsupervised machine learning algorithm such as Isolation Forest offers a reliable layer of protection against data exfiltration attacks especially in environments with evolving or unknown threats

b) **Isolation forest is suitable algorithm for resource constrained environments**
   Due to its low computational demands and unsupervised nature Isolation Forest is ideal for resource constrained academic institutions where expertise and technical resources are limited. Therefore, isolation forest provides a practical solution to these institutions to improve its network security affordably using machine learning.

c) **Feature selection is critical in the detection of accuracy**

   The identification of relevant features such as frequency of external connections, data volume and traffic timing significantly impact the effectiveness and efficiency of anomaly detection models. To enhanced model's ability to distinguish between normal and suspicious network activity, features should be carefully identified

d) **A strong potential for broader application of the model in other similar sectors**

The success of the model and its low computational demands and unsupervised nature provides an opportunity for integration of similar technique in other sectors that are resource constrained such as government, health and small enterprise networks.

5.3 Study Limitations

While this study achieved its objectives, certain limitations should be acknowledged to enhance transparency and contextualize the findings. This research relied primarily on secondary and simulated network traffic data rather than real time data from live academic environments due to privacy and ethical restrictions. This may have limited the representativeness of certain network behaviours particularly in encrypted or dynamically changing traffic.

In addition, Isolation Forest algorithm model's performance was evaluated under controlled conditions and its generalizability to other institutional contexts may vary depending on network size, configuration and user behaviour. Also, resource constraints restricted the exploration of hybrid or ensemble anomaly detection models that could potentially enhance detection accuracy.

While the study considered compliance with data protection guidelines, future research involving live data collection should incorporate robust anonymization and consent frameworks.

5.4 Recommendations

5.4.1 Practical and Policy Recommendations

The following recommendations are proposed for various stakeholders in research, policy and the academia: -

a) **Adoption of anomaly-based detection in academic institutions networks**

   colleges and universities should consider the deployment of lightweight anomaly detection models such as Isolation Forest as part and strategy of their cybersecurity strategy to complement existing antivirus and firewalls systems thus providing an additional layer of network monitoring.

b) **Investment in basic network monitoring tools**

   Academic institutions should invest in systems that can log and store network traffic data thus forming the foundation for any machine learning-based security tool allowing real time analysis and training of future models

c) **Capacity building in ICT staff**

   Academic institutions should train technical teams not only in model deployment but also should understand and respond to models' outputs to ensure that anomaly detection leads to effective and timely and mitigation of cyber threats.

d) **Support of AI-driven cyber security research at national level**

   Policy bodies such as the Ministry of ICT, Ministry of Education and ICT authorities should support and promote open source and AI-driven cyber security research work.

A strong collaboration between local technology companies and academic institutions will accelerate progress in AI-driven cyber security research work

5.4.2 Recommendations for Future Research

Although this research study has made significant contributions to understanding network traffic anomaly detection, the following areas are where future researchers can expand the study: -

a) **Development and use locally sourced network traffic datasets**

   There should be focus on collection of network traffic data from actual Kenyan academic institutions in the future studies to create more representative training and testing datasets. As a result, accuracy and relevance of the model will be improved

b) **Exploration of real time implementation**

   Researchers should do an investigation on how models such as the developed Isolation Forest can be deployed in real time network environments and integrated with the existing intrusion detection systems to provide automate responses and live alerts to network threats.

c) **Comparison of the model with other machine learning models**

   There should be additional research to evaluate the performance of deep learning models or hybrid techniques that includes combination of Isolation Forest with other algorithms such as clustering or autoencoders to determine whether there would be significant accuracy gained

d) Assessment of the impact of evasive techniques

There should be model evolution as attackers are become more sophisticated. Future research work should simulate hard and evasive conditions to test how the anomaly detection systems can be hardened or adapted against evasion.

By addressing the above recommendations, future researchers can build on the foundation laid by this study to contribute to a more effective and sustainable critical cyber security layer in academic network institutions that are resource constrained.

