sciendo INTERNATIONAL JOURNAL ON SMART SENSING AND INTELLIGENT SYSTEMS

Research Article | DOI: 10.2478/ijssis-2025-0043 Issue 1 | Vol. 18 (2025)

Establishing software-defined network for fraud detection on energy fraud and traffic classification test case(s)

Elisha Indarjit, Vipin Balyan and Marco Adonis

Department of Electrical, Electronics, and Computer Engineering, Cape Peninsula University of Technology, Bellville, Cape Town 7535, South Africa

*E-mail: elishaindarjit@gmail.com, vipin.balyan@rediffmail.com, adonism@cput.ac.za

Received for publication January 22, 2025.

Abstract

The increase of fraudsters and fraud attacks on the communication network plays a major role in the loss of revenue, network abuse, and degradation of services. Communication and cloud networks belong to separate companies and consist of handover points from one network to another. The scope is to bring the communication and cloud networks closer and understand network traffic profiles using Software-Defined Networking (SDN) concept. The SDN controller will serve to route the extracted tapped data to a central server instead of making use of the live network to route traffic. To apply policies in an automated method and identify which user's traffic can be a fraud case, and further to send a block signal to the network charging element on the communication network. The traffic flow process will be redesigned using the SDN controller within the architecture for a seamless approach and maintaining the performance of the network. This article explores energy fraud identification and traffic classification based on four test cases to define a new framework for communication and cloud providers to enable the detection and blocking of fraud. The aim of this study is to perform real-time data classification; reporting of fraudulent subscribers, applications, and protocols; real-time analytics; and real-time blocking of data is classified as a fraudulent activity within the mobile data network. The main purpose is to achieve a higher energy/service fraud detection rate than similar work in the field, to provide a solution that integrates within the traditional network, and further blocks the fraud within the network.

Keywords

Fraud, Detection, Communication network, Cloud network, Software-Defined Network

I. Introduction

Fraud detection plays a major role in networks, subscribers, and revenue, and the digital era is exploding with new technologies, such as 5G/6G networks, and requires automated solutions to suppress harmful elements that can arise. The fraud attacks across the network will negatively impact companies' resources and downtime on the communication network. To avoid them, careful steps need to be taken to reengineer the network's framework from Manjula Devi et al. (2024, p. 3) and to seek a solution for the mitigation of fraud attacks.

Currently, there are outdated fraud detection systems but no automated blocking, and previous research work shows high expense and uses hierarchical layers of the infrastructure from Raman et al. (2024, p. 2).

Open Access. Published by Sciendo. CC BY-NC-ND
© 2025 Indarjit et al., This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License https://creativecommons.org/licenses/by-nc-nd/4.0/

Fraud detection systems encounter the significant challenge of rapidly adapting to constantly changing fraud patterns and the evolving strategies of fraudsters, all while quickly identifying new and increasingly complex schemes by Thar and Wai (2024). Also, Software-Defined Networking (SDN) is a concept that is designed on a vertical layer, which oversees the environment, and the study intends to use SDN on an in-line setup for fraud detection.

The main purpose is to achieve a higher energy/service fraud detection rate than similar work in the field, to provide a solution that integrates within the traditional network, and further blocks the fraud within the network. This study will define a new framework for communication and cloud providers to enable the detection and blocking of fraud.

The process of detection is performed by many techniques and each with its manual intervention to initiate the detection, and part of the gap area identified is an old scheme to address developing networks. The requirement for a solution is automated. The area of blocking is mentioned in previous studies/journal papers but not covering the scope for a network design in a communication networks, the requirement is for a solution to perform automated blocking of a fraud user. Another area is the source of data; for an efficient system, it needed to interface with the user data, which brings up the challenge of processing, the requirement for a solution that does not degrade the performance of network nodes.

Another area is the time frame of detection and blocking, and the purpose of a solution is to block the fraudster. The longer the fraudster utilizes the communication network or communication service, the greater the network is occupied with indirect losses. Also, the requirement on correlation of the virtual and physical interfaces to ensure network packet arrive in sequence.

This study examines fraud detection and its schemes provided in Table 1; to gain a better perspective into the problem, this study builds on its critical characteristics:

- SDN location and specification
- Energy fraud identification
- Test case(s) proposals

The new integrated platform will serve the cloud and traditional networks to build fraud detection and blocking solutions and to provide an improved fraud detection scheme than the current methods in Table 1.

This study takes into consideration the impact of fraud on the Service Provider and end user; while digital platforms are scaling and becoming more popular, the fraudster is also adapting to changes by finding loopholes in design principles. The aim is to solve or overcome the technology issues through automation, service evolution, aggregation schemes, and traffic filtering/intelligence.

II. Fraud Elements

a. SDN location and specification

This study shows the SDN controller used for centralized control consisting of software to enforce policies in Figure 1; the SDN controller architecture will handle fraud detection and blocking signals; and it becomes the proposed architecture to perform the study. The SDN layer will encompass two parts of fraud detection in terms of traffic classification and fraud blocking by the integration with the policy charging element. The physical connectivity remains the same; however, the logical connectivity portraits change and overlook three domains in Figure 2:

- Traffic is routed from the probe platform at different sites to the SDN controller, to note the SDN controller host routing tables of the nodes and links connected.
- The SDN controller will perform the traffic classification and fraud detection.

The SDN controller can manage the Metro Network domain, Core Network domain, and Data Centre domain, for each a Virtual Private Network (VPN) and Segment Routing are created and applied, both protocol is of importance since during the process of traffic classification, traffic will be encrypted for the purpose of a safe/protected network, and also changes in routing play a role either via packet loss or packet duplication.

The next layer involved is the protocols used to establish the network, the protocols fall either under service or transport protocols. At which a brief summary is provided on each, the research needs to ensure that traffic is readable and within the correct format for traffic classification. The 4G network is designed with the control and the data plane together, while 4G CUPS designs both planes separately, and further 5G network encrypts the service-based interfaces. Therefore, the research becomes complex, and each sector needs to be addressed or catered for.

Service protocols are formulated with the following:

Table 1: Fraud detection methods

Classification method	Description	Limitation	Reference
XGBoost	The method uses a gradient-boosting framework to build full-scale decision trees and implement parallel decision trees.	Trends to overfit the data.	Sheng and Yu (2022), Bao, (2020)
ADABoost	The technique handles binary classification problems and improves predictability by the conversion of a larger number of weak learners into strong learners.	The requirement of a dataset devoid of most of the noise.	Yulita et al. (2021), Chang and Fan (2019)
Naive Bayes	The method is based on Bayes' theorem to predict the outcome by the probability of occurrence.	The downfall is faced by the zero-frequency issue, which is the missing variable as zero.	Vijay and Verma (2023), Hairani et al. (2021)
Decision Tree Classifier	The algorithm uses classification and regression problems, and it works on a tree-based structure, which serves as the classifier of the dataset.	High in computation of resources and changes in data affect the outcome.	Zulfikar et al. (2018), Indumathi et al. (2021)
Random Forest classifier	The technique allows aggregation of several Decision Tree classifiers to improve the predictive capability of the algorithms.	The issue on the massive amount of computational resources and the time required between periods.	Mishra et al. (2020), Lu et al. (2019)
KNN	The method is used for classification on a distance-based approach to locate all unknown data points.	The downfall does not work on high dimensionality or large records.	Lu et al. (2015), Altay (2022)
Logistic Regression	The techniques are used for dichotomous and dependent variables.	The issue is on overfitting the count of features and recorded observations.	Doss and Gunasekaran (2023), Bheemesh and Deepa (2023)

