

Original Article

Behavior-Based Detection of Abnormal Cloud Identity Activity

Uday Allala*Lead SRE Engineer, USA.*

Received: 01-06-2019

Revised: 01-26-2019

Accepted: 02-03-2019

Published: 02-05-2019

ABSTRACT

With the emergence of cloud computing environments, identity and access controls are playing a more prominent role in order to manage users, applications, and resources, and abnormal identity activity is becoming a major security issue. Known attack detection methods, such as rule or signature-based, work for known attacks but may be ineffective to detect subtle variations in legitimate user behavior that are new and evolving or have not been observed. This paper presents a behavior-based detection method to detect abnormal cloud identity activity by continuously monitoring identity and access events, and building behavioral profiles of users and entities. The proposed solution includes features that represent temporal, geographical, device, session, resource-access, authentication, and privilege aspects to describe normal identity behavior and anomalies from a known behavior pattern. The proposed detection model applies anomaly scoring and adaptive risk assessment mechanisms to determine whether a user's behavior is normal or potentially malicious and malicious identity behavior. The framework is tested with representative cloud identity activities, including both normal and abnormal activities, and is compared with traditional detection methods with accuracy, precision, recall, F1 score, ROC-AUC, false-positive rate, detection latency, and computational overhead. Experimental results show that the behavior-based analysis can enhance the detection of abnormal identity activity while providing appropriate detection performance and minimizing false alarm under different workloads. The key contribution of this research is an integrated behaviour-based framework that connects dynamic behavioural profiling, anomaly detection and identity risk scoring to enable continuous and adaptive detection of identity threats to the cloud.

KEYWORDS

Cloud Security, Identity and Access Management, Behavioral Analytics, User Behavior Analytics, Anomaly Detection, Identity Threat Detection, Machine Learning, Cloud Computing.

1. INTRODUCTION

1.1. Background and Motivation

Cloud computing has revolutionized the way organizations handle their applications, data, and digital resources, and identity and access management (IAM) plays a pivotal role in cloud security. In distributed cloud environments, authentication, authorization, session, and resource access events are generated continuously by users, administrators, applications, service accounts, and automated workloads. [1] Legitimate identity activity can differ significantly from user to user, location to location, device to device, applications to applications and operating time to operating time as cloud infrastructures grow more dynamic and interconnected. This complexity makes it easy for compromised credentials, account takeover, privilege misuse, unauthorized resource access, and other threats related to identity to stay hidden amidst the vast amount of legitimate transactions. To this end, security measures need to go beyond the static concept of access control and continuously assess if the observed identity behavior conforms to the behavioral patterns established.

The traditional approach to identity monitoring is based on pre-defined rules, signatures, fixed thresholds and known indicators of compromise. While they work well for detecting clear-cut security breaches, they can produce too many false positives if there are legitimate changes in behavior, and they are ineffective at detecting new or low-and-slow attacks that do not conform to a pattern. An alternative approach is behavior-based detection, which builds dynamic profiles of normal identity activity and detects the deviations in both temporal, geographic, authentication, device, session, resource-access, and privilege areas. This research problem of distinguishing abnormal cloud identity activity from the legitimate behavioral variation is the subject of this study, and with a low detection latency and manageable computational overhead. The aims of the research are to create a continuous behavior-based detection framework, to engineer meaningful identity behavioral features, to model normal user and entity activity, to calculate behavioral anomaly and risk scores and to assess the effectiveness of detection under various workloads and attack scenarios. The key contributions are an integrated behavioral profiling and anomaly detection framework, adaptive identity risk assessment, and a comprehensive evaluation methodology for measuring detection accuracy, false-positive behaviors, latency, and scalability.

2. BACKGROUND AND THEORETICAL FOUNDATIONS

2.1. Cloud Identity and Access Management

Cloud Identity and Access Management (IAM) defines the processes for identity authentication, authorization of resource access, privilege management, and security policy enforcement for distributed cloud environments. Cloud platforms produce an abundance of events related to identities, such as attempts to access identity, identify actions, access to resources, changes to permissions, device information, geographic location, and more. [2] Such events serve as a good behavioural representation of the typical user/non-human interaction with a cloud service. Checking for authorization of an identity does not necessarily imply authorization of an action, however. This is especially valuable for compromised accounts, credential misuse, privilege abuse, and account takeover, where attackers can use legitimate credentials to make their actions appear as if they are

coming from the legitimate account holder. Hence, there is a need for ongoing identity activity analysis to supplement traditional IAM controls.

2.2. Identity and Access Activity in Cloud Environments

Static rules and manual monitoring are not enough to detect subtle behavioural deviations in the massive number of cloud identity events, their velocity and variety. To overcome this challenge, User and Entity Behavior Analytics (UEBA) builds behavior profiles based on authentication, temporal, geographic, device, session, resource access, API, and privilege attributes. Anomaly detection based on identity is then used to detect deviations from these profiles to detect potentially abnormal activity. Behavioral profiling offers a moving snapshot of normal identity behavior and risk-based assessment gives more weight to deviations as they become more serious, contextually more serious, and more cumulative. Using machine learning, such as supervised, unsupervised or semi-supervised learning, can also enable it to learn complex behavioral patterns and identify anomalies that do not fit into pre-defined rules. Combined, behavioral profiling, anomaly detection, adaptive risk scoring and machine learning offer a theoretical basis for ongoing detection of abnormal cloud identity activity, while minimizing detection sensitivity, false-positive reduction and computational impact.