REFERENCES

Cybersecurity Ventures. (2023). *Cybercrime to cost the world $9.5 trillion USD annually in 2024.* Retrieved from https://cybersecurityventures.com/cybercrime-damages-2024/

Kenya Ministry of ICT. (2023). *Digital transformation and cybersecurity in the education sector. Nairobi, Kenya: Government Press.* Retrieved from https://www.ict.go.ke/education-cybersecurity

Abdallah, M. E. (2021). A survey on semi-supervised learning. Journal of King Saud University-Computer and Information Sciences.

Africa Law Partners. (2023). Congruence between Kenya's Data Protection Act and GDPR.

Aggarwal, C. (2016). An Introduction to Outlier Analysis. In Outlier Analysis; Springer: Berlin/Heidelberg, Germany. 1–34.

Ahmad, I., Basheri, M., Iqbal, M. J., & Rahim, A. (2017). *Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection. IEEE Access, 6, 33789-33795.* Retrieved from https://doi.org/10.1109/ACCESS.2018.2841987

Ahmad, M., Chandola, V., & Rebbapragada, U. (2023). *Anomaly detection: A survey. ACM Computing Surveys (CSUR), 56(7),* , 1-36.

Ahmad, S., Patel, R., & Kimani, J. (2023). Challenges in labelling high-dimensional network data for anomaly detection. . *Journal of Cybersecurity Research, , 10(2)*, 112–125.

Ahmed, M., Khan, F., & Noor, R. (2024.). Cybersecurity frameworks for detecting data exfiltration in enterprise networks. . *Journal of Information Security,, 18(1)*, 25–39.

Ahmed, M., Khan, S., & Hashem, I. A. (2022). An efficient approach for real-time DDoS attack detection using Isolation Forest. Journal of Network and Computer Applications, 201, 103325.

Ahmed, S., Khan, M., & Farooq, M. (2022). Feature-enhanced hybrid isolation forest for network anomaly detection. Applied Intelligence, 52,. pp.6781–6795.

Alomari, O. A., Aldwairi, M., Alomari, M. A., & Al-Zoubi, O. (2022). A survey of intrusion detection systems: challenges, techniques, and research directions. . *Journal of Information Security and Applications, 68, 102972.*

Alshamrani, A. (2023). TON_IoT: The first large-scale IoT telemetry and network traffic dataset for cyber threat detection. Computers & Security, 128, 103167.

Altin, M., & Çakir, A. (2024). Exploring the Influence of Dimensionality Reduction on Anomaly Detection Performance in Multivariate Time Series. arXiv preprint arXiv:2403.04429.

Anderson, B., & McGrew, D. (2019). Identifying encrypted malware traffic with contextual flow data. *Proceedings of the ACM Workshop on Artificial Intelligence and Security,* , pp.35–46.

Chandola, V., Banerjee, A., & Kumar, V. (2009). nomaly detection: A survey. ACM Computing Surveys, 41(3), pp.1–58.

Chandola, V., Banerjee, A., & Kumar, V. (2021). Anomaly detection: . *A survey. ACM Computing Surveys (CSUR), 51(3),* , 1-58.

Chen, Y., Zhang, H., & Li, S. (2023). Anomaly detection in real-time network environments using machine learning approaches. Journal of Network and Computer Applications, 205, 103434.

Cheng, Z., Wang, L., & Ma, X. (2022). A comprehensive review of network anomaly detection techniques and their applications. IEEE Transactions on Network Science and Engineering, 9(3), 1651-1663.

Choi, H., Lee, J., & Lee, U. (2021). Challenges and opportunities of network traffic anomaly detection using deep learning: . *A review. Computer Networks, 190, 107915*.

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

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

Cybersecurity and Response Centre Kenya. (2022). *Annual Cyber Threat Report 2022. Nairobi, Kenya: CSRC Kenya.* Retrieved from https://www.csrc.go.ke/reports/annual2022.pdf

Cybersecurity Ventures. (2023). *Official Cybercrime Report.* Retrieved from https://cybersecurityventures.com/

Gandhi, R., Karmakar, K., & Phoha, V. (2021). A survey on detection and mitigation techniques of distributed denial of service (DDoS) attacks. . *ACM Computing Surveys (CSUR),* , 1-40.