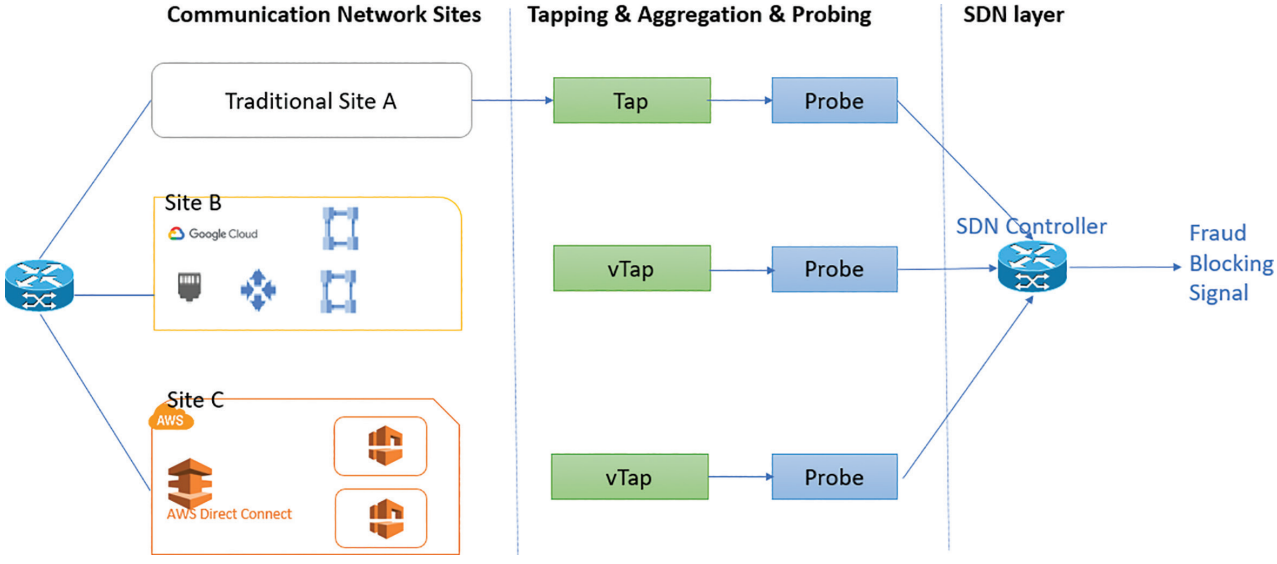


Figure 1: Software-defined network controller architecture. SDN, Software-Defined Networking.

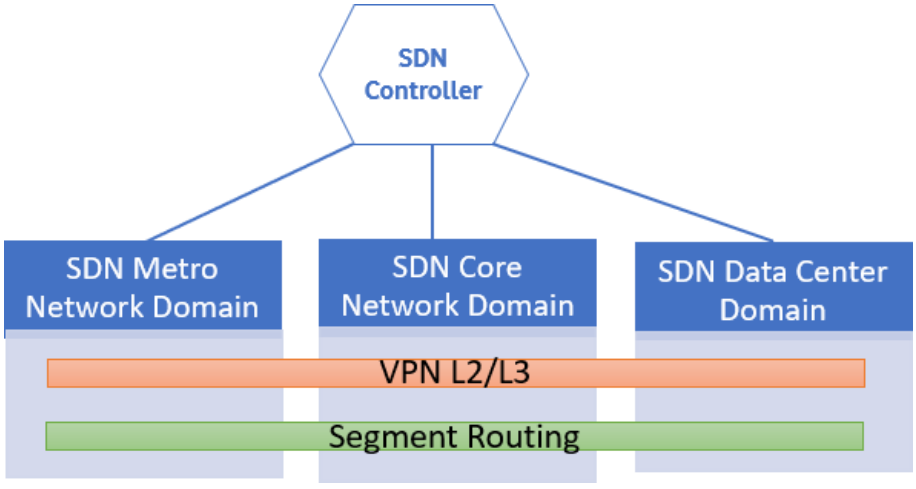


Figure 2: SDN controller overlooking Metro domain, Core domain, and Center domain. SDN, Software-Defined Networking; VPN, Virtual Private Network.

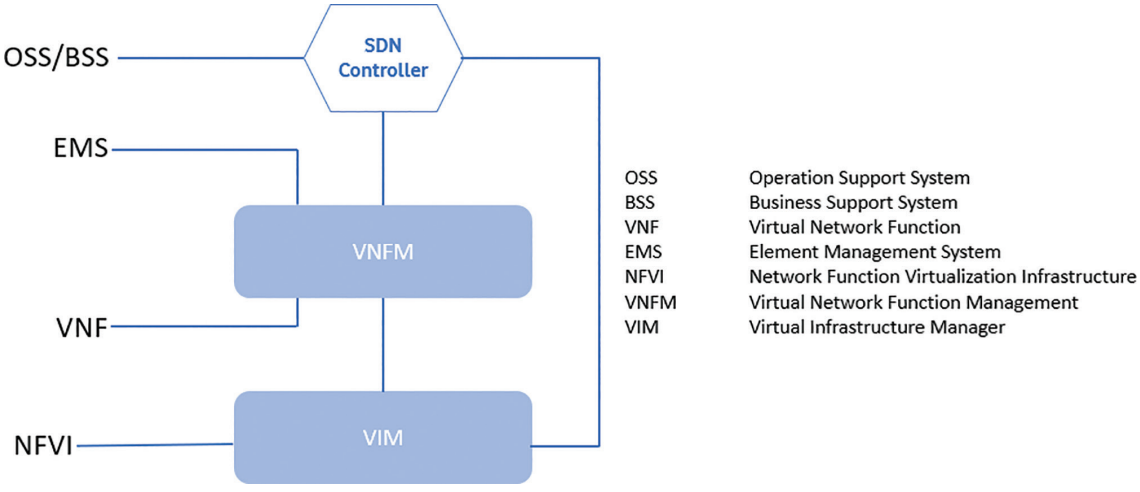


Figure 3: Layered Software-Defined Network on a virtual platform for service provisioning. VNF, Virtual Network Function.

- Layer 2 VPN builds a topology of point-to-point connections that connect end users.
- Layer 3 VPN, the Customer's Edge (CE) switch, or router must be configured to exchange traffic with the Provider's Edge (PE) switch/router.

Transport protocols are formulated with the following:

- Inter-Domain Traffic Protocols

These protocols manage the flow of data between different administrative domains or networks. Examples include:

- Border Gateway Protocol (BGP): The primary inter-domain routing protocol used to exchange routing information between different autonomous systems on the internet.

The SDN controller is built by a vendor or organization to provision services from the network. Since the design is based on virtual infrastructure, the layered architecture in Figure 3 provides end-to-end service provision, implements automatic network deployment and unified O&M, and meets other requirements.

The SDN layered virtual network adopts an open architecture to help simplify and efficiently operate and maintain the NFV networks.