3. RELATED WORK

3.1. Rule-Based Identity Anomaly Detection

Rule based identity anomaly detection looks for suspicious cloud activities based on the predefined security policies, thresholds, signatures, and access-control rules. Common rules identify events like multiple failed authentication, logins from restricted locations, unusual privilege modification, access to unauthorized resources and violations of defined access policies. [3] These methods are relatively easy to implement, easy to compute and can detect very specific and known security violations. However, it requires a lot of manual rules to be set in order to be able to detect subtle changes in behaviour, previously unknown attacks and legitimate behaviour that is abnormal only in certain contexts for a user. Cloud workloads and identity behaviors can also change frequently, leading to a need for ongoing rule maintenance, which can result in higher false positive rates and lower detection adaptability.

3.2. Machine Learning-Based Anomaly Detection

A more adaptive solution is to use machine learning anomaly detection, which learns patterns from historical identity and access activity and detects anomalies that deviate from the learned behavior. [4] There are supervised and unsupervised and semi-supervised methods for classification of known malicious and legitimate activities, and unsupervised and semi-supervised methods for identification of anomalous activities in the absence of abundant labeled attack data. Multiple behavioral aspects such as authentication frequency, access time, resource usage, geographic location, device characteristics, privilege activity, etc. can be included in models like Random Forest, Support Vector Machine, Isolation Forest, clustering algorithms, and gradient-based classifiers. While machine learning can be more flexible in detecting patterns than static rules, this flexibility relies on the quality of the data, how the data is represented as features, the configuration of the model, and

the ability to learn and adapt to new behavioral patterns. Furthermore, the behavior of the clouds can be dynamic, resulting in a need for ongoing model updates for maintaining consistent anomaly detection.

3.3. Research Gaps

Previous studies have shown that rule-based, statistical, machine learning and deep learning techniques can be effective for the detection of abnormal identity activity, but there are some drawbacks. Traditional detection models rely on specific indicators or attack signatures and many machine learning models rely on large off-line datasets, but do not necessarily reflect the ongoing evolution of legitimate identity usage. While deep learning can capture complex patterns, it can also require more computational power and decrease the interpretability, making it harder to deploy in real-time cloud security operations. There is a clear need for research to build an integrated behavior-based detection system that continuously builds a dynamic identity profile, assesses multi-dimensional behavioral deviations, adapts detection thresholds and integrates anomaly detection with risk-based prioritization. This is required to increase the detection of subtle identity threats without increasing false positives, be responsive in real-time, scale to large numbers and be easy to operate.

4. PROBLEM FORMULATION AND THREAT MODEL

4.1. Problem Definition

Cloud identity environments produce a constant stream of authentication, session, [5] API usage, resource access and privilege-related events, and it's essential to differentiate between normal identity behavior and potentially malicious. An identity activity event is represented as $(e_i = \{u, t, l, d, s, r, p, a\})$ where (u) represents the identity, (t) the time stamp, (l) the identity's location, (d) the device used, (s) the session context, (r) the resource accessed, (p) the privilege information, and (a) the action performed. The goal of the proposed approach is to build a dynamic behavioral profile of each identity and decide if a new event or series of events is significantly different from the normal behaviour. The detection problem is then stated as finding abnormal identity activity with minimal false positive, detection latency and computational overhead. The anomaly assessment is then used to produce an identity risk score for continuous monitoring and prioritization of potentially compromised identities.

4.2. Cloud Identity Activity Model

The cloud identity activity model is a stream of time-dependent events that represents the interactions between users and non-human identities and cloud services. There are contextual attributes for each event, including authentication type, access time, location of source, device or application information, session duration, API activity, resource type, access frequency, and privilege level. An identity (u) can be characterized by its behavioral profile $(B_u = \{T_u, G_u, D_u, S_u, R_u, P_u\})$, where, and (T_u) , (G_u) , (D_u) , (S_u) , (R_u) , and (P_u) denote temporal, geographic, device, session, resource-access, and privilege behavior, respectively. This profile is continually compared to a series of observed events in order to measure behavioral deviation. This model allows the detection process to take into account the context and history of identity activity,

not evaluate individual accesses, and thus better identify abnormal activity in dynamic cloud environments.

4.3. Threat Model

The adversary in the threat model is assumed to get legitimate credentials or use an existing identity, and then use legitimate authentication methods to interact with cloud resources. [6] The attacker can try to steal credentials, take over an account, escalate privileges, access resources without authorization, use excessive permissions, log in from an unusual location or at an unusual time, and/or try to remain anonymous to the legitimate user. The proposed detection framework does not rely entirely on the detection of static access-control violations, but assumes that attacks can also occur as deviations from temporal, geographic, device, session, resource, authentication, and privilege behavior. The first security goal is then to identify changes in the user's behavior and build up of risk, while avoiding false alarms from legitimate changes in user activity.

4.4. Detection Requirements

The successful behavior-based cloud identity detection system should offer real-time monitoring, multi-dimensional behavior analysis, adaptive anomaly detection, and risk prioritization of alerts. The framework should be able to detect isolated abnormal events, as well as correlated sequences of suspicious activities, and have high precision, recall and F1-score. It should also provide low latency processing for real-time or near real-time identity monitoring, and be scalable as the number and speed of cloud events grow. The detection system should also be flexible enough to detect new patterns of legitimate activity, filter out false-positives, and differentiate between high-risk deviations and low-risk variations in activity. Finally, it should deliver understandable risk data to security operations. The requirements set the groundwork for the assessment of the proposed framework on the aspects of detection accuracy, false positive and false negative rates, latency, scalability and computational efficiency.

