Deep Science Publishing
https://doi.org/10.70593/978-93-49307-51-3_2

Chapter 2

Systematic review of Identity and Access Management (IAM)

Esther Chinwe Eze

University of North Texas, United States

2.0 Introduction

The importance of data cannot be overemphasized in today's world. Businesses and companies today are dependent and functional on data and information. It is safe to say that data is the engine of every business regardless of whatever sector they fall in. The identity and access of users within and outside the organization are dependent on data. Data manages the identity and access of users. For securing user data, many companies have implemented the utilization of Identity and access management (IAM). The IAM system allows users to access data based on their credentials (Thakur et al., 2021).

As of now, businesses and corporate companies of various sizes are beginning to utilize and understand the need for implementing an IAM framework. Over the years, business leaders and more especially the IT department in several organizations have been placed under increased regulatory and organizational pressure to make sure that corporate resources are protected by way of controlling access to corporate resources and tracking user privileges (Phil & Gittlen, 2020). As a result of this, corporate businesses can no longer rely on manual and error-prone processes to assign and track user access to resources hence the adoption of IAM. IAM helps automate access control and auditing of corporate resources (Phil & Gittlen, 2020). While the implementation of IAM systems has provided organizations with the appropriate level of security through the use of policies, authentication, validation, and privileges within and outside organizations, the emergence of innovative and sophisticated threats and attacks in recent times has sparked up concerns that question the efficiency of IAM in terms of security. Thus, the risk of attacks posed on IAM is significantly higher than before. This section focuses on different types of overviews of IAM, IAM attacks, the investigation of IAM attack detection techniques, and the revolutionization of IAM with Machine Learning (ML).

2.1 Identity and Access Management (IAM)

A system that defines and manages the roles and privilege access of individual network entities (users and devices) to a range of cloud and on-premises applications is referred to as an Identity and access management system (IAM). Typically, a User can be customers, partners, and employees and devices can be computers, smartphones, routers, servers, controllers, and sensors (Strom, 2021). One of the most crucial jobs for IT security departments is the process of controlling user identity and access to corporate resources like networks, applications, and most especially data. The implementation of IAM poses as the best solution to this challenge. The utilization of IAM systems has helped ensure that the identity and access to the cloud, hybrid environment, and on-premises are at the right level of security (Thakur et al., 2021).

In the IT environment, the concept of an IAM system is about giving access privileges, defining and managing the roles of users controlling the access of those users, and executing the condition where the users are either getting access to or denied access (Banday & Marajo, 2017). The solution IAM provides is making sure a system or network authenticates the identity of a user against a set of specified credentials given. Depending on the system being accessed, these can range from a simple username and password to digital certificates, physical tokens, biometric passwords (such as fingerprints, iris scans, or facial recognition), or a combination of these factors (Multi-factor Authentication) (Logon Box, 2017). Usually, the sensitivity of resources being accessed as well as the impact that can arise if these resources fall into the hands of unauthorised users determines the strength of the authentication. While a stronger authentication (preferably multiple factors) is required to gain privileged access to an administrative account or classified and proprietary data, public or general information might require little or no authentication (LogonBox, 2017).

Today many companies like Microsoft Azure, IBM IAM System, Okta's identity and access management, Oracle Identity Cloud Service, Centrify Identity Service, RSA SecurID Suite (Dell company), etc provide IAM services that offer the facility to on-boarding and off-boarding of users, which results in providing secure access to systems and applications automatically (Thakur et al., 2021). From the study of Thakur et al. (2021), a basic IAM system architecture is shown in Figure 1 below.

2.2 Concept of IAM

The key task of an IAM system is to perform the following: identification, authentication, and authorization (De Groot, 2019). When a user tries to access a system or resource, they are expected to enter or provide their authorized login credentials for identification purposes. The process of authentication allows the user's credentials and identity to be verified and if the authentication process is successful, the authorization process is initiated by IAM. Authorization, on the other hand, is the process where the system decides whether the authenticated user has permission or not to perform whatever action they have requested (Miller, 2020). The meaning behind this is that only the intended user is required to have access to any information or resources like computers, software, hardware, information, IT resources, or a required task to perform.