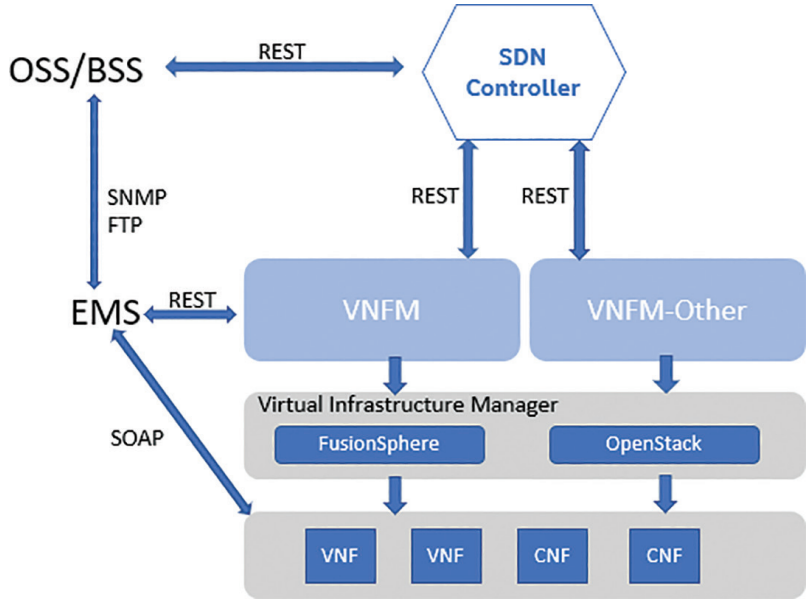


Figure 4: Detailed layered Software-Defined Network setup. CNF, Container Network Function; SDN, Software-Defined Networking; VNF, Virtual Network Function.

- For service provisioning,
- To provide network service life cycle management, and,
- Flexible applicability

Figure 4 shows a detailed layered Software-Defined Network architecture, showing the protocols used to communicate and its components:

SDN—To manage the network service life cycle, including service deployment, scale in/out, and termination, and provide cross-DC resource scheduling capability.

OSS/BSS—Serves as the support system of the live network and manages the Virtual Network Functions (VNFs) and Container Network Functions (CNFs).

EMS—Serve as the node management system, to monitor network outages at the Software as a Service and Infrastructure as a Service layer, and to maintain the capabilities of the VNFs and CNFs.

VIM—Serves as the infrastructure layer management system of FusionSphere and FusionManager.

VNFM—To manage the VNF life cycle for deployment, scale in/out, and termination.

The research explores the 4G and 5G networks; the 5G network is built using VNF and CNF; and on CNF deployment, the following metrics are taken:

To prepare the deployment environment, the following steps occurred:

- Configuring the switch underlay network and installing and pre-configuring.
- Deploy and install SDN software, PaaS, and Life Cycle Management.
- Manually create virtual private clouds (VPCs), host groups, and disk types on FusionSphere.
- Create tenants and projects and modify credentials on PaaS.

Integration challenges: Implementing advanced algorithms often requires significant investment in technology and training, which can be a barrier for some organizations. From the test case(s) in Section "Test Case(s) proposal", it was identified that the cost to expand the solution due to growing capacity plays a negative role in meeting the initial investment, though different techniques can be used to mitigate the cost factor. Such as:

Internet protocol pooling: The process is to configure on the Tapping platform a selection of Internet Pools at each site within a communication network and to circular the Internet pools by save on capacity expansions, let's take an example, site a had 40 Internet pools, to ensure the capacity processing

limits are met, 20 Internet pools are configured within a map rule for month 1 and the next 20 Internet pools are configure for month 2. Each Internet pool has an equal capacity. The process of alternating the pools allows a huge saving on capacity expansions.

b. Energy fraud identification

A fraudster seeks to bring harm to the smart grid and dismantle the operational aspects for their personal agenda, and their objective is to utilize the smart grid without payment or cause harm to the infrastructure. For the purpose of the paper, two main aspects are studied: energy abuse and application abuse. Both environments are of large scale and with many access points consisting of multiple systems, and this study focuses on the communication network and identifies the aggregation point.

From the aggregation point, the traffic can be sampled to further identify traffic profiles for each environment. On energy abuse and service abuse, based on a sample of users, the average trend on vlan() is shown in Figure 5, and the peak utilization is important and needs to be understood.

a and b indicate normal utilization peaks occurring twice in a 7-day count,
c indicates failover of traffic, a network is built on redundancy, failover occurs for a certain duration of time, and the capacity integrated serves to accommodate user traffic.

Figure 6 shows the footprint of abuse on the network over a 7-day average per period, d shows fraud on the network, and the characteristics of fraud are:

The utilization exceeded the highest peaks over the 7-day period,
The high utilization over a longer period compared to other active sessions,
The changes within data usage considering hard stops,
The policy of random peak usage.

It was found that when fraud is detected, it can be two categories:

Ghost traffic, The Tx is indicated as 0. The sample is shown from the testing,

[local]GGPS03# show port util table
------ Average Port Utilization (in mbps)
------ Port Type Current 5min 15min
Rx Tx Rx Tx Rx Tx ----- ----------
---- ------- ------- ------- ------- ------- ------- 5/13 10G
Ethernet 3002 0 2855 0 2817 0

On the actual fraud traffic, the TX and RX both have a value. Fraud requires to be looked at within a month, and the pattern of a fraudster can be categorized by random sampling.

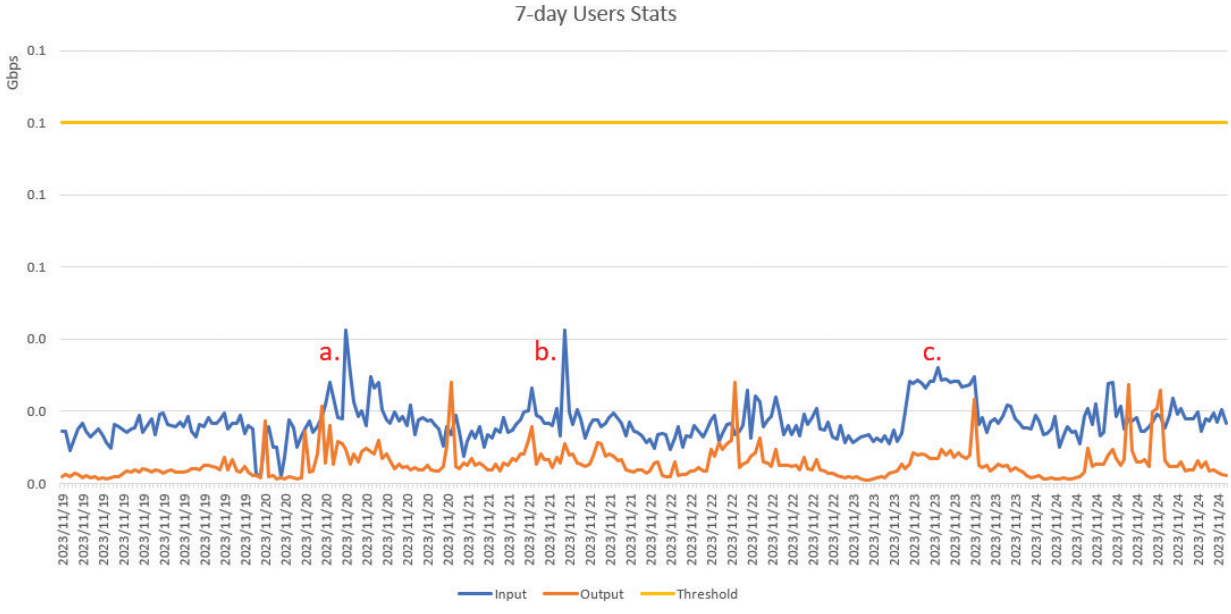


Figure 5: 7-day user stats on average of a sample of ().

