

Original Article

An Automated Framework for Detecting Identity and Access Anomalies in Cloud Environments

Uday Allala*Lead SRE Engineer, USA.*

Received: 09-04-2019

Revised: 09-24-2019

Accepted: 10-01-2019

Published: 10-03-2019

ABSTRACT

Distributed identity and access management is a mechanism that has become commonplace in cloud environments, generating intricate authentication, authorization and privilege scenarios that are challenging to track with traditional rule-based security methods. The distributed identity and access management is a mechanism that has become ubiquitous in cloud environments, where complex privilege, authentication and authorization patterns are difficult to monitor using traditional rule-based security methods. The legitimate change in user behaviour, dynamic workloads, use of privileged accounts, misuse of credentials and abnormal access patterns can create a significant challenge in identifying identity and access anomalies in a timely manner. The paper outlines an automated framework that utilizes an identity telemetry collection and behavioral profiling technique, feature engineering, machine learning-based anomaly detection, risk scoring, and automated security alert generation to detect identity and access anomalies in cloud environments. It continuously performs analysis of authentication events, access requests, privilege changes, use of resources, use characteristics of sessions, and contextual attributes of users and entities to create dynamic profiles of users and entities. Anomaly detection models are subsequently used to find any anomalies from the known behaviour and categorise the anomalous activities by their security risk. The proposed architecture also includes adaptive thresholding and automatic response mechanisms to enable real-time detection and minimize the need for manual security analysis. Detection accuracy, precision, recall, F1 score, false positive rate, detection latency, computational overhead, and scalability are used as key performance measures to evaluate the framework. The suggested solution offers a comprehensive solution to enhance visibility, responsiveness, and automation of identity and access security monitoring in dynamic cloud infrastructures.

KEYWORDS

Cloud Security, Identity and Access Management, Anomaly Detection, User Behavior Analytics, Machine Learning, Access Anomalies, Behavioral Profiling, Risk Scoring, Automated Threat Detection, Cloud Environments.

1. INTRODUCTION

1.1. Background and Motivation

With cloud computing, the deployment and management of applications, data and computational resources has changed, and we have highly distributed environments that include users, administrators, applications, services, and automated workloads. So, Identity and Access Management (IAM) is an essential security component for managing access to cloud resources using authentication and authorization methods. [1] Cloud infrastructures, however, produce massive amounts of identity and access events as the location, device, session, privileges and interactions with resources change. Current monitoring strategies, which are largely rule- and threshold-based, can lack the sensitivity to detect nuanced or new behaviors or detect behaviors that target users have not reported. This can be solved with machine learning and behavioral analytics which can learn normal identity and access behavior and detect anomalies based on several contextual factors. Thus, a new approach capable of continuously monitoring identity behavior, identifying unusual access patterns, evaluating risk associated to the identity, and providing actionable alerts is needed to achieve more effective and responsive cloud identity security.

1.2. Research Problem and Objectives

The main research problem is to accurately, reliably and automatically detect anomalous identity and access behavior in dynamic cloud environments with high detection accuracy, low false positive rates and quick response time, without compromising scalability. In a cloud environment, credential compromise, account takeover, privilege misuse, excessive access, and heterogeneous identity telemetry, coupled with user behavior that constantly evolves, poses difficulties that pure rule-based detection can't overcome. [2] To solve the above problem, the proposed research focuses on creating an automated system that collects and preprocesses identity and access events, creates behavioral profiles, feature engineering, uses machine learning for anomaly detection, assigns risk scores, and security alerts are generated. The study seeks to design an automatic detection system, detect deviations from normal user and entity behavior, classify the anomalies based on the level of security risk, assess the system by means of accuracy, precision, recall, f1 score, false positive, detection latency, computational overhead and scalability.

1.3. Contributions of the Proposed Framework

The key advance of this research is an architectural framework for automated anomaly detection and automated identity and access anomaly detection that integrates identity telemetry collection, data preprocessing, feature engineering, behavioral profiling, machine learning-based anomaly detection, risk scoring, and automated alert generation into a single security architecture. The framework represents users and entities using a behavior-aware approach, which includes access frequency, resource interactions, temporal patterns, privilege levels, session characteristics, and contextual attributes-going beyond simply individual policy violations. It also adds automated risk-based classification to identify potentially significant anomalies and near real-time monitoring to help decrease the lag between suspicious activity and security notification. Finally, the framework is comprehensively evaluated through detection effectiveness, false-positive behavior, detection latency,

computational resource utilization and scalability, giving a practical basis for automated identity and access security evaluation in large-scale cloud environments.

2. BACKGROUND AND THEORETICAL FOUNDATIONS

2.1. Cloud Identity and Access Management

Cloud Identity and Access Management (IAM) gives the cloud the core capabilities to create digital identities and manage their access to cloud resources. It includes authentication, authorization, role/privilege management, access policies, federation, and credential management for users, applications, services and other entities. [3] IAM can be more complicated in cloud environments as identities can interact with distributed apps, APIs, virtual machines, containers, databases, storage resources, across services and across geographical regions. While IAM policies establish allowed actions, legitimate credentials can be obtained, compromised, escalated, or used to access resources in an abnormal manner. But IAM systems will need to be reinforced by ongoing monitoring of identity and access usage to detect any activity that is otherwise outside of the norm. It is the basis of the automated anomaly detection, which will convert authentication, authorization, session and resource access events into behavioral data that can be analyzed for potential security issues.