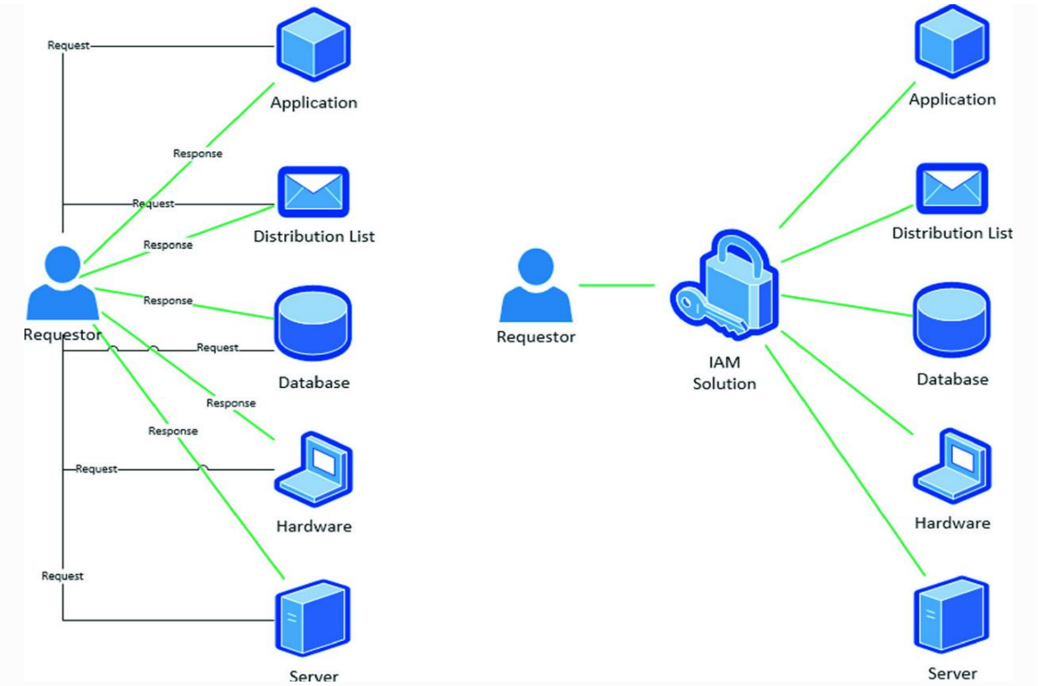

Figure 1: A Basic IAM System Architecture (Thakur et al., 2021).

2.2.1 How IAM Works

Traditionally, a typical identity and access management (IAM) system covers these four basic elements (Strom, 2021):

- An identity repository or directory that contains the personal data of the users which the system uses to define the users.
- A set of tools used for adding, modifying, and deleting that data (related to access lifecycle management)
- A system that controls and enforces user access
- A system for auditing and reporting

User access regulation usually involves authentication methods that verify the identity of a user or device including digital certificates, passwords, smartphone software tokens, and hardware. A more modern approach to regulating user access includes biometric elements (Strom, 2021). As a result of today's complex computing environment, identity and access management systems usually integrate elements of biometrics, artificial intelligence machine learning, and risk-based authentication to meet the heightened security threats (Ganesh The Awesome, 2020).

2.2.2 IAM Tools and Services

Managing an IAM system can be a complex task to handle especially at an enterprise level. However, there are tools and solutions designed to help manage it. The following are a few tools (not limited to these):

- **Single Sign-On (SSO):** Provides a solution that helps reduce the need for users to use multiple credentials. This allows a user to log in once and authentication to other internal applications and systems is guaranteed automatically (Miller, 2020).
- **Multi-Factor Authentication (MFA):** This provides an extra layer of authentication. The MFA authentication process combines something the user knows (like a password) with something the user has (like a security token or OTP) or something that's part of the user's body (like biometrics) (De Groot, 2019). For instance, a user logs in using their primary account, and a unique code is been sent to the user's smartphone or security token for verification of identity (Miller, 2020).
- **Risk-Based Authentication:** A dynamic solution that runs an algorithm to calculate the given risk of a user performing a specific action. If the risk score is too high, the action is blocked and the IT team is notified (Miller, 2020).
- **User Provisioning:** Automated systems that allow you to quickly create new enterprise accounts for users and assign them to roles and groups through a front-end interface (Miller, 2020).