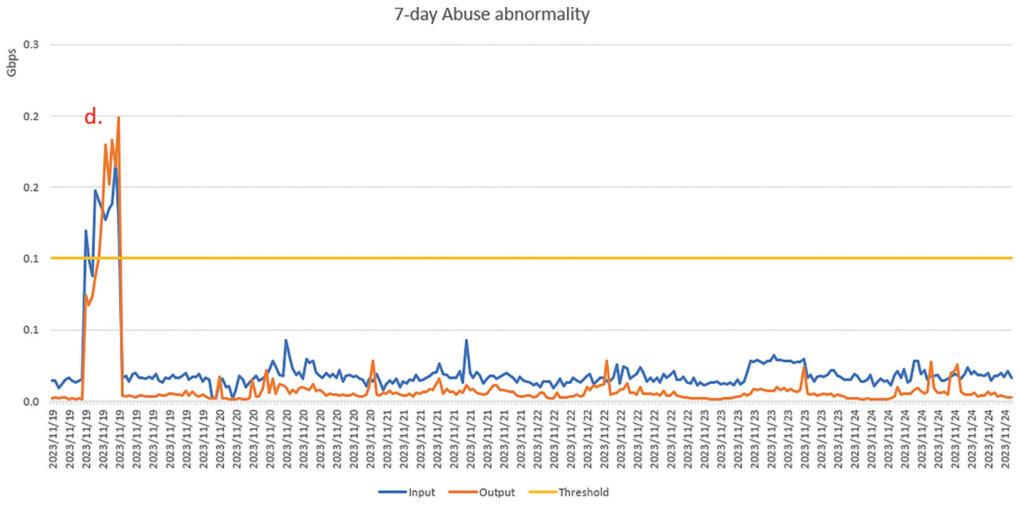


Figure 6: 7-day user stats showing network abuse.

c. Test case(s) proposal

c.i. Test case 1: Fraud detection with an unpaid token on the communication network

Energy sector, communication sector, and service application sector.

To use traffic classification and identify the patterns of a fraudster for each sector. To create an algorithm by machine learning for the detection.

Fraud detection with an unpaid token on the network requires bringing the network's subscriber platform and the billing platform together. The mapping of the subscriber to their billing becomes a fundamental point for the research and its implementation. The first step is to explore what components are active on the billing platform in Figure 7.

To showcase Test Case 1 and replicate the concept, many tools were explored to bring the billing platform and subscriber usage together, and the software tool used to demonstrate the testing by the name of Alteryx Designer was incorporated, making use of its components. The data used are collected from Call Detail Records (CDR), which holds the subscriber ID, data usage, and details of the session of the Home Location Register (HLR). The results mapped record of Subscriber_Rated_Billed and identified incorrectly billed subscribers due to platform abuse.

Example for Test Case 1:

- Subscriber ID: 1162565599
- Rated: 20.72355
- Billed: 1
- Fraud: 19.72355

c.ii. Test case 2: Fraud detection utilizing the network

To evaluate the effects of fraud users for traffic classification and build an algorithm to automate the detection.

To build an accurate traffic classification model, the following considerations need to be taken:

- Inaccurate traffic classification, from the complexity of different devices and firmware version upgrades, Service Providers make use of encryption to preserve customer traffic,
- Misclassification, considering traffic classification rules, which are static and do not appear to be automated, the flexibility to classify on port number, source-destination address, and domain names.
- Complex encryption algorithm, when traffic is encrypted, it becomes complex to perform classification, considering encryption by random numbers, hash functions, and other techniques.
- Traffic processing problem, the process to de-encrypt traffic can take much more processing power.
- Traffic congestion, subscribers that use multiple applications at the same time makes it much more challenging to classify and record the traffic for each application.

To address the encryption challenge, the data require to be formatted and then de-encrypted on the

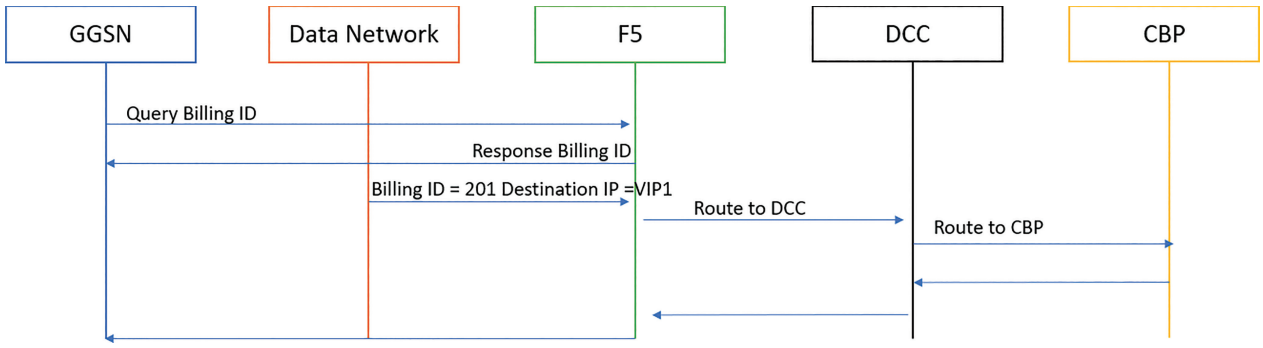


Figure 7: Billing flow—High level.

The components are shown below

CBP	Convergent Billing Point implements rating, charging, and accounting functions and supports both online charging and offline charging, also providing real-time QoS control, and this is triggered based on the threshold configured in the tariff.
DCC Proxy	The function supports the specific demands of Diameter Charging—a dedicated online mediation instance, to take control of any internal routing of Diameter traffic based on subscriber.
GGSN	Manages data sessions and integrates with Online Charging System for real-time data rating and charging.
F5	F5 Distributed Cloud Services Billing service enables a subscriber to understand usage reports, quotas, and pricing, obtain usage reports, and switch between subscription plans.

The test case needs to extract if the subscriber is billing correctly by performing checks from the Convergent Billing Point, where the Convergent Charging System (CCS) resides.

communication network, the process to extract the Key and initialization vector by obtaining the encryption key and IV from the context (usually established during the authentication process). To apply AES Decryption by using the AES algorithm to decrypt the encrypted traffic. The same key and IV used for encryption are required for decryption. In addition, removal of Padding, it should be removed after decryption.

The test case requires the evaluation of traffic or utilization in a network, the process to analyze the traffic, cater on the requirements for positive fraud case, using previous work to formulate and compare a higher detection rate, and apply the algorithm on a software tool. The test case required to create a new framework to solve the deployment.

