ISSN: 2754-6659

Journal of Artificial Intelligence & Cloud Computing

SCIENTIFIC Research and Community

Research Article

Open Access

Revolutionizing Role-Based Access Control: The Impact of AI and Machine Learning in Identity and Access Management

Shanmugavelan Ramakrishnan

USA

ABSTRACT

Identity and access management is the bedrock of cybersecurity. Identity access to digital resources is governed by its techniques, procedures, and rules, which also define the breadth of identity permission over those resources. Some new cyberattack or data breach pops up in the news every week. Many data breaches occur due to inadequate security measures, software flaws, human mistake, malevolent insiders, or the abuse of access and privileges. An improved access control system is possible with the use of AI methods. In order for organisations to better handle authentication and access control in order to reduce cyber risks and other IAM difficulties, studies into artificial intelligence in IAM are necessary. With an eye towards AI's potential uses in identity and access management - more especially in the areas of privilege monitoring, administration, and control - this research investigates the nature of the connection between AMIS and AI. To better understand how AI works in minimising recognised IAM issues, this study aimed to present evidence from the relevant literature. This study's results show how AI reinforces identity and access management, which helps with automating procedures, keeping up with technology advances, and reducing the prevalence of cyber threats. One way to accomplish this is by using a binary classification system for security access control, which takes the PDP problem and turns it into a yes/no question. In order to create a distributed, effective, and accurate policy decision point (PDP), a vector decision classifier is also built using the supervised machine learning technique. Kaggle-Amazon access control policy dataset evaluated performance by comparing the proposed mechanism to previous research standards in terms of performance, duration, and flexibility. Given that the PDP is not in direct contact with the PAP, the proposed approach accomplishes a high level of secrecy in relation to access control requirements. In conclusion, PDP-based ML can manage massive access requests, execute many major policies simultaneously, and have a 95% accuracy rate, all without policy conflicts, with a response time of about 0.15 s. The security of access control can be enhanced by making it more responsive, flexible, dynamic, and dispersed.

***Corresponding author**

Shanmugavelan Ramakrishnan, USA.

Received: July 12, 2023; **Accepted:** July 18, 2023; **Published:** July 25, 2023

Keywords: Access Control, Artificial Intelligence, Access Management, Identity Management, AI Techniques

Introduction

In the extremely interconnected and globalised work environment of today, individuals and organisations engage in continual interaction with one another. Internet connectivity has become an integral aspect of everyday life, as it connects hundreds of millions of systems all over the world. These systems are powered by a wide range of hardware and software technologies, and they are used to deliver communication and commercial services.

In spite of the fact that organisations are growing more productive and efficient, they run the danger of becoming vulnerabilities to data breaches and other forms of cyber attacks. It can be a difficult issue for many organisations to determine which personnel should be provided access to a specific piece of information. If these organisations choose to ignore this information, it could leave their systems open to attack [1].

Access management, which is also known as access control, is the process of verifying, authorising, and holding accountable the identification of an individual in the event that they are permitted access to resources [2].

The access control system is responsible for determining whether or not the appropriate individual is gaining access to the resources when defined access management is carried out. Access control, which includes the four basic processes of identity, authentication, authorization, and accountability, is the cornerstone of computer security.

It is also the foundation of computer security. According to Damon, the procedure comprises being responsible for the enforcement of access control regulations in order to guarantee a timely and efficient response to any requests for accessing corporate resources [3]. It is the primary purpose of access control to enhance the secrecy and security of information technology assets, also known as the enhancement of the resource's confidentiality.

An effective access control system is critical for preventing unauthorised users from gaining access to the network. Integrated Access Management (IAM) guarantees that end users have a positive user experience that is in line with the suitable security policies. For the purpose of protecting end-users from identity theft and privacy issues, policymakers have enacted a number of legislation that require businesses to install identity and access management (IAM) systems. In addition to these restrictions, the idea of justice is also important.

Frameworks for identity and access management (IAM) are built on three fundamental categories: identity, access, and directory

services. Research has proven that the numerous vendor technology solutions used to support identity and access management (IAM) demonstrate the significance of IAM [3]. Figure 1 shows the four main types of components that make up identity and access management (IAM): authentication, authorization, user management, and a single repository for users.