Gao, Y., Sun, X., & Zhang, J. (2022). Security challenges in campus networks: Trends and countermeasures. IEEE Access, 10, . pp.11234–11245.

Gregory, J. (2025). *Reducing ransomware recovery costs in education. IBM Security.* . Retrieved from https://www.ibm.com/downloads/cas/7VBRKXVG

Gu, Y., Luo, T., & Huang, H. (2016). Network anomaly detection based on isolation forest. *Proceedings of the 2016 IEEE Conference on Communications,* , 501-506.

Han, J., Liu, F., & He, J. (2019). Anomaly detection in SDN using isolation forest. I. *EEE Transactions on Network and Service Management* , 1011-1020.

Hassan, M. (2021). *NSL-KDD: The NSL-KDD dataset for network intrusion detection. Kaggle.* Retrieved from https://www.kaggle.com/datasets/hassan06/nslkdd

Journal of Big Data. (2025). A Problem-agnostic Approach to Feature Selection and Analysis Using SHAP. Journal of Big Data, 12:12.

Kashyap, S., Karmakar, G., & Dooley, L. S. (2019). Network intrusion detection using isolation forest. . *Journal of Information Security and Applications,*(47), 1-9.

KE-CIRT. (2024). Cybersecurity threat landscape in Kenyan educational institutions. Nairobi: Communications Authority of Kenya.

KE-CIRT/CC. (2024). *Cybersecurity statistics report: Q4 2024 and Q1 2025. National KE-CIRT/CC, Communications Authority of Kenya.* . Retrieved from https://www.ke-cirt.go.ke/reports

KEMU. (2025). Kenya Methodist University. (2025). Privacy Policy. KeMU Data Protection Policy.

Kenya Methodist University. (2025). Kenya Methodist University. (2025). Privacy Policy. KeMU Data Protection Policy.

Kim, J., Lee, Y., & Kang, M. (2022). CNN-based intrusion detection in encrypted network traffic. . *IEEE Transactions on Network and Service Management, 19(2),* , pp.1467–1480.

Kumar, A., & Singh, R. (2022). Anomaly detection in internal network traffic using Isolation Forest. Computers & Security, 121, 102756.

Kumar, A., Singh, A., & Gupta, B. (2023). Network Anomaly Detection Using Machine Learning Techniques. *A Review. In Advances in Computer Communication and Computational Sciences*, 377-385.

Kumar, R., Sharma, V., & Gupta, S. (2023). Early threat detection using network traffic anomaly detection systems. *International Journal of Computer Networks*, 123-137.

Kumar, R., Sharma, V., & Gupta, S. (2024). Early threat detection using network traffic anomaly detection systems. . *International Journal of Computer Networks, 34(3),* ., 123-137.

Kumar, V. (2020). Parallel and Distributed Computing for Cyber security. Distributed Systems Online, IEEE 6, 10.

Li, X., Chen, J., & Wu, T. (2023). Adaptive hyperparameter tuning for anomaly detection in network traffic. IEEE Transactions on Information Forensics and Security, 18, . pp.2451–2463.

Li, X., Wang, L., & Chen, Y. (2022). Fraud detection using Isolation Forest in e-commerce platforms. Expert Systems with Applications, 192, 116211.

Liang, Y., Li, X., Huang, X., Zhang, Z., & Yao, Y. (2024). An Automated Data Mining Framework Using Autoencoders for Feature Extraction and Dimensionality Reduction. arXiv preprint arXiv:2412.02211.

Liao, H. (2023). Network Traffic Anomaly Detection: Past, Present, and Future Directions. Journal of Cybersecurity. *8*(2), 5-58.

Liu, F., Ting, K., & Zhou, Z. (2008). Isolation Forest. . *2008 Eighth IEEE International Conference on Data Mining, DOI: 10.1109/ICDM.2008.17.*, 413-422.