5. PROPOSED BEHAVIOR-BASED DETECTION FRAMEWORK



Fig 1 - Proposed Behavior-Based Detection Framework for Abnormal Cloud Identity Activity

5.1. Framework Overview

The proposed behavior-based detection framework offers a continuous detection method that works by analyzing the identity events, building behavioral profiles, measuring deviations, and assigning risk levels to identify abnormal cloud identity activity. [7] The framework is composed of a

series of interdependent phases of acquiring identities, preprocessing, engineering behavioral features, user profiling, building a baseline, anomaly detection, adaptive risk scoring, alert prioritization, and security response. The framework does not rely on static rules to monitor identity actions; instead, it considers the actions' historical and contextual behavior. The entire processing pipeline is engineered to continuously convert raw cloud identity events into meaningful security decisions, while ensuring a high level of detection accuracy, scalability and a low response time.

5.2. Identity Activity Data Acquisition

It is an activity data acquisition layer that fetches identity activity events from identity and access management systems, authentication services, API gateways, cloud audit logs, resource access logs and security monitoring platforms. Information collected in the events includes user identity, time of access, authentication mechanism, source location, device characteristics, session information, resources accessed, API operations, and privilege changes. The heterogeneous events are converted into a common event representation for consistent downstream analysis. Continuous collection allows for capturing changes in identity behavior in near real time and allows historical activity to be used to build accurate profiles and identify deviations from normal activity.

5.3. Behavioral Feature Engineering

Behavioral feature engineering takes raw identity events and turns them into useful features that capture the typical interactions between identities and cloud services. [8] The framework identifies temporal features (including login frequency and access time), geographic features (including location changes and travel distance), device features and session features, resource-access frequency, API activity, authentication patterns and privilege-related characteristics. Individual-event features and sequence level features are both regarded as detection of anomalies or changes in behavior over time. Normalizing features and aggregating them contextually to minimize identities' and workloads impact. The multi-dimensional feature representation allows the anomaly detection engine to differentiate between normal variations in behaviour and potentially malicious identity activity.

5.4. Anomaly Detection Engine

The anomaly detection engine compares the identity profile's behavioral features with the behavior that is received and the normal behavior patterns to see how much the behavior is deviating. [9] Anomaly scores are computed based on deviations for each identity activity sequence for the temporal, geographic, device, session, resource, authentication and privilege dimensions. The detection mechanism may utilize statistical and machine learning approaches to detect both known as well as unknown behavioral anomalies. The engine takes into account many associated deviations, instead of one, to make detection more robust and less prone to false alarms. If the activities are found outside the dynamically established anomaly boundary, they are deemed to be potentially abnormal and sent to the adaptive risk-scoring phase.

5.5. Adaptive Risk Scoring

The adaptive Risk Scoring functionality transforms identified behavioral anomalies and assigns them a continuous identity risk score that indicates the level of suspicious activity and its relevance within the context. All the above factors and the frequency/persistence of abnormal events are used to create a risk score. The detection thresholds are dynamically adjusted in accordance with the changes in the patterns of the legitimate behaviors, so that the framework can be sensitive to new threats without triggering too many alerts in the event of normal change of legitimate behavior. High-risk activities have priority for immediate investigation, while moderate deviations can be further monitored and observed as behavior. This adaptive training activity allows for ongoing identity risk assessment as well as consistent identity alert generation and security response.

6. BEHAVIORAL FEATURE ENGINEERING AND USER PROFILING



Fig 2 - Dynamic User Behavioral Profiling and Baseline Construction Process

6.1. Identity Activity Features

Table 1: Threat Scenarios and Behavioral Indicators

Threat Scenario	Behavioral Indicators	Expected Risk
Account Takeover	New device, unusual location, abnormal login time, unusual resource access	High
Credential Misuse	Valid credentials from unfamiliar location or device	High
Privilege Abuse	Unexpected privilege elevation or administrative activity	High
Insider Threat	Abnormal resource access, unusual data activity, excessive permissions	High
Impossible Travel	Rapid access from geographically distant locations	High
Abnormal Login Activity	Repeated failures, unusual login frequency, unfamiliar access time	Medium-High
Unauthorized Resource	Access to previously unused or restricted resources	High

Access		
Lateral Movement	Sequential access to multiple unrelated cloud resources	High
Threat Scenario	Behavioral Indicators	Expected Risk

Identity activity represents the basic attributes of interactions between cloud identities and protected services. The proposed options collect authentication attempts, successful and failed login attempts, [10] API calls, number of times logged in, requests for access, interactions with resources, changes to permissions and types of identities. These features allow the identity behavior to be represented in a structured manner, and can be used to differentiate between routine activities and suspicious changes from the detection mechanism. The framework can include authentication, access, session, and privilege information to provide a complete behavioral representation of human users and non-human identities like service accounts and automated applications.

6.2. Temporal Behavioral Features

The temporal behavioral features describe the time and the frequency of the behaviors of an identity that involve cloud activities. Examples of indicators are login times, number of accesses, session time, time between session actions, unusual activity hours, and deviation from normal daily/weekly activity. The framework also looks at the temporal behavior of the device and the difference between current activity and historical behavior patterns to detect deviations, such as bursts of activity, accessing during non-working hours or authenticated attempts that are too frequent. Because malicious users may try to conduct their actions at times or frequencies that are drastically different from the legitimate identity's behavior, temporal analysis is especially useful in detecting compromised credentials and account takeover.