The framework in Figure 8 for traffic classification involves categorizing the traffic to manage Billing, Quality of Service, Network Management, and Security. The following pillars are used to approach traffic classification:

The common grounds to host and manage the traffic are:

- Billing: Implement fair usage policies or tiered service plans.
- Quality of Service (QoS): allows traffic to be prioritized traffic for different user experiences (e.g., emergency service or government service)
- Network Management: To manage the network nodes and systems.
- Security: Caters on the detection and to mitigate malicious/unauthorized traffic.

To perform the test case on traffic classification to identify a fraud case, the three areas required to be implemented are Policies, Conditions, and Actions. The use of assessing traffic patterns and intelligence metrics is required to be applied to subscriber traffic. The aim is to filter and mark out fraud subscribers using automation and ensure that the session is blocked. Test case 2 allows the creation of detecting fraudulent traffic from the network, using the tapped user plane traffic, to place Policies, Conditions, and Actions. A policy named Fraud_Detection_High_Utilization permits the identification of traffic above its mean threshold and high utilization over a period of

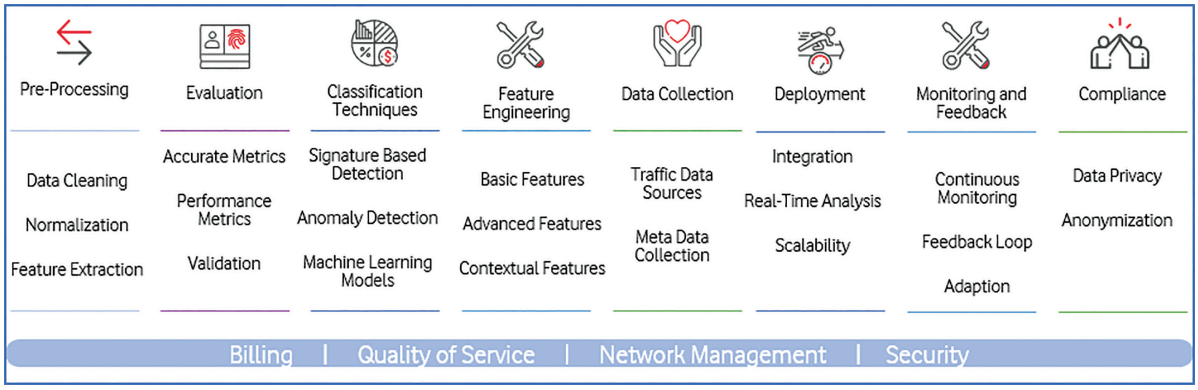


Figure 8: Traffic Classification Framework.

time, and once both Conditions meet the criteria, the configured Action will be applied.

- Policy: Fraud_Detection_High_Utilization
- Condition: Monitor traffic above mean threshold.
- Action: Drop specific traffic classified as fraudulent.

Results:

- Utilization of Traffic: Monitored over a 30-day period.
- Condition: Established to identify high utilization.
- Action: Successfully dropped fraudulent traffic.

c.iii. Test case 3: Fraudulent access on cloud network

The means of using other forms of access is to make use of the cloud environment and its key functions on a virtual machine. The test case will allow the identification of a fraudster on the cloud network access.

A summary of the services of the three main cloud providers is stated in Table 2.

To define the difference between the following clouds and why the importance of a Hybrid cloud, in summary,

The Public Cloud requires no capital expenditure to scale up or expand, the applications running can be easily scaled out, and the cost is dependent on what resources are used: the Private Cloud, the hardware required to be purchased from the start of deployment, the owner has full control of the platform and it easily secures, and also, the owner is responsible for maintenance and upgrades. Additionally, Hybrid Cloud possesses the highest flexibility; on applications, the owner can deploy any applications without being bound by cloud provider limits; and the owner controls the security, legal aspects, and compliance.

VPC is a private network inside a public network (the Internet), and the VPC provides a high level of security and control over its network infrastructure, allowing to define customized network configurations and control access resources and more. The VPC also provides a high level of isolation, allowing to create secure and desolated environments for their applications and services. This makes the VPC a great solution for organizations that run sensitive workloads or comply with security and privacy regulations.

The detection process combines multiple data sources, applied machine learning for real-time anomaly detection and used Hadoop for real-time graph analysis and visualization. Test Case 3 demonstrates the use of Big Data and machine learning for real-time fraud detection in cloud networks. By leveraging Hadoop, distributed SQL query engines, and advanced ML algorithms, the solution provides robust and scalable fraud detection capabilities, ensuring enhanced security and operational efficiency. In terms of Google Cloud, their products are serverless and managed by Google, allowing more time to prepare the data, build the fraud detection model, host online predictions on streaming data, set up the fraud notification, and create operational dashboards. The first stage is to gather the historical data on credit card transactions as training data, containing credit card numbers, transaction amount, merchant information, category, and subscriber demographics; for Test case 3, the following Fraud Model in Figure 9, XGBoost is trained to predict 1 if fraud else 0 based on various features, such as transaction category, amount, and demographics.

Table 2: Overview of cloud key services

Service	Description	AWS	Azure	GCP
Computing	Virtual machines and scalable computing resources	EC2	Virtual machines	Compute engine
Object storage	Storage for unstructured data in objects	S3	Blob storage	Cloud storage
Block storage	Storage for data in blocks, similar to traditional hard drives	EBS	Disk storage	Persistent disk
Database (relational)	Managed relational database services	RDS	SQL database	Cloud SQL
Database (NoSQL)	Managed NoSQL database services	DynamoDB	Cosmos DB	Firestore/ datastore
CDN	Global distribution of content to reduce latency and improve performance	CloudFront	Azure CDN	Cloud CDN
Serverless computing	Running code without managing server infrastructure	Lambda	Functions	Cloud functions
Big data processing	Processing and analyzing large datasets	EMR	HDInsight	Dataproc
Machine learning	Provision of services and tools for machine learning	SageMaker	Machine Learning Studio	AI platform
Identity and access management	Management of users and permissions	IAM	Azure AD	Cloud IAM
Monitoring and logging	Monitoring and logging of applications and infrastructure	CloudWatch	Azure monitor	Stackdriver (operations)
Networking	Management of networks and their security	VPC	Virtual network	VPC
Container orchestration	Management and orchestration of containers	EKS	AKS	GKE
Data warehousing	Storage and analysis of large amounts of structured data	Redshift	Synapse analytics	BigQuery
Backup and disaster recovery	Backup and recovery of data	Backup	Azure backup	Backup

CDN, content delivery network; VPC, virtual private cloud.