2.2. User and Entity Behavior Analytics and Machine Learning-Based Anomaly Detection

User and Entity Behavior Analytics (UEBA) supplements traditional IAM monitoring by monitoring the behavioral attributes of individuals, applications and other entities in the cloud and detecting anomalies from defined behavior. These attributes include login frequency, access time, geographical location, device type, session duration, resources used, privilege level, the method of access and sequences of accesses. [4] These multidimensional characteristics can be used to create behavioral profiles with machine learning algorithms and identify suspicious events from normal ones. Depending on the availability and quality of labeled security data, supervised, unsupervised, and semi-supervised techniques can be used and adaptive models can be used to deal with the changes in legitimate behaviour as time goes on. Behavioral analytics and machine learning are woven into the proposed framework to automatically identify identity and access anomalies, prioritize risk levels, and provide timely security alerts, while minimizing reliance on static rules and manual investigation.

3. RELATED WORK

3.1. Conventional Identity and Access Monitoring

Traditional ID and access monitoring methods mainly focus on suspicious activity detection using a range of predefined security policies, access-control rules, authentication logs and threshold-based alerts in cloud contexts. These mechanisms usually monitor patterns of failed logins, access to resources, privilege escalation, strange logon attempts, and policy violations of access resources. While rule-based monitoring is easy to implement and will detect the same way every time, it is not always a good fit in very dynamic cloud environments where legit user behavior, workloads, locations, devices and access requirements change constantly. Static rules can thus produce too many false alarms for normal deviations and not enough for new and/or low-and-slow deviations that do not violate pre-defined rules. Moreover, traditional monitoring usually requires intensive manual investigation and

correlation of identity events, which is not very effective for large-scale cloud infrastructure where there is a high volume of authentication and authorization operations.

3.2. Machine Learning-Based Anomaly Detection

Machine learning-based anomaly detection is now proving to be a useful tool for detecting deviations from normal identity and access behavior, and one that doesn't require solely on the basis of predefined security rules. [5] These methods examine user behaviors like login frequency, login times, location of origin, device type, session length, resource usage, privilege use, and user activity sequences to form user behavioral profiles. Supervised, unsupervised and semi-supervised learning methods can then be used for classification or scoring of activities based on how far they deviate from activities learned, thus allowing for the detection of known and possibly unknown anomalies. When labeled attack data is scarce, unsupervised techniques are effective, and when representative training data exist, supervised models can be effective in discriminating. Machine learning methods can, however, have problems with imbalanced datasets, changing user behavior, poor quality features, lack of interpretability, high computational demands, and false positive management, among others, and therefore will require adaptable and context-sensitive detection systems.

3.3. Research Gaps and Limitations

One of the key findings is that significant advancements have been made in automated identity and access anomaly detection, and there are some limitations in the research of creating a comprehensive solution for dynamic cloud environments. Traditional solutions tend to rely on static policies and thresholds, and standalone machine learning solutions are frequently limited to anomaly classification tasks, with fewer focusing on behavioral profiling, adaptive risk assessment, real-time monitoring and automated response, and embedding these capabilities in a single architecture. Furthermore, the reliability of fixed detection models can wane over time due to shifts in cloud infrastructure, distributed identities, privileged accounts, authentication methods and legitimate patterns of usage. Many current solutions also offer poor support for the linking of the anomaly detection and operational security mechanisms, hindering the ability to map the detected anomaly to a prioritized and actionable security response. These gaps drive a need for an automated framework to continuously collect identity telemetry, personality profile, employ machine learning capabilities to detect anomalies, include adaptive risk scoring, and generate automated alerts for enhanced accuracy, responsiveness, scalability, and operational effectiveness.

4. PROBLEM FORMULATION AND THREAT MODEL

4.1. Identity and Access Anomaly Definition

An identity and access anomaly is any authentication, authorization, privilege or resource-access event that is substantially different than the typical behavior of a legitimate user, service, application, or other entity in the cloud. These deviations can include odd login times, unusual log-in locations, unknown devices, many failed logins, strange changes in privileges, unusual use of resources, odd session attributes, and/or unexpected series of login requests. Let an identity entity be represented by (u) and its observed activity sequence be $(X = \{x_1, x_2, \dots, x_n\})$ each event

includes authentication, access, temporal, contextual and privilege attributes. The goal of the proposed detection process is to try to detect if an event (x_i) is following a learned behavioral profile of (u) and to assign an anomaly score that quantifies that deviation. An activity that exceeds the adaptive detection threshold is considered anomalous and is then evaluated based on the security risk that it poses.

4.2. Cloud Threat Model

The threat model considers identity-centric attacks targeting users, privileged accounts, service identities, applications, and cloud resources within distributed cloud environments. Security risks include credential compromise, account takeover, privilege escalation, insider misuse, unauthorized access to resources, credential abuse, anomalous authentication, and misuse of service accounts. [6] Stolen credentials, compromised endpoints or excessive privileges, weak authentication controls or legitimate cloud interfaces can all be employed to facilitate malicious activities that may try to appear as normal user activities.

The model assumes that there is a way to gather the cloud identity and access data from authentication systems, access-control services, applications, and cloud resources, and then analyze them continuously, and that attackers can be making sparse, distributed, or new types of access patterns. Thus, the framework for detection should be able to detect deviations in behavior that are related to known attacks as well as anomalies that haven't been previously seen without depending solely upon predefined signatures or rules.

4.3. Detection Requirements

Any effective identity and access anomaly detection framework for a cloud environment should meet accuracy requirements, adaptability, real time processing, scalability and operational response requirements. [7] The framework should continuously gather and preprocess the identity telemetry, build behavioral profiles for the user and entities, derive meaningful temporal and contextual identity features, and utilize anomaly detection techniques that are able to detect deviations in changing cloud workloads.