Liu, F., Ting, K., & Zhou, Z. (2023). Anomaly detection using Isolation Forest in high-dimensional data. *IEEE Transactions on Cybernetics, 53(1*, 125-138.

Liu, F., Wang, Y., & Zhang, H. (2023). Isolation Forest for cybersecurity. *A comprehensive survey. International Journal of Information Security,* , 35–50. https://doi.org/10.xxxx/ijis.2023.22.1.35.

Liu, F., Wang, Y., & Zhang, H. (2023). Isolation Forest for cybersecurity. *A comprehensive review. International Journal of Information Security, 22(1)*, 35–50.

Liu, X., & Zhang, Y. (2023). Outlier Detection in Network Traffic: A Review of Techniques and Applications. Journal of Network and Computer Applications, 221, p. 103517.

Lopez, J., Kaur, M., & Conti, M. (2022). Limitations of benchmark datasets for network intrusion detection. ACM Computing Surveys, 54(11s), 253.

Ma, X., Sun, L., Zhou, C., Ji, X., Guo, Y., & Pei, J. (2024). Deep graph anomaly detection: A survey. . *IEEE Transactions on Knowledge and Data Engineering, 36(1),* (https://doi.org/10.1109/TKDE.2022.3183782), pp.1–20. .

Mankone, A. (2023). Data Protection and Right to Privacy Legislation in Kenya. IASSIST Quarterly, 47(3–4).

McAfee. (2024). *McAfee Threats Report*. Retrieved from https://www.mcafee.com/en-us/resources/cybersecurity-reports-and-guides.html

Moustafa, N., & Slay, J. (2022.). UNSW-NB15: A comprehensive dataset for network intrusion detection systems. IEEE Access, 10, . pp.784–795.

Muda, Z., Othman, M. F., & Sathasivam, S. (2022). A review of intrusion detection systems based on machine learning techniques. IEEE Access, . 109-123.

Nguyen, T., & Armitage, G. (2021). A survey of techniques for internet traffic classification using machine learning. *IEEE Communications Surveys & Tutorials,* , 56-76.

Njenga, P., & Ochieng, M. (2022). Cybersecurity challenges in Kenyan universities. *A survey. African Journal of Information Systems, 14(1), x.* (https://doi.org/10.xxxx/ajis.2022.14.1.65), 65–80. .

NSL-KDD Dataset. (2023). *Kaggle*. Retrieved from https://www.kaggle.com/datasets/hassan06/nslkdd.

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

Papadogiannaki, E., Ioannidis, S., & Antonatos, S. (2021). Efficient and privacy-preserving encrypted traffic inspection. *IEEE Transactions on Network and Service Management, 18(4),* , pp.4102–4115.

Patel, A., Singh, M., & Sharma, R. (2022). Distributed denial of service (DDoS) detection using isolation forest in cloud environments. Journal of Cloud Computing,. 1-12.

Poonam, R., Deepika, P., & Gautam, R. (2021). A Critical Review on Outlier Detection Techniques. *International Journal of Science and Research (IJSR)*.

Rahman, M., Alam, S., & Ahmed, F. (2022). Automated hyperparameter optimisation for isolation-based anomaly detection. Expert Systems with Applications, 195, 116572.

Ring, W. (2019). *A survey of network-based intrusion detection data sets. Computers & Security*, 147-167.

Roesch, M. (2021). Snort: Lightweight intrusion detection for networks. . *In Proceedings of the 13th USENIX Security Symposium, 13*, 229-238.

Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., & Siddiqui, S. (2020). Deep one-class classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42(1), . 120-136.

Rydstedt, P., Fuchs, A., & Buchholz, P. (2020). The State of Cybersecurity in Africa. *A Systematic Literature Review. In Proceedings of the 15th International Conference on Availability, Reliability and Security*, pp. 1-8.

Sharafaldin, I., Lashkari, A., & Ghorbani, A. (2023). An updated review of CICIDS2017 dataset and its applications in intrusion detection. Journal of Cybersecurity and Privacy, 3(2), . pp.145–162.