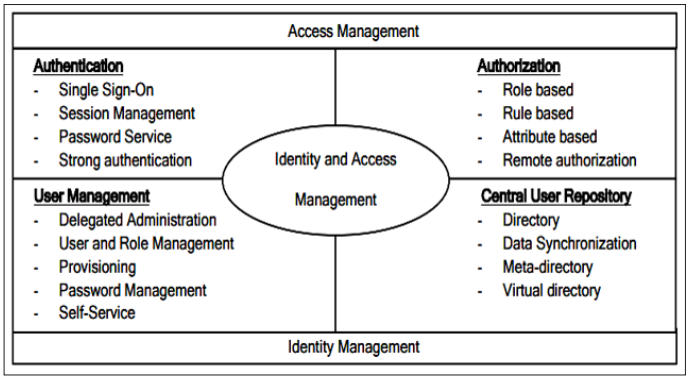

Figure 1: Identity and Access Management Components

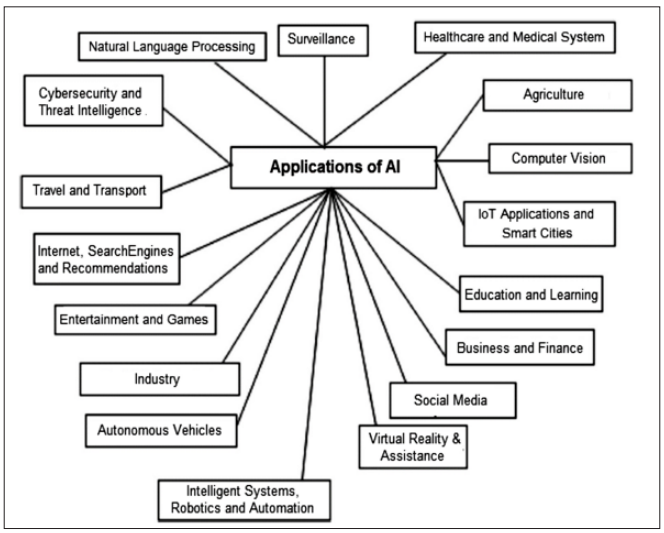

Figure 2: Artificial Intelligence (AI) has a Number of Possible Practical Uses

In recent years, artificial intelligence (AI) methods have been effectively applied to solve several challenges across various application fields. Healthcare, cybersecurity, business, social media, VR/AR, robots, and countless more areas are just a few of the many modern-day application domains. Figure 2 presents a list of prospective application areas for artificial intelligence in the real world.

Literature Review

According to Sudarsan, an essential component of secure digital systems is the administration of digital identities, as well as authentication and authorization technologies [4]. Access management makes use of identification information to validate the identity of an entity, such as an individual or a device, and then authorises permission to access the resource that has been sought [3]. "True" digital transformation, which is the safe, adaptive, and flexible information technology infrastructure that every industry and higher education institution strives to attain, is built on Identity Management, which serves as the cornerstone there.

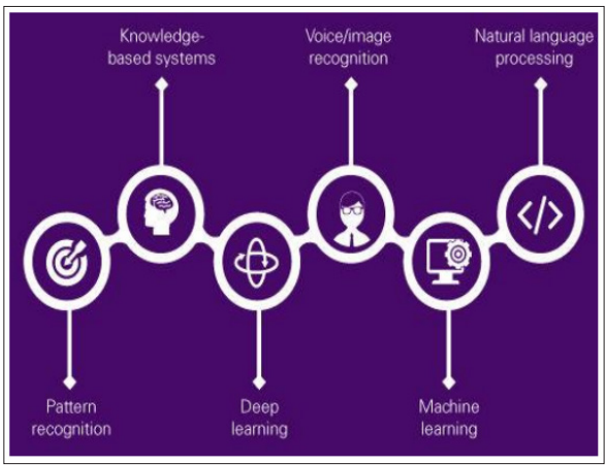

Figure 3: AI Capabilities

Existing approaches to identity management face considerable challenges in terms of privacy and interoperability as a result of the increasing number of consumer and employee accounts and the requirement to manage them in a secure manner. The fact that users are frequently provided access to resources depending on their position within the organisation is one of the challenges that IAM faces. On the other hand, employees are rarely fit to a particular task they are assigned. According to Naik and Jenkins's research from 2020, a big security concern for many organisations operating in the digital environment is finding a way to manage employee access to sensitive apps and data [5].