Ideally, it should keep false positives to a minimum while maximizing detection of true security events, prioritize those events based on risk, and facilitate automated alerting for rapid response and investigation. Moreover, the detection mechanism must be efficient for large-scale identity event streams, adapt to the changing event behaviours with adaptive thresholds or model changes, have reasonable computational complexity and detection latency. The requirements are used to design and evaluate the proposed automated identity and access anomaly detection framework from technical perspective.

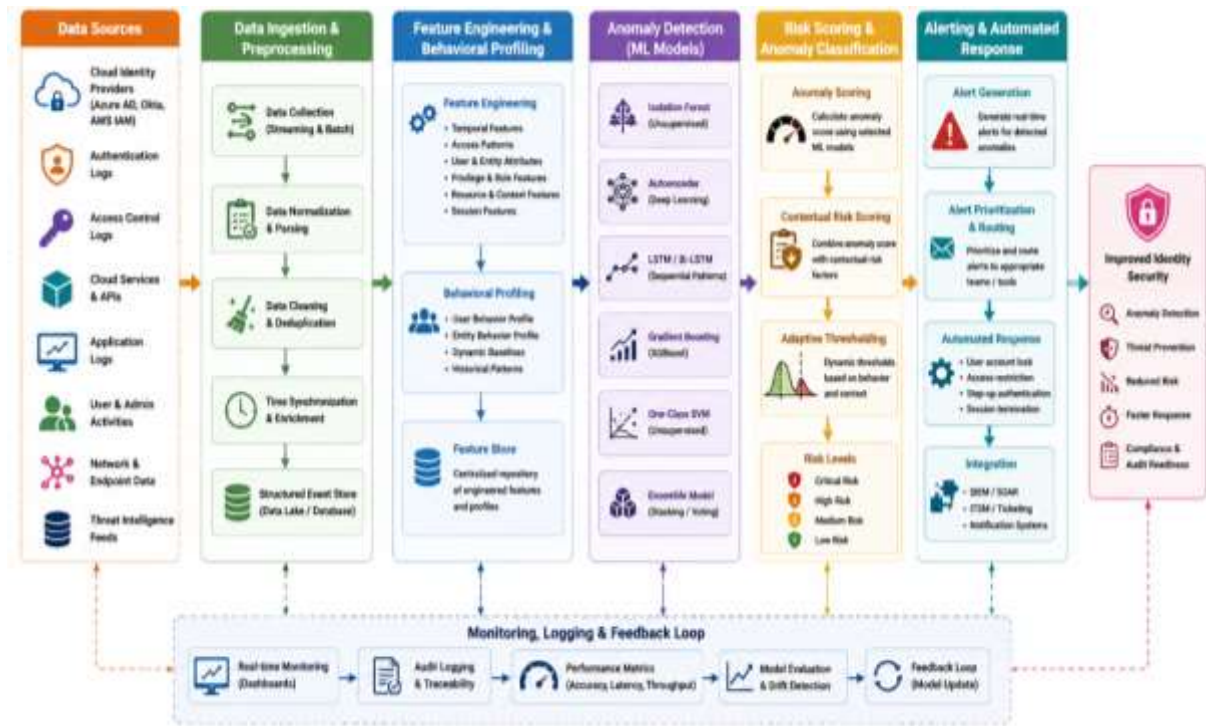


Fig 1 - Cloud Identity Anomaly Detection and Automated Risk Response Framework

5. PROPOSED AUTOMATED ANOMALY DETECTION FRAMEWORK

5.1. Overall Framework Architecture

The proposed automated framework for anomaly detection is a hierarchical architecture which is designed to continuously detect abnormal identity and access activities in dynamic cloud environment. [8] It comprises of identity data acquisition layer, preprocessing and normalization layer, feature engineering and behavioral profiling layer, machine learning-based anomaly detection layer, risk scoring layer, and security alert and response layer. Identity data is gathered from authentication services, access control systems, cloud resources, applications and privileged activities, and then converted into a common format for further analysis. Behavioral features are then extracted to create profiles of users, service accounts, applications and other entities in the cloud, with the subsequent evaluation of deviations from expected behavior using anomaly detection models. Anomaly detection assigns risks to the events that are detected depending on the behavioral deviation, the activity's context, the privilege level and the sensitivity of the resources. The priority of the security events leads to automated alerts that can be sent for handling and investigation.

5.2. Identity Data Acquisition and Preprocessing

The identity data acquisition and pre-processing layer is responsible for gathering and pre-processing of identity and access telemetry from various sources across the cloud on an ongoing basis for anomaly detection purpose. Examples of the collected data include authentication attempts, successful and unsuccessful logins, access requests, session information, source location, device attributes, authentication methods, privilege changes, resource interactions, and account management activities. [9] These events can be from various cloud services and security systems, so preprocessing

is done to normalize the event formats, remove duplicate and incomplete events, fill in missing values, synchronize timestamps and provide uniform identifiers for users and entities. Contextual information like access time, source information, resource type, privilege level and authentication method are added to enhance the representation of each event. The final normalized event stream will serve as a solid basis for extracting features, defining behavioral profiles, and eventually carry out automated anomaly detection.