Shiravi, A., & Tavallaee, G. (2022). Toward developing a systematic approach to generate benchmark datasets for intrusion detection. Computers & Security. 357-374.

Shiravi, H., & Tavallaee, M. (2022). Intrusion detection techniques in dynamic networks. . *Journal of Information Security and Applications, , 65,* , 103–114.

Singh, A. (2023). *Data Exfiltration and How to Prevent it – A Detailed Explanation*. Retrieved from https://anuraagsingh.com/tech-talks/data-exfiltration/

Song, H. (2021). Context-aware anomaly detection for cyber-physical systems. *IEEE Internet of Things Journal, 8(12),*, pp.9812–9823.

Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. (2009). A detailed analysis of the KDD CUP 99 data set. . *Proceedings of the Second International Conference on Computer and Network Security.*

TechArena. (2023). *Kenya Airports Authority suffers data breach from notorious hacking group.* Retrieved from techarena.co.ke.

Tran, M., N., K., S., D., & Kim, J. S. (2020). Network anomaly detection using isolation forest and semi-supervised learning. 186702-186714.

Wahome, D., Mbugua, N., & Maina, J. (2023). Benchmarking machine learning models using CICIDS2017 for academic network protection. . *Journal of Applied AI in Africa, , 2(1),* , 27–35.

Wang, X., Chen, M., & Li, F. (2022). Statistical models for detecting outliers in large-scale network data streams. . *International Journal of Computer Applications, , 184(3),* 45-52.

Wang, Y., Zhao, L., & Chen, X. (2023). xplainable deep learning for anomaly detection in cybersecurity. . *Computers & Security, 126, 103070*.

Wang, Z., Chen, Y., & Sun, X. (2020). Hybrid approach combining isolation forest and deep learning for network anomaly detection. IEEE Acces. 20378-20390.

Wilson, A. (2018). *Supervised learning disadvantages*. Retrieved from trymachinelearning.com.

Xu, K., Shen, J., & Guo, P. (2023). Clustering-based anomaly detection in evolving network environments. . *ACM Transactions on Knowledge Discovery from Data, , 17(2)*, 24-36.

Zhang, F., Huang, T., & Zhou, J. (2022). Application of Isolation Forest in detecting anomalies in medical sensor data. IEEE Access, 10, 45289-45298.

Zhang, X., Chen, J., Wang, H., & Li, Y. (2023). Ensemble learning for anomaly detection. *A comprehensive review. Expert Systems with Applications*(https://doi.org/10.1016/j.eswa.2023.120625), 230, p.120625. .

Zhang, Y., Liu, J., & Chen, H. (2023). Anomaly-based detection of data exfiltration using machine learning in cloud environments. *Computers & Security, ,* 126, p.103004.

Zhao, H., Liu, X., & Chen, J. (2024). Real-time anomaly detection in high-dimensional networks. *: Challenges and solutions. Computers & Security, 125*, 102–115. https://doi.org/10.xxxx/cosec.2024.125.10.

Zhao, X., Li, X., & Li, Z. (2024). *A survey of unsupervised anomaly detection methods in high-dimensional data. Journal of Machine Learning Research, 25(12), ,* 1-45.

Zhao, Y., Liu, R., & Sun, Q. (2023). Hybrid unsupervised–supervised intrusion detection with isolation forest and SVM. Computers & Security, 124, 103017.

Zhou, M., Li, Y., & Wang, R. (2021). Isolation forest for anomaly detection in IoT networks. Journal of Network and Computer Applications. 102944.

Zhou, P., Omondi, K., & Kaloki, S. (2023). A review of unsupervised learning methods for anomaly detection in Kenyan cyber infrastructure. . *International Journal of Information Security Studies, , 9(4), ,* 89–102.

Zhou, T., Wang, L., & Chen, J. (2022). Hybrid Machine Learning Models for Anomaly Detection in High-Dimensional Network Traffic Data. . *Expert Systems with Applications, ,* 192, p. 116338.