6.3. Geographic and Location Features

Geographic features are used to represent the spatial aspects of the identity activity: the source location, the geographic distance between each access, the frequency of the location, and changes in the identity's normal access region. [11] These features are used to detect the suspicious situations, e.g., access from areas not seen before, rapid geographic change, and impossible-travel pattern. Geographic information is analyzed in conjunction with temporal and authentication information, and it is not used as a standalone identifier, which enables the framework to differentiate between valid trips or remote accesses and potentially tampered credentials. This contextual method decreases unwanted notifications due to typical shifts of user position.

6.4. Behavioral Profile Construction

The construction of behavioral profiles is based on a representative type of normal behavior for each cloud identity formed from historical and continually monitored behavioral features. The profile includes temporal, geographic, device, session, resource-access, authentication and privilege characteristics for identity (u), and summarizes their typical ranges, frequencies, relations, and patterns in context. It is constantly updated as more legitimate activity is witnessed, allowing the framework to account for individual variations among users, rather than adopting one standard for

all identities. In detection mode incoming activity is compared to the corresponding profile, which makes it possible to measure behavioral deviations and feed the anomaly detection engine, with identity-specific, contextual information.

6.5. Dynamic Behavioral Baselines

Dynamic behavioral baselines offer a flexible definition of what is being considered normal identity activity as it changes over time and with the changing workloads in the cloud. [12] The framework uses trusted observations to continually update baseline statistics and behavioral patterns instead of a fixed threshold, and avoids that suspicious events have a disproportionate impact on the normal profile. Changes in working times, new devices, different access patterns, different resource requirements are all examples of changes that adaptive baseline mechanisms address. The baseline allows the detection system to detect real behavioral drift and not wrongly flag malicious deviations, thus facilitating the continuous detection of anomalies with a system that is more adaptable, less prone to false alarms, and more effective in detecting anomalies in dynamic cloud environments.

7. BEHAVIOR-BASED ANOMALY DETECTION MODEL

7.1. Normal Behavior Modeling

Normal behaviour modeling is done to create a reference representation of valid identity activities based on the behaviour features and profiles built. For each identity, the model represents the patterns that are expected between each of the following dimensions: temporal, geographic, device, session, authentication, resource-access, and privilege. Activity and contextual relationships can be characterized using statistical distributions and machine learning based representations of the normal range of activity. [13] The learned patterns are continually used to compare incoming identity events to see if the observed behavior is consistent with the profile of an identity. This identity based modeling allows the framework to detect deviations without the assumption that all users behave alike in accessing the system.

7.2. Behavioral Deviation Measurement

Behavioral deviation measurement is a method that measures the difference between an observed identity activity and the normal behavioural profile. A deviation score ($D(X_t, B_u)$) is computed from a number of behavioral dimensions such as temporal, geographic, device, session, resource, authentication and privilege features. Individual variations are adjusted for and standardized to avoid the influence of one feature on the whole assessment. [14] Several or correlated deviations may be given increased weight as evidence of possible misuse of an identity, as long-running abnormal behavior is more likely to be evidence of misuse of an identity than a single deviation. This deviation measure is used as the main input for anomaly scoring and adaptive risk assessment.

7.3. Identity Risk Score Calculation

The identity risk score aggregates behavioral deviation, anomaly severity, event criticality, historical risk, and contextual information into a single risk score to measure potential identity compromise. A generalized formulation is $(R_u = \sum_{i=1}^n w_{iD_i} + C + A + H)$, where (D_i)

are deviations from the norm across the behavioral dimensions, (w_i) are the corresponding weights, (C) is the contextual risk, (A) is the severity of anomalies, and (H) is the relevant historical risk. The score is continually updated as new identity events are seen, allowing the score to be accumulated when a number of suspicious events are seen over the period of a short time. Identities can also be placed in low-risk, medium-risk, and high-risk states, based on established adaptive risk ranges, and then prioritized for investigation and response based on the severity of the behavioral anomaly detected.

8. SYSTEM ARCHITECTURE AND IMPLEMENTATION

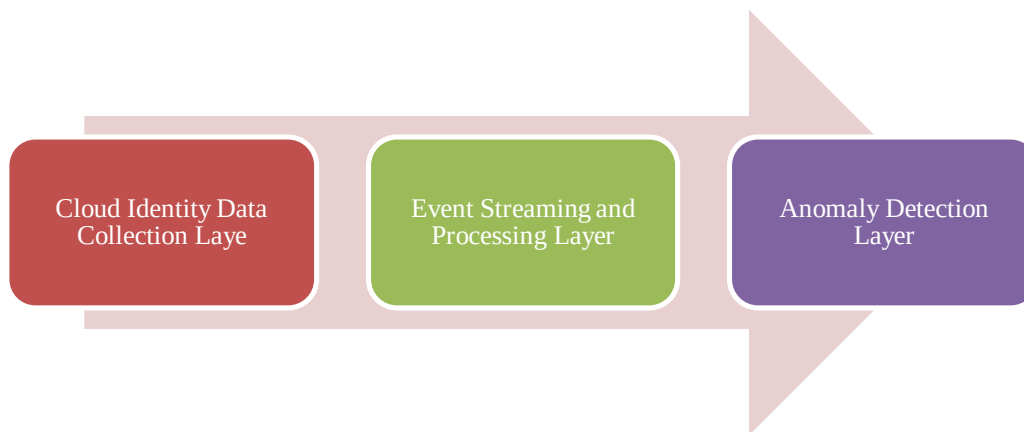


Fig 3 - System Architecture and Implementation

8.1. Cloud Identity Data Collection Layer

The cloud identity data collection layer continually collects data related to identity and access events from cloud IAM platforms, authentication services, audit logs, API gateways, resource-access logs, and security monitoring systems. [15] Applicable attributes of each event are captured including identity information, timestamp, authentication method, originating location, device characteristics, event session information, resources accessed, API operations, and privilege changes. The events are gathered and converted into a consistent format to be confidently passed to the processing pipeline. The continuous data acquisition guarantees that the data for routine and suspicious identity activities are available for behavioral analysis, which allows the framework to keep up to date identity profiles and have the ability to pick up new deviations with little delay.