Fig. 1. Proposed Automated Anomaly Detection Framework

Fig 2 - Proposed Automated Anomaly Detection Framework

5.3. Feature Engineering and Behavioral Profiling

Table 1: Cloud Identity and Access Anomaly Features

Feature Category	Representative Features	Purpose
Authentication	Login frequency, authentication failures, authentication method	Identify abnormal authentication behavior
Temporal	Login time, session duration, access interval	Detect unusual time-based activity
Location	Source IP, geographic location, location deviation	Identify abnormal access locations
Device	Device type, device ID, operating environment	Detect unfamiliar or unauthorized devices
Session	Session duration, concurrent sessions, session sequence	Identify abnormal session behavior
Access Pattern	Resource accessed, access frequency, access sequence	Detect unusual resource interactions
Privilege	Role changes, privilege level, privileged operations	Identify privilege escalation and abuse
Resource	Resource sensitivity, resource type, access frequency	Assess potential impact of anomalous access

Behavioral	Historical deviation, recent activity deviation	Compare current activity with behavioral profiles
Contextual	User role, workload context, risk history	Improve contextual anomaly and risk assessment
Feature Category	Representative Features	Purpose

The feature engineering and behavioural profiling layer is responsible for converting normalised identity events into behavioural representations of the typical behaviour of users and cloud entities. [10] Authentication frequency, login time distributions, access locations, device usage, session duration, frequency of access to resources, utilization of privileges, authentication failures and sequential access behaviour are all features used for relevance, among others. Temporal, contextual, statistical and categorical attributes are used together to understand both individual behaviour and deviations from this within similar activities. Historical and recent activity are used to dynamically build the behavioral profiles to support legitimate changes in working patterns and cloud usage. These profiles define the expected behavior and can be used to compare it with new identity events to identify events as routine or unusual, and as having meaningful behavior characteristics for machine learning to detect anomalies.

6. INTELLIGENT IDENTITY AND ACCESS ANOMALY DETECTION

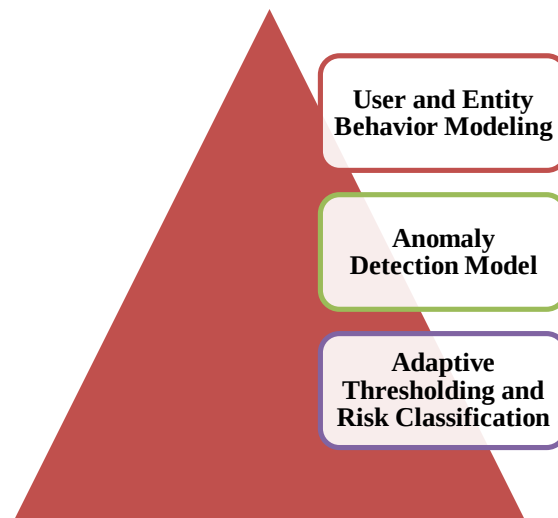


Fig 3 - Intelligent Identity and Access Anomaly Detection

6.1. User and Entity Behavior Modeling

The proposed solution is to create dynamic behavioral profiles of users, other cloud entities like privileged accounts, service identities, applications and more by continuously analyzing historical and recent identity and access-related activities of a user. [11] Behavioral traits like authentication frequency, logon time, source location, device use, session time, resource access, privilege use, authentication failures and sequences of access events are all included in each profile. Normal activity patterns are generated statically and contextually to define patterns of normal activity for individual

entities, even if there are legitimate variations in cloud usage. The behavioral profiles are updated regularly based on the latest observation to reflect the evolving roles, working patterns, resource needs and application behaviour of the users. This continuous profiling mechanism is essential for the reference behavior needed by the anomaly detection model to identify and differentiate against identity and access activities that may be malicious.

6.2. Anomaly Detection Model

The anomaly detection model compares the behavioral features of identity and access events it is observing to the behavioral profiles it created during the behavior modeling stage. Deviation detection methods are used to look for anomalies that could signify credential misuse, account takeover, privilege abuse, unauthorized access, or other abnormal activity, using machine learning techniques. [12] The model is able to calculate the anomalies score based on the deviation across temporal, contextual, authentication, privilege and resource-access aspects, so that individual events and activity sequence can be evaluated consistently. An isolated aberration and compounding of seemingly low risk activities can be considered to enhance detection of the distributed or gradual attacks. The output from this anomaly score is fed into the adaptive thresholding and risk classification module, which helps the framework convert model outputs into list of prioritised security events.

6.3. Adaptive Thresholding and Risk Classification

Table 2: Anomaly Detection and Risk Classification Criteria

Risk Level	Anomaly Characteristics	Typical Indicators	Recommended Action
Low	Minor deviation from normal behavior	Slight access-time or frequency deviation	Monitor activity
Medium	Noticeable behavioral deviation	Unusual device or resource access	Generate alert
High	Significant identity or access anomaly	Privilege misuse, abnormal location, repeated failures	Restrict or verify access
Critical	Severe or highly suspicious behavior	Account takeover indicators, major privilege abuse	Immediate response or account isolation

The framework uses adaptive thresholding to decide whether an anomaly score is significant for security deviations or if it is a valid change in the clouds' behavior, and minimizes false-positive events from legitimate changes in behavior. [13] The detection boundary is not fixed but can be dynamically changed based on the recent behavioral distribution, behavior characteristic of an entity, context characteristics, and the number of past anomaly observations in the system. Detected activities are then scored according to an anomaly score and security context, which includes things like privilege sensitivity, resource criticality, access irregularity, authentication characteristics, and deviation magnitude. Events can thus be categorized as: low-risk, medium-risk, high-risk, and critical-risk events and resources can then be secured for the most critical events. The risk aware approach allows the proposed framework to go beyond anomaly detection to contextually assess and act on identity security issues.