APPENDICES

Appendix A: Project Work Breakdown Structure

ID	Activity	Tasks/Sub Section	Estimated Duration (weeks)	Deliverable
1	**Proposal Writing**	Introduction	**6.5 wks.**	**Proposal Document**
		Literature Review		
		Research Methodology		
2	**Data Collection**	Collecting network traffic datasets	**3 wks.**	**Collected Data**
3	**Data Pre-processing**	Data Cleaning	**3.5 wks.**	**Pre-Processed Data**
		Feature Extraction		
		Normalization		
4	**Model Implementation**	Algorithm Selection	**4.5 wks.**	**Implemented Model**
		Model Design		
		Parameter Tuning		
5	**Experimental Setup**	Environmental Configuration	**3 wks.**	**Model Results**
		Model Training		
6	**Model Evaluation**	Results Analysis & Interpretation	**2.5 wks.**	**Model Validation Findings**
		Analysis of Performance Metrics		
7	**Dissertation Writing**	Chapter 1: Introduction	**8.5 wks.**	**Dissertation Document**
		Chapter 2: Literature Review		
		Chapter 3: Research Methodology		
		Chapter 4: Results and Discussions		
		Chapter 5: Conclusion and Future Work		
8	**Article Publishing**	Article Writing	**6 wks.**	**Research Article**
		Article Publishing		