8.2. Event Streaming and Processing Layer

The event streaming and processing layer is the scalable infrastructure needed to process high volume cloud identity events in real time. Events coming in are ingested, ordered, filtered, normalized, and transformed into analytical records in real-time and passed on to the parts of the behavioural analytics. Event aggregation, time window analysis, session reconstruction, duplicate removal, and contextual enrichment are supported by stream-processing mechanisms with low processing latency. The layer will be designed to be scalable and will be expanded to meet the growing number of identities and the number of cloud workloads, permitting behavioral features to be generated on an ongoing basis without causing significant processing bottlenecks.

8.3. Anomaly Detection Layer

The anomaly detection layer performs analysis of the anonymised behaviour features and uses a comparison to the existing behavioural profile and dynamic baselines to identify current identity activity. [16] It compares variations in time, location, device, session, authentication, resource-access, and privilege and creates a score of anomaly for each identity or activity sequence. Several behavioral anomalies can be correlated to detect more complex or infrequent threats which do not necessarily break traditional access-control policies. If there is a significant deviation with the activity, it is passed to the risk assessment layer, enabling the framework to differentiate between potentially malicious activity and legitimate variations, while supporting ongoing and adaptable detection.

8.4. Real-Time Detection Workflow

The real-time detection workflow covers the entire process of generating identities to responding to security. A new cloud identity event is first captured and pushed to the processing layer where it is normalized, enriched and behavioral features are extracted. The resulting feature vector is then used to compare the behaviour to the behavioural profile and dynamic baseline of the identity, in order to produce behavioural deviation and anomaly scores. The anomaly information is then added with contextual and historical information to produce an identity risk score, which is then prioritized by severity of the risk and passed on to suitable security response mechanisms. This ongoing process allows for the immediate detection, evaluation and response to unusual identity usage, as well as flexibility in adapting to evolving adoption of cloud services.

9. EXPERIMENTAL METHODOLOGY

9.1. Experimental Environment

The experimental setup allows the assessment of the proposed behavior-based detection system in representative cloud identity workloads. It is a cloud-based identity activity processing pipeline with authentication services, access-control components, audit-log sources, event-stream processing, behavioral analytics, anomaly detection, and risk-scoring modules. [17] Experiments include different volume of identity activity and adequate computational capabilities are provided for real-time feature extraction, behavioural profiling, anomaly detection and risk assessment. Each experiment has the same environment setup to allow for a comparison of the proposed approach with baseline detection methods.

9.2. Dataset Description

The dataset represents a set of representative cloud identity and access events with normal and abnormal activities of both users and entities. All events contain details like identity identifier, authentication result, timestamp, source location, device details, session properties, resources accessed, API activity and privilege related operations. Normal activities are the typical authentication and resource access pattern while abnormal activities are activities related to credential compromise, account takeover, privilege misuse, abnormal access location, abnormal resource usage, and suspicious session behavior. The dataset is structured to maintain temporal and

behavioral relationships to enable building identity-specific profiles and assessing continuous anomaly detection.

9.3. Normal and Abnormal Activity Generation

Normal and abnormal identity activities are produced to create controlled experimental scenarios to evaluate detection effectiveness. [18] Normal behavior is based on representative authentication, session, geographic, device, resource-access and privilege patterns, with deviations in abnormal behavior including unusual logon times, unknown locations, fast geographic change, new devices, excessive resource usage, escalation of privileges and unusual API usage.

Variations in behavior are implemented at varying levels to assess the ability of the proposed framework to identify minor legitimate variations and significant security anomalies. The activities generated are sequentialized over time for realistic continuous monitoring and baseline construction of behavior.

9.4. Baseline Methods

The proposed behavior-based detection approach is evaluated and compared with representative baseline approaches to see how effective the proposed approach is. The base detection approaches are rule based, statistical threshold based, traditional machine-learning based anomaly detection and behavioral detection with fixed risk scoring. These methods have a variety of complexities and detection power and can be tested separately to assess the contribution of behavioral profiling, adaptive thresholds and risk assessment.

All the methods are tested on the same set of input activities and experimental conditions, so that the differences in performance can be tested under the same conditions and only depend on the activity to be detected and the methodology used for the detection.

9.5. Evaluation Metrics

To measure the ability of each method to distinguish normal and abnormal identity activity, detection performance is measured through accuracy, precision, recall, F1-score and ROC-AUC. The false-positive rate and false-negative rate are added to evaluate the operational reliability of security alerts and the detection latency evaluates the responsiveness of the framework under real-time workloads. The efficiency of deployment is also measured by computational overhead which includes processing time and use of resources. These metrics offer a comprehensive analysis of the effectiveness of detection, reliability of alerts, real-time responsiveness, and scalability.