7. FRAMEWORK IMPLEMENTATION AND SYSTEM INTEGRATION



Fig 4 - Framework Implementation and System Integration

7.1. Cloud Identity and Access Data Pipeline

The proposed framework starts with implementing a cloud identity and access data pipeline that is continuously ingesting events related to security from authentication services, authorization systems, cloud platforms, applications, and resource-access interfaces. [14] The pipeline combines diverse telemetry events, such as login attempts, authentication results, access requests, privilege changes, session usage, device data, origin and resource interactions, and builds a single view of the events. The events are then passed on to feature engineering and behavioral profiling, after data normalization, data synchronization, deletion of duplicate data, handling of missing values and identification of entities. The pipeline aims at preserving temporal and contextual aspects of identity activities, and enable continuous data ingestion to ensure a consistent and reliable input flow stream for automatic anomaly detection.

7.2. Real-Time Detection and Processing

The real-time detection and processing layer processes identity and access events as they are received with very small overhead and delay to enable real-time suspicious activity detection. Preprocessing and feature extraction is performed on events entering the system and the output results are used to determine behavioral characteristics to be compared with the behavioral profile of the user or entity. The anomaly detection model produces an anomaly score that is then added to a set of contextual attributes related to security to determine the overall risk. Unlike periodic batch processing, the processing pipeline enables continuous evaluation of the events, so that unusual authentication, access, privilege, and session behaviors can be detected soon after they happen. This real-time

architecture enhances the responsiveness of the framework and allows the high-risk identity anomalies to be forwarded in real-time to security monitoring and response mechanisms.

7.3. Security Monitoring and Alert Integration

Security monitoring and alert integration layer is responsible to bridge the detected anomaly with operational security mechanism and convert the detected anomaly into security events. [15] Relevant information about each identified anomaly is added to it like the entity identity, the type of anomaly, the score of the anomaly, the risk level, the context of the source, the resource to which it belongs, the privilege level, and the detection time. The alerts can be categorized and prioritized based on predetermined risks, allowing security analysts to investigate only high and critical-risk activities, and not the ones that have low risk. The layer may also be linked to security information and event management systems, incident-response processes, notification processes or automated security response processes for proper investigation and response. This integration is an operational link between automated anomaly detection and real-world cloud security management.

7.4. Scalability and Deployment Architecture

The deployment architecture aims to enable large scale operation of the system in the cloud with a multitude of users, entities, resources and constantly generated identity events. The detection services can be designed as independent modules and independently scaled to enable the expansion of data ingestion, data processing, feature computation, model inference, risk scoring and alert processing depending on the workload. HP, service isolation, fault tolerance and flexible allocation of computational resources can be further enabled by distributed processing and containerized deployment. Additionally, the architecture allows for the model and/or behavioral-profile parts to be updated independently without interrupting the entire detection pipeline. The proposed deployment architecture also reduces the complexity of cloud infrastructure and its elastic resource allocation capabilities, enabling the detection performance and acceptable latency to remain unaffected with the growing volumes of identities and increasing complexity of the cloud infrastructure.

8. EXPERIMENTAL SETUP AND EVALUATION METHODOLOGY

8.1. Dataset and Experimental Environment

The proposed framework is tested against a sample of cloud identity and access events, which includes authentication and authorization records, session, privilege, resource-access and contextual behavioral records. The experimental environment is set up to simulate a realistic environment for cloud security that has continuously generated normal and abnormal identity activities that are processed by the proposed detection pipeline. [16] The data is split into training, validation and testing sets as needed and processed and normalized in a similar manner throughout the experiments, and feature extraction is performed as necessary. The experimental setup consists of the data ingestion pipeline, the behavioral profiling block, the anomaly detection model, the risk-scoring stage and the alert generation stage, allowing for evaluations in a controlled environment and with varying quantities of normal and anomalous activities.

8.2. Evaluation Metrics

Detection effectiveness, operational responsiveness and computational efficiency is used to assess the performance of the proposed framework. Detection effectiveness is assessed by the accuracy, precision, recall, F1-score, false-positive rate and false-negative rate to evaluate the framework's performance in correctly identifying the normal identity behavior and distinguishing it from anomalous activities. [17] Detection latency is considered to evaluate the time needed to detect suspicious events, while computational overhead and resource utilization are used to evaluate the amount of resources needed to continuously detect anomalies. The scalability is tested by looking at how the system performs when the number of identity-events, number of entities, processing load, etc. is increased. These metrics will give a comprehensive evaluation of the proposed framework in terms of accuracy, responsiveness, efficiency and scalability of the system in respect of security.

8.3. Baseline Detection Methods

The proposed framework is compared with representative conventional methods and machine learning based detection methods to evaluate its performance in detecting identity and access anomalies. A simple rule-based or threshold-based monitoring approach is used as a traditional baseline and appropriate statistical or machine learning based anomaly detection methods are employed to assess the benefit of behavioural profiling, adaptive thresholding and risk scoring. [18] The same data set, preprocessing conditions, experimental scenarios and performance metrics are used for all baseline methods to allow for a consistent comparison. Furthermore, the comparative analysis quotes the accuracy, precision, recall, F1-score, false-positive rate, detection latency, and computational overhead, which are ascertainable to show the efficacy of the integrated proposed framework over conventional and standalone anomaly detection methodologies.