2.2.3 Features of IAM Solutions

Identity and access management systems are considered to be an invaluable tool to many enterprises today. The process of maintaining data security is too complex to be handled manually, therefore access to data must be governed by using flexible and granular control methods. For this reason, enterprises need to seek IAM solutions that offer features designed to focus on today's multidimensional access needs (Identity Management Institute, 2022). Listed below are some key features of an IAM solution.

2.2.3.1 Equipped For Emerging Security Trends

In cybersecurity and IAM, one thing that never stops evolving is trends, especially in the aspect of security. Current trends are predictors of more changes that will happen in the future, from the slow demise of passwords to the implementation of zero-trust security. As users begin to access systems in different new ways and new devices begin to appear on the market, the need for an adaptable and responsive IAM approach will be required for businesses. Thus, in choosing an IAM solution, the thought of flexibility is key. IAM solutions should be able to handle future enterprise access needs and not just the business needs of today. Future access needs should cover user access behaviors and technology with the sensitivity to identify and protect against new threats (Identity Management Institute, 2022).

2.2.3.2 Fast Incident Alerts and Responses

Without proper IAM solutions, a breach activity can go undetected for months in a network. Undetected breaches can result in crippling consequences that could be detrimental to enterprises and so for this reason, the solution would be to implement tools that can detect and prevent the escalation of suspicious behaviors. Once a possible breach activity is detected, an IAM's job is to respond automatically with the right interim defense while an alert is been sent to the security/IT team. When the right IAM solution is deployed it makes monitoring and responding to suspicious behavior easy for the security team to detect thereby reducing the risk of breaches (Identity Management Institute, 2022).

2.2.3.3 Numerous Identity Verification Options

When it comes to identity verification methods, each has drawbacks, and some of these drawbacks are yet to be discovered. A good IAM solution should be able to offer flexible login options that integrate multiple methods of identity verification. One of the most common methods for access control is multi-factor authentication (MFA) which combines other factors like passwords, one-time passwords, biometrics, authenticator apps, and email links. Multiple methods are required when handling sensitive data or in situations where privileged access is necessary and high-risk access is required (Identity Management Institute, 2022).

2.2.3.4 Mobile-Ready Access Control

The "anywhere", "anytime" nature of mobile device usage makes the creation of secure access parameters necessary. With the number of user-owned devices accessing networks, many enterprises are still coming to grips with this. The number of user-owned devices poses significant security concerns (Identity Management Institute, 2022).

2.2.3.5 Compatibility and Integration

When introducing an IAM solution or any new software in an organization there are chances that the new solution or system may create possible conflicts between platforms.

Thus, an IAM technology should be tested to make sure it is compatible across all platforms to prevent possible problems from arising. Alternatively, there may be a need to integrate an existing system to support a newer version of an IAM technology this can be done by upgrading the existing system to support the newer IAM technology.

2.2.3.6 Comprehensive Analytics

Identity analytics plays an integral role in revealing how users access and interact with networks by providing essential information that can help clarify the roles and access policies of users. Analytics can help expose breach activities by revealing a direct correlation between the identity of the user and the security incidents. The information obtained from analytics provides enterprises with information on how to improve the security framework. Analytics also plays an important role in compliance, this explains how important analytics plays in IAM technology (Identity Management Institute, 2022).

2.3 Possible IAM Threats and Attacks

The common forms of attacks on IAM are password attacks. Password-based authentication is one of the most popular forms of user identity verification, as popular as this method is, it tends to have many weaknesses. These weaknesses can result in a corporate and personal data breach if not handled properly.

Listed below are some of the most common password-like threats and attacks on IAM:

2.3.1 Phishing

Phishing is rated as one of the most popular methods of attack because most times it favors the attacker hence why it is commonly used by attackers. According to Verizon's (2017) data breach investigation, 80% of data breaches leveraged stolen or weak credentials/passwords (Biscoe, 2017). With a phishing attack, the attacker only needs to gain access to only a few accounts or probably one of the admin accounts to compromise the entire organization. Even in a well-trained organization, many employees fall prey to phishing attacks (Fisher, 2018).