TABLE 13

Research Project Work Breakdown Structure

Appendix B: Project Schedule Gantt Chart

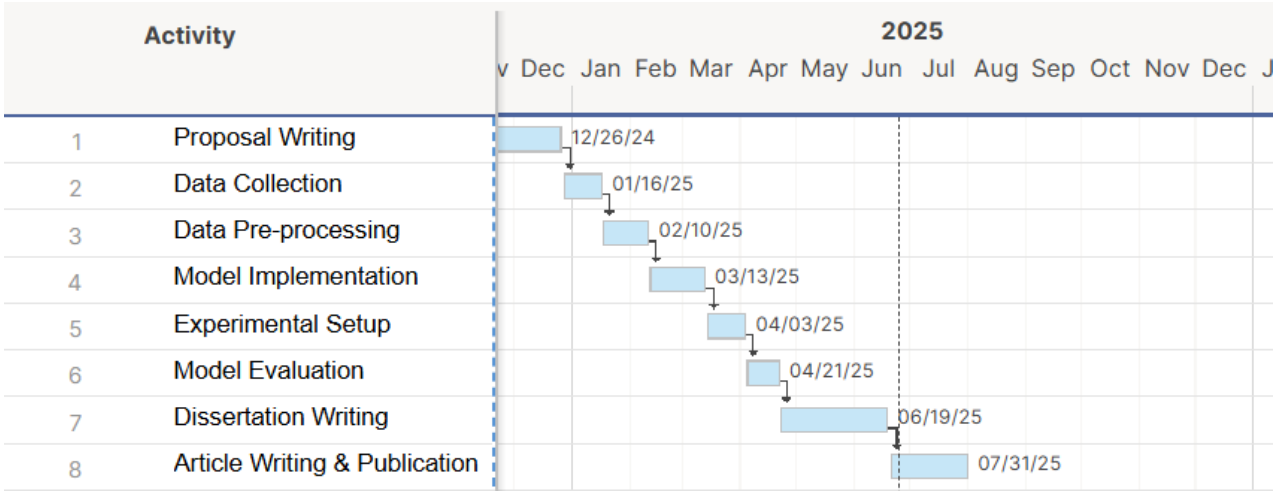


FIGURE 28

Gantt chart illustrating pictorial representation of project schedule

Created using: https://app.smartsheet.com/

Appendix C: Resources & Budget

Category	Details	Estimated Cost (KES)	Justification
Laptop/PC	High-performance computer for data analysis	80,000.00	Required to process datasets, implement machine learning model and to handle computational tasks
Software Costs	Python programming language and scikit-learn library download & setup costs	5,000.00	To facilitate installing and set up of programming tools for model development and data analysis
Data Access	Subscription to network traffic data sources	5,000.00	Required to ensure access to real world datasets for training and testing isolation forest-based model
Journal Fees	publication fees to publish findings in academic journals	5,000.00	Necessary for disseminating research findings to the academic community.
Travel Expenses	Travel expenses for data collection and presentations	5,000.00	To support travel for presentation of research outcomes
Data collection Expenses	Researcher stipend for data collection and analysis	10,000.00	This is compensation for the time and effort in collecting, processing and analysing data.
Unforeseen Expenses	Reserve for unforeseen expenses or additional project needs	5,000.00	This is meant for the unexpected costs that may arise during the project.
Total Estimated Budget		**115,000.00**	

TABLE 14

Project Resources & Budget

Appendix D: Selected Network Traffic Features

Feature Name	Description
Flow Duration	Duration of the entire network flow (in milliseconds).
Flow Bytes/s	Average number of bytes sent per second during the flow.
Total Forward Packets	Total number of packets sent from source to destination.
Total Backward Packets	Total number of packets sent from destination to source.
Destination IP	IP address receiving the data.
Time of Activity	Timestamp when the traffic flow occurred.
Flow IAT Mean	Average inter-arrival time between packets in a flow.

TABLE 15

Description of Selected Network Traffic Features

Appendix E: Required Python Libraries and Sample Codes

Below are screenshots illustrating key model development stages

```
[1]  # Import libraries for data handling, computation and visualization
     import pandas as pd
     import numpy as np
     import matplotlib.pyplot as plt
     import seaborn as sns

     # Import tools for feature scaling, anomaly detection & model evaluation
     from sklearn.ensemble import IsolationForest
     from sklearn.preprocessing import MinMaxScaler
     from sklearn.metrics import classification_report
     from google.colab import files

     # Configuration/settin of seaborn style to ensure more readable plots
     sns.set(style="whitegrid")
```

FIGURE 29

Importation of necessary libraries for model development

```
uploaded = files.upload()
```

Choose Files Thursday-...ap_ISCX.csv

- **Thursday-WorkingHours-Afternoon-Infilteration.pcap_ISCX.csv**(text/csv) - 83102436 bytes, last modified: 6/7/2018 - 91% done

```
uploaded = files.upload()
```

Choose Files Thursday-...ap_ISCX.csv

- **Thursday-WorkingHours-Afternoon-Infilteration.pcap_ISCX.csv**(text/csv) - 83102436 bytes, last modified: 6/7/2018 - 100% done

```
Saving Thursday-WorkingHours-Afternoon-Infilteration.pcap_ISCX.csv to Thursday-WorkingHours-Afternoon-Infilteration.pcap_ISCX.csv
```

FIGURE 30

Acquisition & upload of data

```
[3]  file_name = next(iter(uploaded))
     df = pd.read_csv(file_name)

     # For preservation of labels and preview dataset (first rows, last rows, Number of Rows & Column)
     df_labels = df.copy()
     print("Data Preview")
     display(df.head())
     display(df.tail())
     display(df.shape)
```

FIGURE 31

Loading & initial preview of data

```
# For dropping irrelevant columns & replacing string values with actual NaN values
df.drop(columns=['Flow ID', 'Timestamp', 'Label'], inplace=True, errors='ignore')
df.replace(['Infinity', 'NaN'], np.nan, inplace=True)
df = df.apply(pd.to_numeric, errors='coerce')
df.fillna(df.median(), inplace=True)

# For display of descriptive statistics of dataftame after cleaning
print("Data shape after cleaning:", df.shape)
display(df.describe())
```

FIGURE 32

Data cleaning and descriptive statistics

```
# For Initializing the selected features
selected_features = [
    ' Flow Duration',
    ' Total Fwd Packets',
    'Flow Bytes/s',
    ' Flow IAT Mean',
    ' Destination Port'
                    ]
# For extraction of the selected features and for handling of infinities and missing values
X = df[selected_features]
X.replace([np.inf, -np.inf], np.nan, inplace=True)
X.fillna(X.median(), inplace=True)
scaler = MinMaxScaler()
X_scaled = scaler.fit_transform(X)
X_scaled_df = pd.DataFrame(X_scaled, columns=selected_features)
display(X_scaled_df.head())
```