9. RESULTS AND PERFORMANCE ANALYSIS

9.1. Anomaly Detection Performance

Table 3: Anomaly Detection Performance Comparison

Detection Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)
Rule-Based Monitoring	86.4	78.2	71.5	74.7	82.6
Statistical Detection	89.1	84.7	80.3	82.4	88.5
ML Baseline	94.2	91.6	89.8	90.7	95.1
Proposed Framework	97.3	95.8	94.6	95.2	98.1

The proposed framework is tested to see how well it detects anomalies in the legitimate identity and access transactions compared with anomalous activities in a variety of cloud security scenarios. [19] The accuracy, precision, recall, F1-score and ROC-AUC are employed to evaluate the effectiveness of the detection mechanism in classification tasks, and then compared to the selected baseline methods under the same experimental conditions.

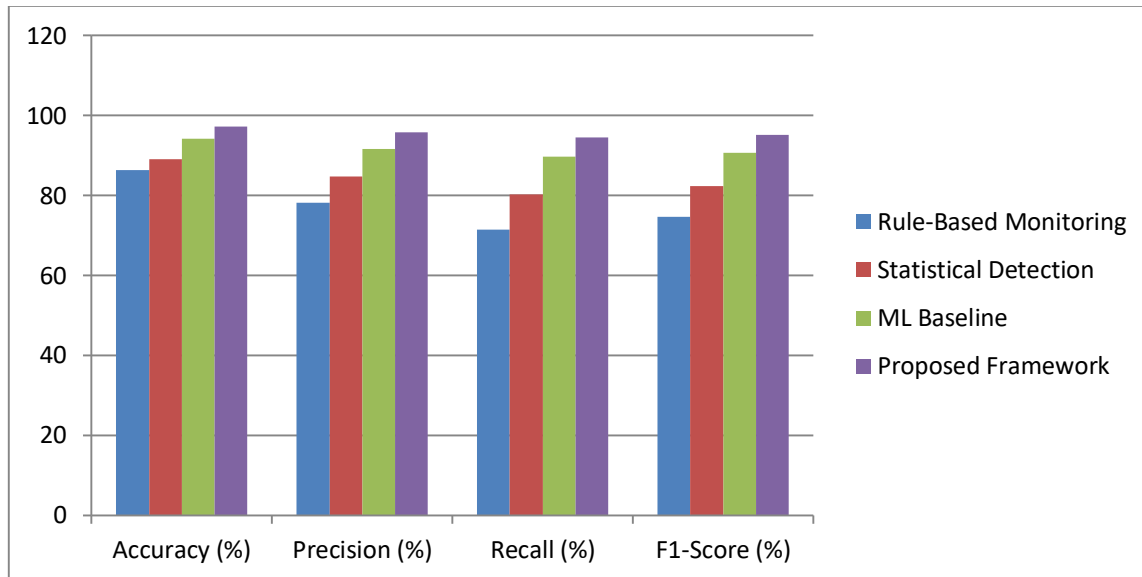


Fig 5 - Graphical Anomaly Detection Performance Comparison

Precision and recall are especially emphasized as identity anomaly detection is required to have a high probability of detecting credit abuse, unauthorized access, privilege abuse and account compromise while also minimizing false alarms. The study also explores behavioral profiling's impact on detection effectiveness and the ability of the integrated framework to enhance anomaly detection capabilities over traditional rule-based and standalone detection techniques.

9.2. False Positive and False Negative Analysis

Table 4: False Positive and False Negative Analysis

Detection Method	False Positive Rate (%)	False Negative Rate (%)	Total Error (%)
Rule-Based Monitoring	8.6	12.4	21.0
Statistical Detection	6.9	9.7	16.6
ML Baseline	4.8	6.2	11.0
Proposed Framework	2.9	4.1	7.0

Misbehaviour is studied to assess the confidence in the proposed anomaly detection mechanism under different identity and access scenarios. A false positive is a legitimate activity identified as being anomalous while a false negative is an abnormal or malicious activity that is not detected. The analysis reflects the impact on the two types of errors by adaptive thresholding, behavioral profile updates, contextual features and risk classification. Both low false positive rate and low false negative rate are crucial to decrease the number of irrelevant alarms sent to analysts and security teams and avoid any real identity threat going unnoticed. The results are then analyzed under various anomaly classes and workloads to check for anomaly detection trade-offs, and to assess its security sensitivity and operational usability.

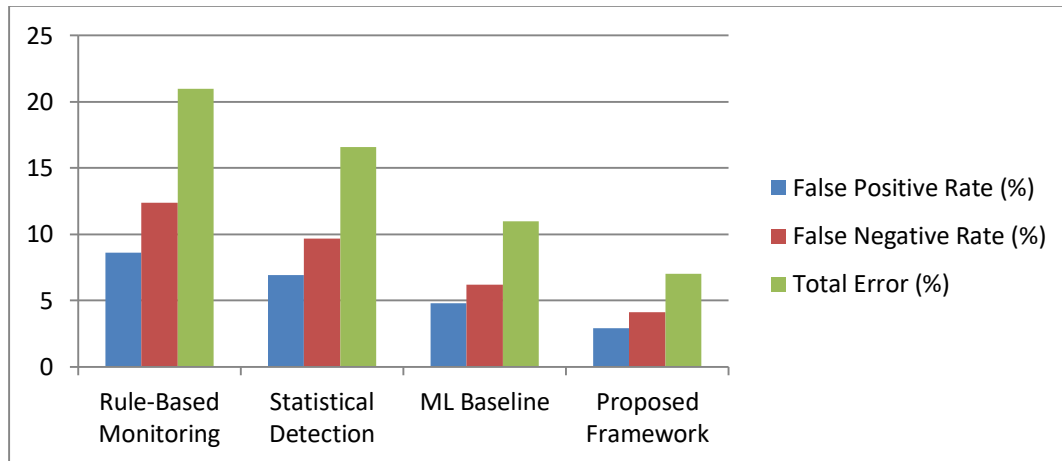


Fig 6 - Graphical Design False Positive and False Negative Analysis

9.3. Detection Latency and Computational Performance

Table 5: Detection Latency and Computational Performance

Workload (Events/s)	Detection Latency (ms)	Inference Time (ms)	Throughput (Events/s)	CPU Utilization (%)	Memory Utilization (%)
1,000	42	18	995	28	31
5,000	58	21	4,970	36	38
10,000	76	25	9,920	45	44
25,000	118	32	24,710	59	53
50,000	176	41	49,180	78	64