10. RESULTS AND PERFORMANCE EVALUATION

10.1. Overall Anomaly Detection Performance

Table 2: Overall Anomaly Detection Performance Comparison

Detection Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)
Rule-Based Detection	84.6	82.1	78.5	80.3	81.7

Statistical Thresholding		87.3	85.6	82.9	84.2	85.8
Isolation Forest		91.2	89.4	88.1	88.7	90.6
Random Forest		93.5	92.1	91.3	91.7	94.2
Proposed Framework	Behavior-Based	96.4	95.2	94.7	94.9	97.1

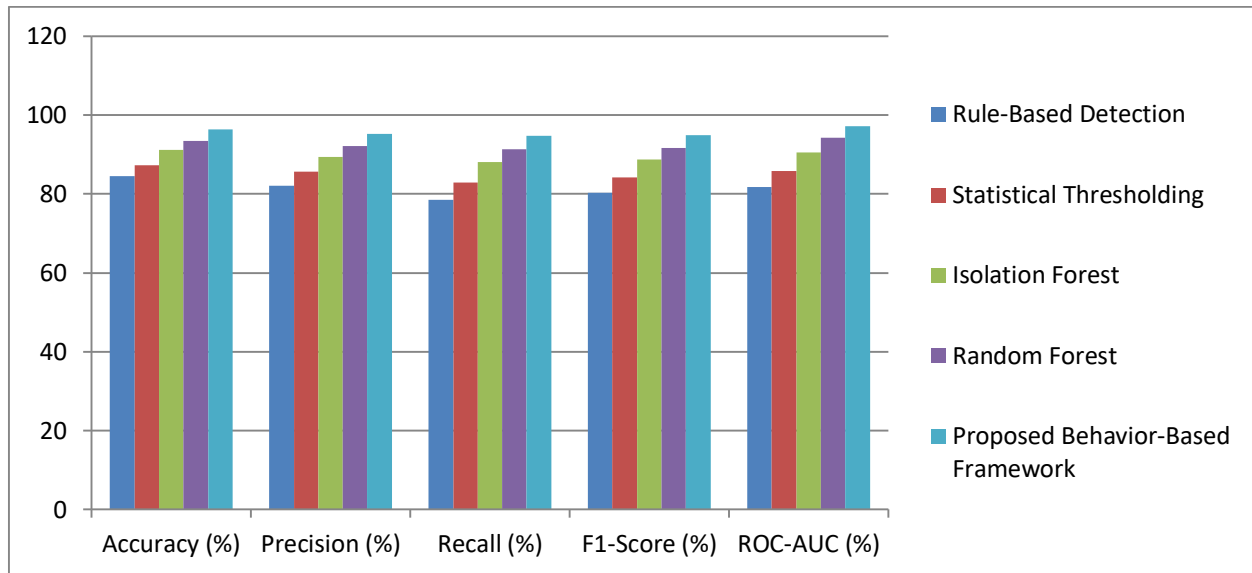


Fig 3 - Graphical Overall Anomaly Detection Performance Comparison

The overall performance of the proposed behavior-based detection framework is assessed by the accuracy, precision, recall, F1 score, and ROC-AUC. The proposed solution delivers more effective detection results than rule-based, statistical, and traditional machine-learning solutions, thanks to its multi-dimensional behavioral characteristics, user profiles, anomaly detection, and adaptive risk scoring. The results show that continuous behavioral analysis can achieve better identification of abnormal cloud identity behaviors, while keeping the proportion of false-positive and false-negative results well balanced. The overall performance comparison of the evaluated detection approaches is shown in Table V.

10.2. Precision, Recall and F1-Score Analysis

Precision, recall, and F1-score offer a comprehensive evaluation of how the framework performs to accurately identify abnormal identity activities while minimizing the number of false alerts. The proposed approach will likely continue to provide high precision, because multiple correlated behavioral deviations will not generate an alert, while high recall will ensure that it will detect the various abnormal identity scenarios, including credential misuse, account takeover, privilege abuse, and unusual resource access. The resulting F1-score illustrates the overall completeness of detection and reliability of the alerts.

10.3. Comparison With Baseline Methods

The proposed framework is evaluated with rule-based detection, statistical thresholding, Isolation Forest and Random Forest with the same experimental setup. While conventional rule-

based methods offer lower detection capability as they rely on predetermined patterns, statistical methods and machine learning are more effective in detecting anomalies using data analysis. The proposed framework delivers the highest overall performance due to the ability to add extra contextual data in the form of dynamic behavioral profiling and adaptive risk scoring to help distinguish legitimate behavioral changes from suspicious identity activity.

10.4. Detection Performance under Different Workloads

Table 3: Detection Performance Under Different Cloud Identity Workloads

Workload (Events/s)	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Latency (ms)
1,000	97.1	96.3	95.8	96.0	82
5,000	96.8	95.9	95.2	95.5	96
10,000	96.4	95.2	94.7	94.9	118
25,000	95.9	94.6	93.8	94.2	167
50,000	95.3	93.9	93.1	93.5	241

The framework is tested with progressively higher cloud identity-event workloads to see if the detection quality is stable as the number of events increases. The results of experimental workloads between 1,000 and 50,000 events/s show that accuracy, precision, recall and F1-score gradually decline as the workload increases but stay at an effective detection level. The results prove the proposed architecture can process a high volume of identity activity and maintain good detectability of anomalies with reasonable latency. The detection performance is summarized in various workloads.