IAM methods that are more intelligent are associated with a lower level of security risk, according to a number of studies. Experts are of the opinion that artificial intelligence has the potential to be a strategically important potential driver of the continuous technological improvement. This is because it can serve as a tool for intelligence augmentation (IA), rather than a substitute for human thought and reasoning processes [6].

Umurerwa and Lesjak report that AI technologies are frequently used by organisations due to the opportunities they present [7]. Reports indicate that despite the growing interest in artificial intelligence (AI), a significant percentage of organisations are still in the period of data collection, and only a small number of organisations have really implemented AI. Additionally, the paper emphasises continuing research into how artificial intelligence (AI) might be incorporated into business models, and it emphasises that organisations may face dangers if they choose to ignore AI.

Additionally, the paper emphasises continuing research into how artificial intelligence (AI) might be incorporated into business models, and it emphasises that organisations may face dangers if they choose to ignore AI. Furthermore, naively accepting artificial intelligence might lead to a variety of issues that are associated with AI [8]. Artificial intelligence technologies can be difficult to comprehend and to comprehend. In order to fully embrace artificial intelligence as a technological innovation, a constant digital transformation will be required. As a consequence of this, the effectiveness of the influencing elements can be improved by boosting digital organisational preparedness through research (such as this thesis) and by remaining current with technology advancements.

Building models through the use of algorithms is the focus of machine learning, a subfield of artificial intelligence that combines computer science and mathematics. Models of machine learning are able to generate predictions or decisions that are equivalent to those made by the human brain when they are trained using data. Additionally, the accuracy of the output will rise in proportion to the number of learnings that are performed [9].

Artificial intelligence (AI) encompasses a wide range of disciplines, some of which include robotics, computer vision, expert systems, genetic algorithms, reinforcement learning, neural networks, fuzzy logic, planning and scheduling, and machine learning (ML). Machine learning is the most widely used subfield of cybersecurity, according to Versola [10].

Supervised, unsupervised, semi supervised, and reinforcement learning are only a few of the many varieties of machine learning. Instead of being explicitly coded, these machine learning types use statistical approaches to let computers "learn" from data. When aimed at a particular task rather than a general objective, it is capable of producing the best results. In the case of machine learning, for instance, its applications include the detection of intrusions or viruses, as well as the identification of users based on biometrics [11].

Mathematical methods that enable the integration of sensory chains, grouping, intensification, reasoning coding, and decision studying make up machine learning. Classification algorithms, which seek out harmful code and software, anomaly detection algorithms, which look for unusual or harmful traffic, and correlation algorithms, which connect signals from various systems, are all part of its utility [12].

Over the course of time, the system continues to develop in terms of its speed and efficiency in identifying and responding to unforeseen risks [13]. According to the Ellen MacArthur Foundation, artificial intelligence has the ability to handle complexities and improve understanding of huge quantities of information. It may also be considered as a supplement to human capabilities that promotes more efficient learning from feedback.

A number of fields in science and technology have made use of AI-based methods because of their capacity to examine massive amounts of data. Their applications have been found in numerous security-related sectors, including encryption, trust, attack detection, privacy, and Zhang, Guo, Bosri [14-16].

Methodology

The researchers set out to explore how AI and ML are changing the face of industry-specific identity and access management systems and built-in AI solutions for these systems. In this study, a qualitative research method and a research design that was simplified from a meta-analysis approach were utilised in order to present a concise and detailed overview of findings from qualitative studies that were conducted in the same research area, to develop an in-depth analysis, to gather comprehensive data through a variety of data collection methods, and to answer the research questions that were posed for this study.

Data Analysis

During the course of the literature review, the researcher concentrated on works that had been subjected to peer review and that set forth the components of how artificial intelligence and machine learning revolutionise identity and access management systems. It was decided to establish a fundamental topic description, which had headings that described the major discoveries that were made. In the analysis of the data, the concepts that surfaced as a result of the study questions are taken into consideration, along with the ways in which those concepts are connected to the research and theoretical frameworks. A descriptive interpretative technique to analysis is utilised in a qualitative meta-analysis, and the following steps are included in this approach:

- There are domains that are assigned to the data that was collected.
- There is a definition of meaning units.
- Comparison of meaning units and clarification of the core of meaning units that are similar are the processes that result in the generation of concepts.
- These notions can be further categorised or classified based on comparisons and differences, which are defined by the meaning units that are contained within the data.
- There is a summary of the most important findings, which frequently makes use of visual aids or storytelling.