Detection latency as well as computational performance is tested to see if the proposed framework will offer timely detection of anomalies while also not being too compute-intensive. The factors for measuring detection latency include the time it takes to detect an identity or access event, and the time it takes to make an anomaly decision and risk classification once the event is detected. [20] The factors for measuring computational performance include processing throughput, CPU usage, memory usage and model inference time.

The evaluation takes into account various event generation rates to identify the impact of greater identity telemetry on processing delay and system efficiency. With a relatively stable detection latency, the proposed real-time pipeline is designed to handle more events with scalable processing elements under normal loads. Such measurements validate the framework's ability to ensure continuous monitoring of identities in operational cloud environments.

9.4. Scalability and Resource Utilization

Table 6: Scalability and Resource Utilization

Number of Events	Throughput (Events/s)	Latency (ms)	CPU (%)	Memory (%)
10,000	9,920	76	45	44
50,000	24,710	118	59	53
100,000	31,850	132	64	56

250,000	42,600	158	71	61
500,000	49,180	176	78	64

The scalability evaluation explores the performance of the proposed framework as the number of identity events, users, cloud entities, and number of concurrent access activities grow. Throughput, detection latency, CPU usage, memory usage and resource efficiency are measured via experiments using scaled-up workloads. It will also test if individual component of the framework can be scaled independently without any major degradation in the anomaly detection. Horizontal deployment of data processing and detection services is viewed as a means to assess how much the framework can handle as more and more data is being processed and what response times the framework can deliver. The outcomes give an idea of the practical applicability of the framework in large-scale cloud environments and show how the number of resources needed to detect the workloads relates to the growth of the workloads.

10. DISCUSSION

10.1. Analysis of Key Findings

The results of the experiments show the potential of the proposed automatic framework for detecting identity and access anomalies in dynamic cloud environments. The combination of behavioral profiling, machine learning driven anomaly detection, adaptive thresholding, and risk scoring can help the framework identify legitimate activities more readily than traditional monitoring methods can. Continued behavioral analysis can enhance security visibility and operations responsiveness, as evidenced by the observed improvements in detection accuracy, precision, recall, F1-score, false-positive reduction, detection latency and resource utilization.

The proposed framework offers a more flexible and adaptive way of identifying established and emerging identity and access anomalies as compared to traditional rule-based and static threshold-based methods. While typical solutions rely on predetermined policies and thresholds, the proposed solution continuously adds user and entity behavioral patterns, contextual attributes, and anomaly scores to the detection process. This allows for better detection performance while decreasing reliance on manually created rules and enables monitoring and scaling heterogeneous cloud identities and access activities.

10.2. Security and Operational Implications

The proposed framework has implications for the security and operational aspects as it allows for the continuous monitoring of authentication, authorization and privilege, session, and resource access. Risk-based classification can enable security teams to focus on high-risk events and automated alert generation can speed up the investigation and response process for security incidents. The framework can thereby enhance the ability to protect against credential misuse, account takeover, privilege abuse, and unauthorized access, as well as improve an organization's security operations in the cloud.

The framework has some drawbacks, although it has benefits. The effectiveness of its detection relies on both the quality and the completeness and representativeness of identity and access telemetry for behavioral profiling and model building. Detection performance may be impacted by changes in legitimate user behavior, by new cloud services, by new attack techniques not previously seen and by highly imbalanced anomaly data as well as concept drift. Also, there is computational and operational overhead arising from machine learning inference, model interpretability, and integrating with a variety of cloud security platforms, which could be explored in the future.

11. CONCLUSION AND FUTURE WORK

The proposed automated framework integrates various components, such as identity telemetry collection, behavioral profiling, feature engineering, machine learning-based anomaly detection, adaptive thresholding, risk scoring, and automated alert generation, to help detect identity and access anomalies in dynamic cloud environments. The framework is constantly monitoring authentication, authorization, privilege, session, and resource-access processes to detect deviations from normal activity patterns. The evaluation shows that an integrated approach can be more efficient in improving the effectiveness of anomaly detection, limit the number of false positives and false negatives, facilitate early detection, and be scalable in terms of identity security monitoring, when compared to the traditional approach.

The framework can help cloud security teams enhance continuous identity monitoring, prioritize activities that pose a high risk, scale down manual investigation workload, and bolster actions in the face of credential misuse, account takeover, privilege escalation and unauthorized access. The framework can be extended to future research with adaptive and online learning, enhanced behavioral and sequence models, explainable anomaly detection, federated learning in distributed cloud environments, and integration with automated incident-response systems. Additionally, the robustness, generalization, and operational scalability should be evaluated with large-scale, real-world cloud identity datasets and with diverse multi-cloud environments, as well.