10.5. Detection Latency Analysis

Table 4: Detection Latency and Computational Resource Utilization

Workload (Events/s)	Throughput (Events/s)	Latency (ms)	CPU (%)	Memory (%)	Processing Overhead (%)
1,000	1,000	82	18.4	31.2	7.8
5,000	5,000	96	27.6	36.5	9.4
10,000	10,000	118	35.8	42.7	11.2
25,000	25,000	167	49.3	51.8	14.6
50,000	50,000	241	67.5	64.2	18.9

Detection latency is the amount of time to process an identity event and generate an anomaly or risk decision. The results demonstrate that latency rises as workload grows due to the increased number of events that need to be processed and analyzed; however, the proposed architecture based on the streaming paradigm ensures appropriate near real-time response time. Lower workloads allow for a behavioral analysis with significantly lower latency, whereas higher workloads add in a moderate amount of delay that is not much of a burden for continuous security monitoring. Table VII shows latency in addition to throughput and computational resource utilization.

10.6. False Positive and False Negative Analysis

Analysis of false positive and false negative rates to assess the confidence in abnormal identity detection. For false positives, legitimate changes in work schedules, work locations, devices, or resources requirements could be mistaken for malicious activity, while false negatives are abnormal activities that are not detected. The proposed adaptive baseline and correlated contextual risk-scoring mechanisms mitigate false positives by looking at historical behavior and multiple correlated indicators, while multi-dimensional anomaly detection boosts sensitivity to subtle malicious behavior and lowers false negatives.

10.7. Risk Scoring Performance

The effectiveness of the adaptive risk-scoring mechanism is measured through the ability to distinguish between low, medium and high-risk identity activities. The risk score combines factors like behavioral deviation, event criticality, privilege level, resource sensitivity, historical behavior, and abnormal-event persistence. A key objective of the results is to see that high-risk activities, like account takeover, privilege escalation and unauthorized access to sensitive resources can be given significantly higher scores than routine identity events. This prioritization enables the security teams to concentrate on investigation resources on the most important threats.

11. DISCUSSION

The proposed behavior-based detection model shows that the continuous monitoring of identity activity is more valuable than only relying on static access policies and/or attack signatures. The framework can detect behavioral anomalies related to account takeover, credential abuse, privilege escalation, abnormal logon activity, and abnormal resource access by examining temporal, geographic, device, session, authentication, resource-access, and privilege-related characteristics. The profile of the specific behaviour of the identity together with the multi-dimensional anomaly score allows the system to differentiate between normal and suspicious activity and to detect attack patterns which are not typical. In addition, dynamic behavioral baselines enable the framework to be flexible enough to reflect legitimate shifts in user activity, such as variations in working hours, access locations, devices, and resource requirements, which helps make it more suitable for an ever-evolving cloud environment.

The use of adaptive thresholds, contextual risk scoring and persistent event processing increases the usefulness of the anomaly detection in terms of operation, reducing the number of alerts for minor or regular variations in the behavioral pattern. Real-time processing can immediately evaluate suspicious identity activities and prioritize them based on the perceived risk, aiding in quick investigation and response. The modular design also scales by enabling the event collection, behavior analysis, anomaly detection and risk assessment modules to handle larger volumes of identity events without the need for the other modules to grow. On a more practical security standpoint, the framework can be used alongside current cloud IAM, SIEM, security-response tools to offer constant visibility of identity risks and prioritize high-risk actions to be investigated. The features of these capabilities make behavior-based detection a great fit for cloud environments where identities, workloads, resources and access patterns are constantly shifting.

12. SECURITY AND OPERATIONAL ANALYSIS

The proposed detection framework based on behavior rather than on account details enhances identity security by detecting deviations that are linked to account takeover, privilege abuse, insider threats, and credential misuse. Compromised identities might have unusual logon times, logon from unusual geographic locations, use of new devices, abnormal session attributes, access of excessive resources, or use of unexpected privileges. These behavioural attributes can be combined with an identity-specific history of the individual to distinguish between valid and suspicious use. This can be especially useful for identifying attacks that evade traditional authentication mechanisms, as the attacker could have a legitimate identity and behave in a manner that does not align with the identity owner.

The ability to correlate behavioral characteristics across identities, sessions, resources and access sequences can also aid the detection of lateral movement, especially if an identity starts accessing resources that are unrelated and shows up with new privileges and access patterns. Continuous identity risk assessment uses behavioral deviations, severity of events, historical context and sensitivity of resources to keep up to date with the risk state of each identity. The risk information can then be combined with cloud security operations, such as SIEM platforms, security monitoring systems, incident-response workflows, and access-control mechanisms, to enable high-risk identities to be investigated or the risk mitigated with adaptive security actions. Thus, the framework offers an operational layer that is used in conjunction with traditional IAM controls and converts ongoing identity activity into security intelligence.

13. LIMITATIONS AND FUTURE RESEARCH

There are several issues with the proposed framework regarding the quality of the datasets, behavioural variations and the ability of detecting new attacks. The quality and diversity of the cloud identity datasets used can affect the results of these experiments and impact their reliability and generalizability. Furthermore, the construction of behavioral profiles and detection of anomalies may be impacted by incomplete, noisy, imbalanced, and/or misidentified identity events. Another challenge with behavioral drift is that legitimate user behaviors may evolve over time as a result of new roles, devices, locations, workloads, or organizational processes. Dynamic baselines can be used for gradual changes in behavior, but rapid behavior changes can be misinterpreted as anomalies. Moreover, very novel attacks, which cause small variations in the way the attacks are conducted, might be hard to recognise without extra contextual data or threat intelligence.

The cost of real-time behavioural analysis can add computational and infrastructure burden, especially in large cloud deployments with lots of identity events. Scales processing resources and efficient model management is essential for continuous feature extraction, profile updates, anomaly scoring, and risk assessment. Additionally, privacy and data-governance considerations should be taken into account as behavioral monitoring could involve sensitive data, including location, device, access patterns, and user activity history. To this end, future studies should focus on privacy-preserving behavioural analytics, federated and distributed learning, explainable anomaly detection, online learning for behavioural drift and sophisticated deep-learning and/or graph-based models for

complex identity relations. The framework should also be tested with larger scale real-world datasets, various cloud platforms, new attack scenarios, and real-world deployments over an extended period of time in future studies to assess robustness, scalability, and security effectiveness.