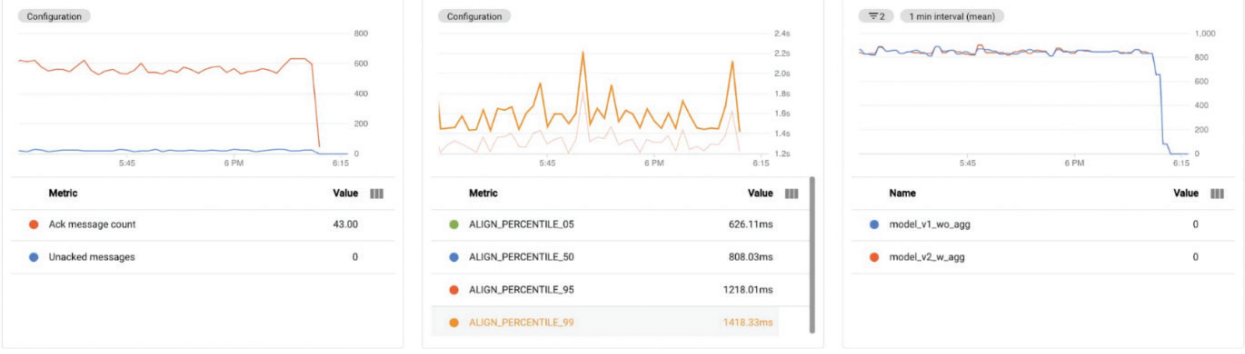


Figure 9: Online Cloud Fraud prediction.

c.iv. Test case 4: DDoS attack

To identify when fraudsters attempt to limit the network's availability by identifying their pattern and scoping a proposed deep-learning technique for detection.

An attack by a fraudster or group of fraudsters in attempt to make the network's resource temporarily or indefinitely unavailable to its intended users. This is achieved by several methods, but the basic principle is to overload the target server or something in the network path to the server so that it becomes unavailable to legitimate users.

The target is on:

- Network Link Capacity (1–100 Gbps link—easily filled these days),
- Firewall session table or connection per second capacity,
- Load Balancer or Web Server session table or connection per second capacity,
- DDoS, the DNS server that serves as the authoritative answer for the domain in question.

A typical attack is performed using botnets, which are either under the control of the attacker or hired for the duration of the attack. Botnets consist of 100s, 1,000s, or 10,000s of zombie computers. These are Internet-connected devices that have been compromised and can be remotely issued with commands by their command-and-control servers.

The ultimate goal of these attacks is varied and could be any of the following:

- Blackmail—A demonstration of the ability to take down a site, followed by a ransom email. (Gaming, banking, and e-trading sites all employ robust DDoS mitigation services.)
- State-sponsored/terrorist—During conflicts, it is advantageous to take down opponents' websites.
- Personal—online computer gamers use DDoS to win against their opponents.
- "Just because I can"—DDoS attacks can be purchased easily from darknet sites.

It can be identified by the types of attack:

- Bandwidth versus packets per second to occupy the specified capacity.
- Reflection/Amplification attacks.
- Slow attacks—Slow loris: some attacks appear to be real connections, but they gradually fill the server's connection table.

Figure 10 is an example of the ATLAS summary report of all the attack activities as seen by users of Arbor Peakflow equipment over a period of 24 hr. The largest attack was 168.74 Gbps in bandwidth, and the max attack rate was 25.88 Mpps/25.88 million packets per second.

The largest attack seen so far is about 500 Gbps (this was in December 2014 and took out several ISPs around Hong Kong).

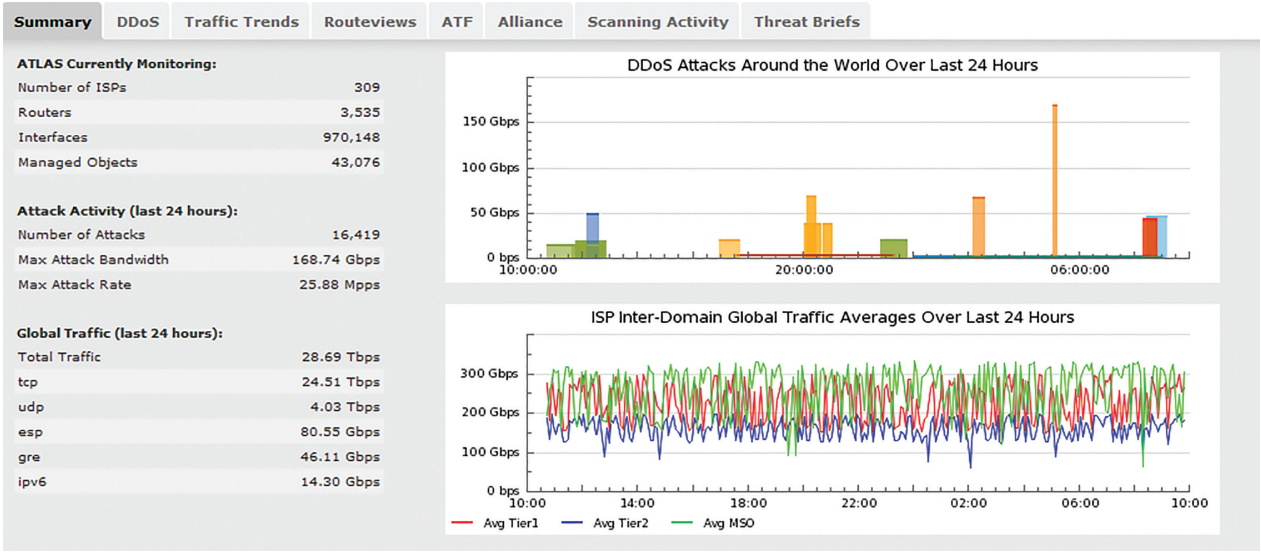
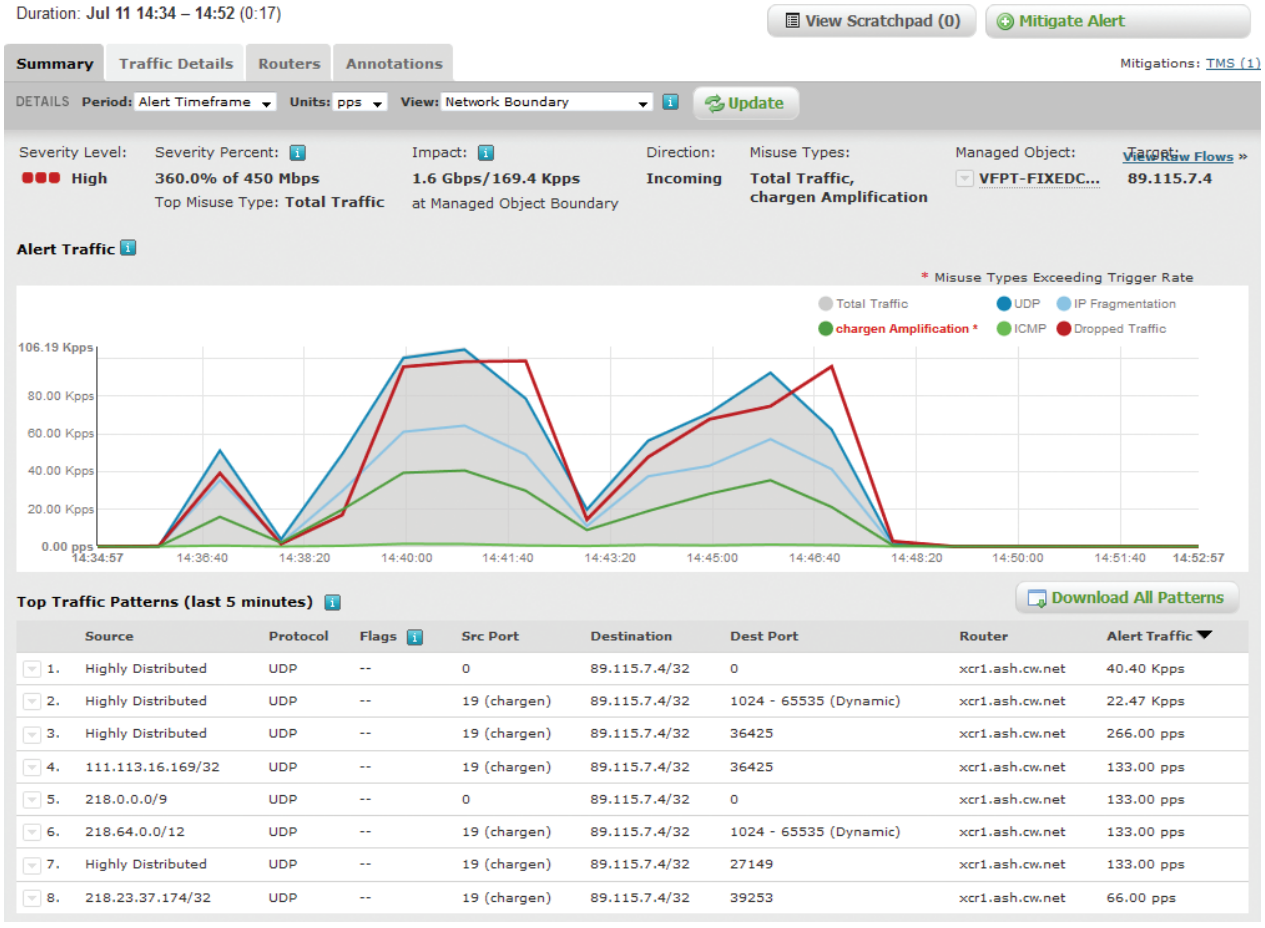