How Phishing works

First, the Attacker obtains email addresses and designs a generic call-to-action phishing message that seems relevant (like a fake Google login page) to the people it is sent to. The phishing message is distributed, and thereafter the attacker waits to see who enters or login with their credentials. The attacker collects the credentials and uses the stolen credentials to access the data they want or impersonate that identity for a more targeted attack on a high-value employee (Fisher, 2018).

2.3.2 Spear Phishing

A more targeted form of phishing is used to send messages to a targeted list. This type of phishing involves thorough research (using social media or web presence) on the victim and focuses on a small number of victims to escape the automated filter. The message appears to be more personal with a malicious call-to-action, unlike a normal phishing attack (Fisher, 2018).

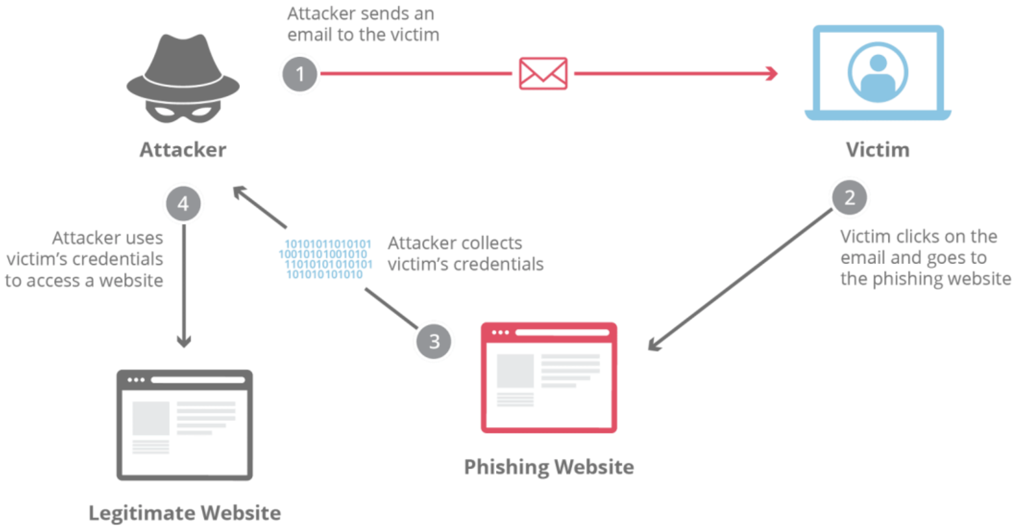
Attacker sends an email to the victim
1
Attacker
Victim
4
2
Attacker uses victim's credentials to access a website
1010101101010110010101001010110101010101011010101010
Attacker collects victim's credentials
Victim clicks on the email and goes to the phishing website
3
Phishing Website
Legitimate Website

Figure 2: Phishing Attack (Cloudflare, 2021).

How Spear Phishing Works

First, targets are carefully picked, and then the attacker crafts a phishing message that appears to be legitimate by pretending to be a colleague and referencing a topical situation. The victim is compelled to enter credentials to view the message. Once the credentials are collected, the attacker uses it to access sensitive data or maybe implement the next phase of the attack (Fisher, 2018).

2.3.3 Brute Force Attack (BFA)

Attackers use brute force attacks to obtain credentials to gain unauthorized access to a system. A brute force attack involves guessing and combining usernames and passwords. BFA is a simple method of attack but highly successful.

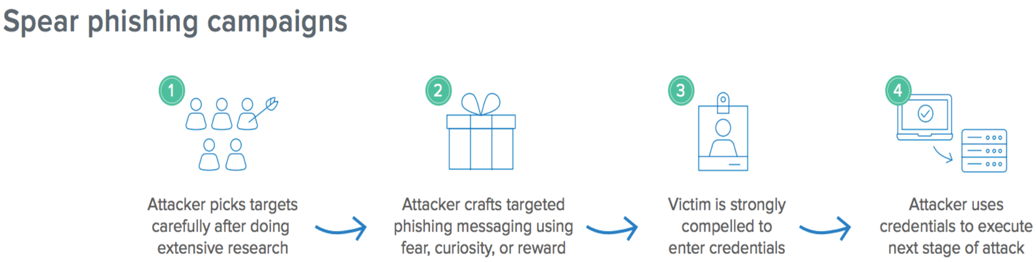
Spear phishing campaigns
1
Attacker picks targets carefully after doing extensive research
2
Attacker crafts targeted phishing messaging using fear, curiosity, or reward
3
Victim is strongly compelled to enter credentials
4
Attacker uses credentials to execute next stage of attack

Figure 3: Spear Phishing Attack (Fisher, 2018).