14. CONCLUSION

This research proposed an abnormal cloud identity detection method based on behavior analysis by continuously analyzing the behavior of identity and access. The framework supports identity activity collection, behavioral feature engineering, dynamic user profiling, anomaly detection, adaptive risk scoring, and real-time alert generation to detect deviations from the behavior models. The analysis illustrates the value of the temporal, geographic, device, session, resource-access, authentication and privilege-related characteristics for detecting identity-based threats such as account compromise, credential misuse, and privilege abuse. The key features are an integrated behavioral detection architecture, dynamic behavioral baselines, multi-dimensional anomaly assessment and continuous identity risk evaluation.

The proposed approach offers a practical complement to traditional cloud IAM and security monitoring methods that rely on pre-defined access policies and/or known attack signatures, by continually analyzing the behavior of the identities. The adaptive profiling and risk-based prioritization features help to detect more accurately and minimise unnecessary security alerts, allowing for prompt investigation of high-risk activities. While there are still some issues to address such as behavioral drift, diversity of datasets, computational challenges, privacy, and new attacks, the framework provides a basis for scalable and adaptive identity monitoring in the cloud. For future work, extensions with privacy-preserving learning, explainable detection, online learning, and advanced behavioral models could enhance and complement the continuous identity threat detection in the changing cloud context.

REFERENCE

- [1] Zheng, R., Chen, J., Zhang, M., Wu, Q., Zhu, J., & Wang, H. (2018). A collaborative analysis method of user abnormal behavior based on reputation voting in cloud environment. *Future Generation Computer Systems*, 83, 60-74.
- [2] Kim, T., Choi, Y., Han, S., Chung, J. Y., Hyun, J., Li, J., & Hong, J. W. K. (2012, April). Monitoring and detecting abnormal behavior in mobile cloud infrastructure. In *2012 IEEE network operations and management symposium* (pp. 1303-1310). IEEE.
- [3] 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.
- [4] Almulla, S. A., & Yeun, C. Y. (2010, March). Cloud computing security management. In *2010 Second International Conference on Engineering System Management and Applications* (pp. 1-7). IEEE.
- [5] Alcaraz Calero, J. M., Edwards, N., Kirschnick, J., Wilcock, L., & Wray, M. (2010). Toward a multi-tenancy authorization system for cloud services. *IEEE Security & Privacy*, 8(6), 48-55. <https://doi.org/10.1109/MSP.2010.194>
- [6] 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.
- [7] Mangiuc, D. M. (2012). Cloud identity and access management-A model proposal. *Journal of Accounting and Management Information Systems (JAMIS)*, 11(3), 484-500.
- [8] 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.

- [9] Kaur, H., Singh, G., & Minhas, J. (2013). A review of machine learning based anomaly detection techniques. arXiv preprint arXiv:1307.7286.
- [10] Mazzawi, H., Dalal, G., Rozenblatz, D., Ein-Dorx, L., Niniox, M., & Lavi, O. (2017, April). Anomaly detection in large databases using behavioral patterning. In 2017 IEEE 33rd International Conference on Data Engineering (ICDE) (pp. 1140-1149). IEEE.
- [11] Casakin, H., Hernández, B., & Ruiz, C. (2015). Place attachment and place identity in Israeli cities: The influence of city size. *Cities*, 42, 224-230. <https://doi.org/10.1016/j.cities.2014.07.007>.
- [12] Gong, S., & Xiang, T. (2011). *Visual analysis of behaviour: from pixels to semantics*. Springer Science & Business Media.
- [13] 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.
- [14] Aldribi, A., Traore, I., & Moa, B. (2018). Data sources and datasets for cloud intrusion detection modeling and evaluation. In *Cloud computing for optimization: foundations, applications, and challenges* (pp. 333-366). Cham: Springer International Publishing.
- [15] Puketza, N. J., Zhang, K., Chung, M., Mukherjee, B., & Olsson, R. A. (1996). A methodology for testing intrusion detection systems. *IEEE Transactions on Software Engineering*, 22(10), 719-729.
- [16] Shiravi, A., Shiravi, H., Tavallae, M., & Ghorbani, A. A. (2012). Toward developing a systematic approach to generate benchmark datasets for intrusion detection. *computers & security*, 31(3), 357-374.
- [17] Joshi, K., Raj, A., & Janakiram, D. (2017, December). Sherlock: Lightweight detection of performance interference in containerized cloud services. In 2017 IEEE 19th International Conference on High Performance Computing and Communications; IEEE 15th International Conference on Smart City; IEEE 3rd International Conference on Data Science and Systems (HPCC/SmartCity/DSS) (pp. 522-530). IEEE.
- [18] Ho, C. Y., Lai, Y. C., Chen, I. W., Wang, F. Y., & Tai, W. H. (2012). Statistical analysis of false positives and false negatives from real traffic with intrusion detection/prevention systems. *IEEE Communications Magazine*, 50(3), 146-154.
- [19] Liao, H.-J., Richard Lin, C.-H., Lin, Y.-C., & Tung, K.-Y. (2013). Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications*, 36(1), 16-24.
- [20] 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.
- [21] Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials*, 18(2), 1153-1176. <https://doi.org/10.1109/COMST.2015.2494502>