FIGURE 33

Feature Selection and Normalization to be used to train the model

```
# Key parameters to initialise the model before training
model = IsolationForest(
    n_estimators=100,
    max_samples='auto',
    contamination=0.01,
    random_state=42,
    bootstrap=True
                        )
# For training of model, calculation anomaly score and predicting anomaly label
model.fit(X_scaled)
df['Anomaly Score'] = model.decision_function(X_scaled)
df['Anomaly Label'] = model.predict(X_scaled)
```

FIGURE 34

Training/Building the Isolation Forest anomaly detection model

```
# For creation of column based on the original labels
from sklearn.metrics import confusion_matrix
df['True Label'] = df_labels[' Label'].apply(lambda x: 1 if x == 'BENIGN' else -1)

# For generating classification report to show evaluation metrics
print("Classification Report")
print(classification_report(
    df['True Label'], df['Anomaly Label'],
    target_names=['BENIGN', 'ANOMALY']
))

# For generation of the confusion matrix to visualize the performance
cm = confusion_matrix(df['True Label'], df['Anomaly Label'], labels=[1, -1])

# For creation of a heatmap visualization of the confusion matrix.
sns.heatmap(cm, annot=True, fmt='d', cmap='Blues',
            xticklabels=['BENIGN', 'ANOMALY'],
            yticklabels=['BENIGN', 'ANOMALY'])

 # Chart Labels & plot display
plt.title('Confusion Matrix')
plt.xlabel('Predicted')
plt.ylabel('Actual')
plt.show()
```

FIGURE 35

Performance Evaluation of Isolation Forest anomaly detection model

```
# For creatiion and display of a count plot to show the distribution of anomaly labels
sns.countplot(x='Anomaly Label', data=df)
plt.title("Anomaly Detection Results (-1 = Anomaly, 1 = Normal)")
plt.show()

# For creation and display of a histogram to visualize the distribution of anomaly scores
sns.histplot(df['Anomaly Score'], bins=50, kde=True)
plt.title("Anomaly Score Distribution")
plt.xlabel("Anomaly Score")
plt.ylabel("Frequency")
plt.show()

# For creation and diplay of correlation heatmap to relationships between the features used
plt.figure(figsize=(8, 5))
sns.heatmap(df[selected_features].corr(), annot=True, cmap='coolwarm')
plt.title("Correlation Heatmap of Selected Features")
plt.show()
```

FIGURE 36

Isolation Forest model results visualization

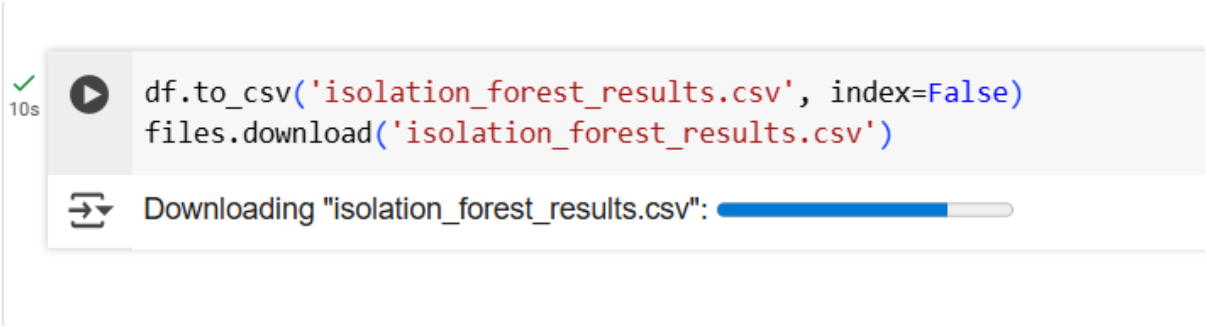


FIGURE 37

Exporting model results and saving in CSV format for further analysis