How Brute Force Attack Works

In the past years, many attackers have performed brute force attacks manually which can be time-consuming, today nearly all BFA are done by bots which makes the attack faster and more effective. In most cases, the attackers usually have a list of commonly used credentials or real credentials of users that have been obtained from a security breach. The bots systematically try the attacker's list of credentials and notify the attacker when access is successful**.**

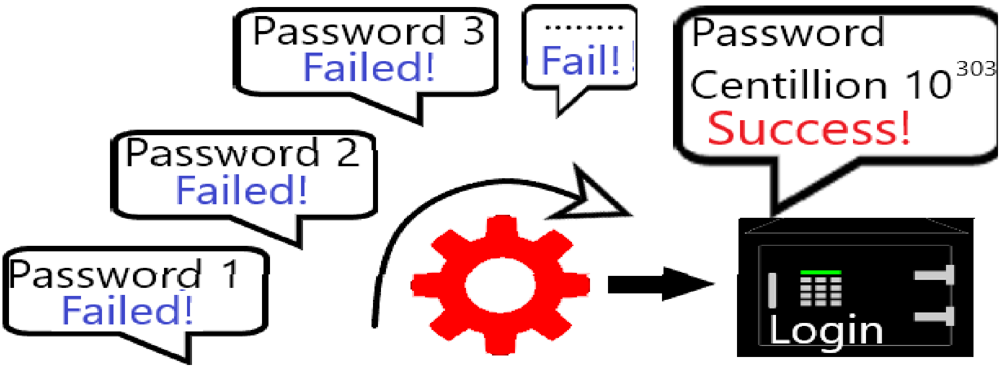

Figure 4: Brute force Attack Illustration (Fajar, 2020).

2.3.4 Credential Stuffing

This is a form of brute force attack that takes advantage of a list of compromised user credentials to gain unauthorized access to a system. This type of attack leverages bots for scaling and automation. There is usually a high chance of breached credentials successfully logging in on another service. Credential stuffing is a rising threat to IAM systems.

How Credential Stuffing Works

Firstly, an attacker sets up a bot that can log in to multiple user accounts in parallel automatically, while using fake IP addresses. The bot then runs an automated process to check if stolen credentials can successfully log in. The bot monitors for successful logins and notifies the attacker if there is a successful login. Once successful, personal information like credit cards or other valuable data from the compromised accounts is obtained. The account information and credentials are retained for future use and the next phase of the attack is executed (Miller, 2020).

2.3.5 Man-in-the-Middle Attack (MITM)

MITM attack is a highly targeted attack and if executed correctly, a man-in-the-middle attack can result in a full take of credentials and data-in-transit (Miller, 2020). MITM attack happens when an attacker intercepts the communication between systems and manipulates data without the knowledge of the provider and relying party. The attacker pretends to act as a provider and the relying party by mimicking the communication between them. MITM attacks are targeted to steal personally identifying information such as credentials, account information, and financial data including credit card numbers and bank details. The stolen information can be used for identity theft, illicit password changes, and unapproved fund transfers (Indu et al., 2018).

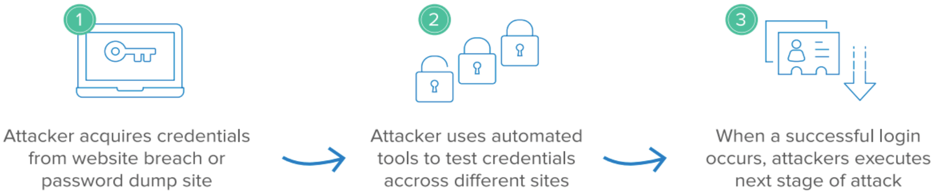
1
Attacker acquires credentials from website breach or password dump site
2
Attacker uses automated tools to test credentials accross different sites
3
When a successful login occurs, attackers executes next stage of attack

Figure 5: Credential Stuffing (Fisher, 2018).