Figure 10: ATLAS showing activities within 24 hr.



	Source	Protocol	Flags	Src Port	Destination	Dest Port	Router	Alert Traffic
1.	Highly Distributed	UDP	--	0	89.115.7.4/32	0	xcr1.ash.cw.net	40.40 Kpps
2.	Highly Distributed	UDP	--	19 (chargen)	89.115.7.4/32	1024 - 65535 (Dynamic)	xcr1.ash.cw.net	22.47 Kpps
3.	Highly Distributed	UDP	--	19 (chargen)	89.115.7.4/32	36425	xcr1.ash.cw.net	266.00 pps
4.	111.113.16.169/32	UDP	--	19 (chargen)	89.115.7.4/32	36425	xcr1.ash.cw.net	133.00 pps
5.	218.0.0.0/9	UDP	--	0	89.115.7.4/32	0	xcr1.ash.cw.net	133.00 pps
6.	218.64.0.0/12	UDP	--	19 (chargen)	89.115.7.4/32	1024 - 65535 (Dynamic)	xcr1.ash.cw.net	133.00 pps
7.	Highly Distributed	UDP	--	19 (chargen)	89.115.7.4/32	27149	xcr1.ash.cw.net	133.00 pps
8.	218.23.37.174/32	UDP	--	19 (chargen)	89.115.7.4/32	39253	xcr1.ash.cw.net	66.00 pps

Figure 11: Summary report of High Alert.

How do we know when there is an attack? How do we know when to mitigate an attack?

It all starts with Managed Objects (MOs)—(formerly called zones—the terms are interchangeable). A MO is essentially the definition of the network object(s) that is to be protected. The MO is a primary building block of the DDoS mitigation service.

When a High Severity alert is generated in Figure 11, the system performs a number of configuration notification actions.

Typically, three actions are automatically performed when a High Severity alert is raised.

- An email is sent to the relevant support teams and customer,
- Logs the alert to Syslog,
- Send an SNMP-TRAP to the VCHS Netcool Servers, and,
- This is used to generate a Remedy ticket that is auto-routed to the relevant support team.

Conclusion

This study proposes research from various studies that showcase their scope and results achieved in aiming to build and improve a new solution that performs network fraud detection and blocking. The new integrated platform will serve the cloud and traditional networks to build fraud detection and blocking solutions and to provide an improved fraud detection scheme. Using a hybrid build of sites for connectivity consists of virtualization for network flexibility and to map SDN technology to a re-designed fraud detection framework.

This study looks at hybrid cloud deployment to show the services offered by well-known cloud providers. The aim is to solve or overcome the technology issues through automation, service evolution, aggregation schemes, and application filtering/intelligence.

The use case of Smart Energy abuse and Service application abuse has been researched and tested,

proposing a new mitigation scheme for protecting the user. The focus is to reduce the rate of fraud in a network, and the solution builds a new platform by using the existing systems.

The below are summary points applied to the Test case(s) on methods:

Rule-Based Systems: These systems use predefined rules to identify fraudulent activities. For example, a rule might flag transactions over a certain amount or from unusual locations. Despite being straightforward, these systems can be easily circumvented by sophisticated fraudsters. On Test Case 1: Fraud detection with unpaid token on the Communication Network uses Alteryx Designer software tool and makes use of its components. The data used are collected from CDRs, which holds the subscriber ID, data usage, and details of the session. The process flow uses input data from the CDR file, which is converted into excel format, and the two inputs used are the subscriber "Rated" information and subscriber "Billed" information. The method is able to compare and filter out the fraud identified.

Machine Learning Algorithms: A more advanced methods, including supervised learning (such as decision trees and support vector machines) and unsupervised learning (such as k-means clustering), can identify complex patterns in large datasets. These algorithms learn from historical data and can improve over time but require substantial data and computational power. Test Case 2: Fraud detection utilizing the network applies an algorithm based on Policy, Condition, and Action, and the algorithm identifies traffic patterns on thresholds configured over a monitoring time. The Condition values can be per-defined to enable accurate automation of fraud identification.

Anomaly Detection Algorithms: These algorithms identify outliers in data that deviate from established patterns, often using methods, such as isolation forests or autoencoders. They are effective in identifying new or evolving fraud patterns but may produce false positives. Test Case 3: Fraudulent Access on Cloud Network presents its method to detect fraud on a cloud network using Ingest, Detect, Investigate, Protect, and Impact.

Deep Learning: Neural networks, particularly those designed for sequential data (such as LSTMs), can detect intricate patterns in transaction sequences. They are powerful but often seen as "black boxes," making interpretation and transparency a challenge. Test Case 4: DDoS attack, it uses both flow and deep packet inspection (DPI) technologies and to provide macrolevel and microlevel visibility, allowing to identify threats and to improve the performance of the network.

In this review, a comprehensive study was conducted to evaluate the methods and impact for fraud detection and blocking.

References

Altay, O. (2022). Performance of different KNN models in prediction English language readability. *2022 2nd International Conference on Computing and Machine Intelligence (ICMI)*, Istanbul, Turkey, pp. 1-5.

Anupriya, E., Kumaresan, N., Suresh, S., Dhanasekaran, S., Ramprathap, K, and Chinnasamy, P. (2022). Fraud Account Detection on Social Network Using Machine Learning Techniques. *2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC),* Bhubaneswar, India, pp. 1-4.

Bao, J. (2020). Multi-features Based Arrhythmia Diagnosis Algorithm Using Xgboost. *2020 International Conference on Computing and Data Science (CDS)*, Stanford, CA, USA, pp. 454-457.