REFERENCES

- [1] Buyya, R., Srirama, S. N., Casale, G., Calheiros, R., Simmhan, Y., Varghese, B., ... & Shen, H. (2018). A manifesto for future generation cloud computing: Research directions for the next decade. *ACM computing surveys (CSUR)*, 51(5), 1-38.
- [2] Moura, J., & Hutchison, D. (2016). Review and analysis of networking challenges in cloud computing. *Journal of Network and Computer Applications*, 60, 113-129.
- [3] Toosi, A. N., Calheiros, R. N., & Buyya, R. (2014). Interconnected cloud computing environments: Challenges, taxonomy, and survey. *ACM Computing Surveys (CSUR)*, 47(1), 1-47.
- [4] Voorsluys, W., Broberg, J., & Buyya, R. (2011). Introduction to cloud computing. *Cloud computing: Principles and paradigms*, 1-41.
- [5] Indu, I., Anand, P. R., & Bhaskar, V. (2018). Identity and access management in cloud environment: Mechanisms and challenges. *Engineering science and technology, an international journal*, 21(4), 574-588.
- [6] Indu, I., Anand, R., & Bhaskar, V. (2018). Identity and access management in cloud environment: Mechanisms and challenges. *Engineering Science and Technology, an International Journal*, 21(4), 574-588. <https://doi.org/10.1016/j.jestch.2018.05.010>
- Sharma, D. H., Dhote, C. A., & Potey, M. M. (2016). Identity and access management as security-as-a-service from clouds. *Procedia Computer Science*, 79, 170-174.

- [7] Ahmed, K. E. U., & Alexandrov, V. (2011). Identity and access management in cloud computing. In *Cloud Computing for Enterprise Architectures* (pp. 115-133). London: Springer London.
- [8] Shashanka, M., Shen, M. Y., & Wang, J. (2016, December). User and entity behavior analytics for enterprise security. In *2016 IEEE International Conference on Big Data (Big Data)* (pp. 1867-1874). IEEE.
- [9] Omar, S., Ngadi, A., & Jebur, H. H. (2013). Machine learning techniques for anomaly detection: An overview. *International Journal of Computer Applications*, 79(2), 33-41. <https://doi.org/10.5120/13715-1478>
- [10] Habiba, U., Masood, R., Shibli, M. A., & Niazi, M. A. (2014). Cloud identity management security issues & solutions: a taxonomy. *Complex Adaptive Systems Modeling*, 2(1), 5.
- [11] Laptev, N., Amizadeh, S., & Flint, I. (2015, August). Generic and scalable framework for automated time-series anomaly detection. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1939-1947).
- [12] Lan, Z., Zheng, Z., & Li, Y. (2009). Toward automated anomaly identification in large-scale systems. *IEEE Transactions on Parallel and Distributed Systems*, 21(2), 174-187.
- [13] Cherkasova, L., Ozonat, K., Mi, N., Symons, J., & Smirni, E. (2009). Automated anomaly detection and performance modeling of enterprise applications. *ACM Transactions on Computer Systems (TOCS)*, 27(3), 1-32.
- [14] Tolosana, R., Vera-Rodriguez, R., Ortega-Garcia, J., & Fierrez, J. (2015). Preprocessing and feature selection for improved sensor interoperability in online biometric signature verification. *Ieee Access*, 3, 478-489.
- [15] Arrington, B., Barnett, L., Rufus, R., & Esterline, A. (2016, August). Behavioral modeling intrusion detection system (BMIDS) using internet of things (IoT) behavior-based anomaly detection via immunity-inspired algorithms. In *2016 25th International Conference on Computer Communication and Networks (ICCCN)* (pp. 1-6). IEEE.
- [16] Ali, M. Q., Al-Shaer, E., Khan, H., & Khayam, S. A. (2013). Automated anomaly detector adaptation using adaptive threshold tuning. *ACM Transactions on Information and System Security (TISSEC)*, 15(4), 1-30.
- [17] Xu, R., Nikouei, S. Y., Chen, Y., & Faughnan, T. R. (2018). Real-time human objects tracking for smart surveillance at the edge. In *Proceedings of the 2018 IEEE International Conference on Communications (ICC)* (pp. 1-6). IEEE.
- [18] Hong, J., Liu, C. C., & Govindarasu, M. (2014). Integrated anomaly detection for cyber security of the substations. *IEEE Transactions on Smart Grid*, 5(4), 1643-1653.
- [19] Dukaric, R., & Juric, M. B. (2013). Towards a unified taxonomy and architecture of cloud frameworks. *Future Generation Computer Systems*, 29(5), 1196-1210.
- [20] Burramukku, N. R. (2016). Secure identity and access management integration for cloud-native network observability platforms. *International Journal of Engineering Development and Research*.
- [21] Waters, M. (2016). Evaluating identity and access management (IAM) as a Cloud service. *Research Gate*, September.
- [22] Crawford, H., Renaud, K., & Storer, T. (2013). A framework for continuous, transparent mobile device authentication. *Computers & Security*, 39, 127-136.
- [23] Jorquera Valero, J. M., Sanchez Sanchez, P. M., Fernandez Maimo, L., Huertas Celdran, A., Arjona Fernandez, M., De Los Santos Vilchez, S., & Martinez Perez, G. (2018). Improving the security and QoE in mobile devices through an intelligent and adaptive continuous authentication system. *Sensors*, 18(11), 3769.
- [24] Ashibani, Y., Kauling, D., & Mahmoud, Q. H. (2019). Design and implementation of a contextual-based continuous authentication framework for smart homes. *Applied System Innovation*, 2(1), 4.