How MITM Attack Works

An attacker intercepts a network connection, often by leveraging tools to mimic a legitimate wifi access point (such as Starbucks Wifi). If data is encrypted, the attacker may attempt to decrypt data by tricking the user into installing a malicious certificate or other technique. If the attack is successful before the initial authentication, the credentials may be stolen as the attacker is monitoring all the user inputs. Alternatively, the attacker steals the session token and can authenticate into the account and execute the next stage of their attack (Miller, 2020).

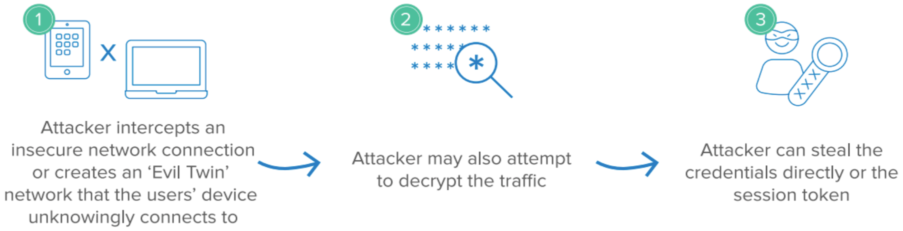
1
Attacker intercepts an insecure network connection or creates an 'Evil Twin' network that the users' device unknowingly connects to
2
Attacker may also attempt to decrypt the traffic
3
Attacker can steal the credentials directly or the session token

Figure 6: Man-in-the-Middle Attack (Fisher, 2018).

2.4 IAM Attack Detection Techniques

In the past years, there has not been much research on IAM attack detection techniques but there have been numerous researches on related attack detection techniques like malware and network attack detection techniques. These techniques are similar to IAM attack detection. Understanding the concept of attack detection is key, therefore a brief introduction of attack detection will be discussed before exploring the various techniques available. An attack(s) could be seen as an attempt to bypass the security policies that have been put in place to secure a system. Bypassing security policies gives attackers easier access to breach sensitive information from a system and some cases destroy the system (Wu et al., 2020). In the attempt to deal with cybersecurity attack threads, many researchers have proposed many solutions and among these solutions, attack detection is proven to be one of the most effective ways of combating attacks (Xu et al., 2019). Attack detection offers a complete and dynamic security mechanism for monitoring, preventing, and resisting attacks. An attack detection mechanism would specifically collect information by monitoring networks, behaviors and patterns, system status, and the usage of systems which would help in detecting unauthorized usage of system users and attacks of external attackers on the system (Wu et al., 2020).

2.4.1 Machine Learning

IAM systems are becoming increasingly popular as they are essential to safeguarding organizations. The downside of IAM systems is that they are highly vulnerable to many cyber threats. When addressing IAM attacks, one significant detection technique that stands out in this area is Machine Learning (ML)(Aboukadri et al., 2024). For context, Machine learning (which is a subset of artificial intelligence) techniques use algorithms to identify and address security problems (Sampath, 2022). Theoretically, machine learning excels at being able to detect and react to intrusions, particularly those involving IAM. Its detection techniques involve analyzing various patterns in data that seem hidden from traditional techniques, checking for anomalies, and responding to attacks in real time (George, 2023). Machine learning makes use of different techniques like supervised and unsupervised learning, however, for this research, only the supervised learning technique will be applied.

2.4.1.1 Supervised Learning Technique

Supervised learning is a machine learning technique where the models learn based on labeled data. It uses an algorithm trained on input-output pairs to predict and classify the labeled data as well as their accuracy and performance measures (Mohr & van Rijn, 2022). It is commonly used for classification (a way to categorize data into certain labels) and regression (continuous value prediction). In supervised learning, once a satisfactory performance is reached the learning process ends (Mohr & van Rijn, 2022). The research will use four major supervised learning classifiers to assist IAM in detecting attacks. The classifiers are Random Forest (RF), Naive Bayes (NB), K-nearest-neighbour (K-NN), and Support Vector Machine (SVM).