Bheemesh. R.K, and Deepa. N. (2023). Accurate SMS Spam Detection Using Support Vector Machine in Comparison with Logistic Regression. *2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*, Chennai, India, pp. 1-5.

Bindu, P., Mishra, R., and Thilagam, P.S. (2018). Discovering spammer communities in twitter. *Journal of Intelligent Information Systems*, vol. 51, no. 3, pp. 503–527.

Cao, N., Shi, C., Lin, S., Lu, J., Lin, Y.,-R., and Lin C,-Y. (2015). Targetvue: Visual analysis of anomalous user behaviors in online communication systems. *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 280–289.

Chang, K.-C, and Fan, C.-P. (2019). Cost-Efficient Adaboost-based Face Detection with FPGA Hardware Accelerator. *2019 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW)*, Yilan, Taiwan, pp. 1-2.

Doss, L.M, and Gunasekaran, M. (2023). Evasion and Poison attacks on Logistic Regression-based Machine Learning Classification Model. *2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*, Chennai, India, pp. 1-8.

Gonzalez, P.A. (2005). Fraud detection method for mobile telecommunication networks, EP.

Gori, M., G. Monfardini, and F. Scarselli. (2005). A new model for learning in graph domains. in

Proceedings. *2005 IEEE International Joint Conference on Neural Networks*. IEEE.

Gurajala, S., White, J., Hudson, H, and Matthews, J. (2016). Profile characteristics of fake Twitter accounts. *Big Data Society*, pp. 1–13.

Hairani, H., Anggrawan, A., Wathan, A.I., Latif, K.A., Marzuki, K, and Zulfikri, M. (2021). The Abstract of Thesis Classifier by Using Naive Bayes Method. *2021 International Conference on Software Engineering & Computer Systems and 4th International Conference on Computational Science and Information Management (ICSECS-ICOCSIM)*, Pekan, Malaysia, pp. 312-315.

Indumathi, N., Ramalakshmi, R, and Ajith, V. (2021). Analysis of risk factors in the Firework Industries: Using Decision Tree Classifier. *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Greater Noida, India, pp. 811-814.

Juszczyszyn, K & Kolaczek, G. (2019). Complex Networks Monitoring and Security and Fraud Detection for Enterprises. *2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, pp. 124-125.

Manjula Devi, C., Gobinath, A., Padma Priya, S., Adithiyaa, M., Chandru, M.K and Jothi, M. (2024). Next-Generation Anomaly Detection Framework Leveraging Artificial Intelligence for Proactive Credit Card Fraud Prevention and Risk Management. *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kamand, India, pp. 1-6.

Mishra, S., Mallick, R.K, and Gadanayak, D.A. (2020). Islanding Detection of Microgrid using EMD and Random Forest Classifier. *2020 International Conference on Computational Intelligence for Smart Power System and Sustainable Energy (CISPSSE)*, Keonjhar, India, pp. 1-5.

Molloy, I., Chari, S., Finkler, U., Wiggerman, M., Jonker, C., Habeck, T., Park, Y., Jordens, F and Schaik, R. (2016). Graph analytics for realtime scoring of cross-channel transactional fraud," in *Proceedings of International Conference on Financial Cryptography and Data Security*, pp. 22–40.

Niu, K., Jiao, H., Deng, N., and Gao, Z. (2016). A Real-Time Fraud Detection Algorithm Based on Intelligent Scoring for the Telecom Industry. *2016 International Conference on Networking and Network Application(NaNA),* Hakodate, pp. 303-306.

Park, N., et al. (2019). Estimating node importance in knowledge graphs using graph neural networks." in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*.

Peng, L, and Lin, R. (2018). Fraud Phone Calls Analysis Based on Label Propagation Community Detection Algorithm. *2018 IEEE World Congress on Services, San Francisco*, CA, 2018, pp. 23-24.

Raman, R., Kumar, V., Pillai, B.G., Rabadiya, D., Divekar, R and Vachharajani, H. (2024). Detecting Credit Card Fraud: A Comparative Analysis of KNN, Random Forest, and Logistic Regression Methods. *2024 Second International Conference on Data Science and Information System (ICDSIS)*, Hassan, India, pp. 1-5.

Sahoo, P. K., Mishra, S., Panigrahi, R., Bhoi, A. K., & Barsocchi, P. (2022). An Improvised Deep-Learning-Based Mask R-CNN Model for Laryngeal Cancer Detection Using CT Images. Sensors, 22(22), 8834.

Shearer, C. (2000). The CRISP-DM model: The new blueprint for data mining," *J. Data Warehousing*, vol. 5, no. 4, pp. 13–22.

Sheng, C and Yu, H. (2022). An optimized prediction algorithm based on XGBoost. *2022 International Conference on Networking and Network Applications (NaNA)*, Urumqi, China, pp. 1-6.

Smruthi, N. Harini. (2019). A Hybrid Scheme for Detecting Fake Accounts in Facebook. *International Journal of Recent Technology and Engineering (IJRTE)* ISSN: 2277-3878, 7:5S3.

Tarmazakov, E, I. and Silnov, D, S. (2018). Modern approaches to prevent fraud in mobile communications networks. *2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, Moscow and St. Petersburg, Russia. pp. 379-381.

Teng, H., Wang, C., Yang, Q., Chen X, and Li, R. (2023). Leveraging Adversarial Augmentation on Imbalance Data for Online Trading Fraud Detection. *in IEEE Transactions on Computational Social Systems*.

Thar, K.W and Wai, T.T. (2024). Machine Learning Based Predictive Modelling for Fraud Detection in Digital Banking. *2024 IEEE Conference on Computer Applications (ICCA)*, Yangon, Myanmar, pp. 1-5.

Vijay, V, and Verma, P. (2023). Variants of Naïve Bayes Algorithm for Hate Speech Detection in Text Documents. *2023 International Conference on Artificial Intelligence and Smart Communication (AISC)*, Greater Noida, India, pp. 18-21.

Xuan, S., et al. (2018). Random forest for credit card fraud detection. in *2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*. IEEE.

Yulita, I.N., Paulus, E., Sholahuddin, A, and Novita, D. (2021). AdaBoost Support Vector Machine Method for Human Activity Recognition. *2021 International Conference on Artificial Intelligence and Big Data Analytics*, Bandung, Indonesia, pp. 1-4.

Zhao, M, and Song, H. (2022). Semantic Analysis based on Artificial Intelligence—Prediction of Telecom Fraud. *2022 4th International Conference on Frontiers Technology of Information and Computer (ICFTIC)*, Qingdao, China, pp. 924-927.

Zhong, R., Dong, X., Lin, R. and Zou, H. (2019). An Incremental Identification Method for Fraud Phone Calls Based on Broad Learning System. *2019 IEEE 19th International Conference on Communication Technology (ICCT), Xi'an*, China, pp. 1306-1310.

Zhou, J et al. (2023). FraudAuditor: A Visual Analytics Approach for Collusive Fraud in Health Insurance. in *IEEE Transactions on Visualization and Computer Graphics*.

Zulfikar, W.B., Gerhana, Y.A, and Rahmania, A.F. (2018). An Approach to Classify Eligibility Blood Donors Using Decision Tree and Naive Bayes Classifier. *2018 6th International Conference on Cyber and IT Service Management (CITSM)*, Parapat, Indonesia, pp. 1-5.