Also, to make sure it's legit, the study uses loads of safety features including reliability tests (such external validation and cross-validation). This aspect of qualitative analysis furnishes in-depth details regarding contextual circumstances or sheds light on the causes of events that are usually overlooked by quantitative research. Results are utilised to understand the pros and cons of the research participants' (here, AI-based solution-using companies) AI feature integration into their IAM system.

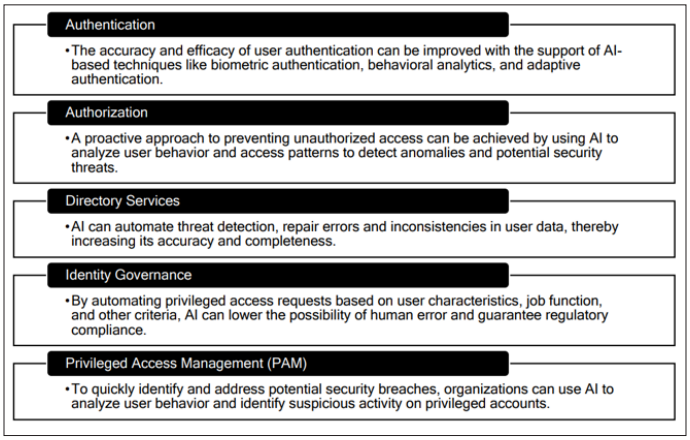

Figure 4: AI Integrated into IAM Components

Proposed Model

In this section, we outline the proposed policy decision point (PDP) for ML-based access control and compare it to previous approaches. We also highlight the differences between the two. More dynamic, distributed, effective, and accurate access decisions are also going to be the focus of this investigation. In addition, we will address the question.

Comparison of Traditional PDP in Access Control vs Proposed PDP-Based ML Technology

The following part will go over the function of the policy decision point (PDP) in conventional access control. Furthermore, we will detail the ways in which our study aims to address the shortcomings of traditional methods by enhancing their adaptability, dynamism, and safety. Decisions on access at policy decision points (PDP) typically include performing logical operations on the request until it is confirmed to comply with the given control policies, as shown in Figure 5. Below is an explanation of the technique.

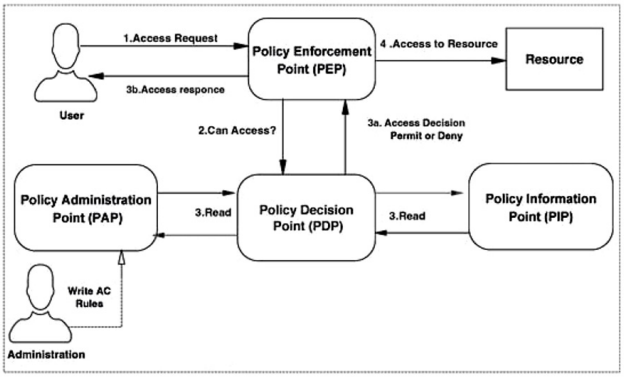

Figure 5: Conventional Approach to Access Control using Policy Decision Points (PDPs)

- The necessary resource can only be obtained if the Policy Decision Point (PDP) receives a request from the Policy Enforcement Point (PEP).
- The PDP is tasked with handling access requests after they are received from the PEP, namely by resolving their properties. At the policy administration point (PAP), it is critical that this match the policy associated with the request.
- The Policy Administration Point (PAP) is responsible for providing a final response, which involves inquiring about the relevant policy set from among all policy sets and then returning the combination policies to the policy determination point. After that, the user is granted direct access to the resource that was specified, provided that the final response is permitted. On the other hand, if the outcome is rejected, the user will not be granted access to the information that was specified.

The proposed framework is an adaptive mapping strategy that includes multiple permission determination architectures to improve the performance of access control decision-making in a BYOD environment. Here are the main aspects that outline the basis of the proposed model.

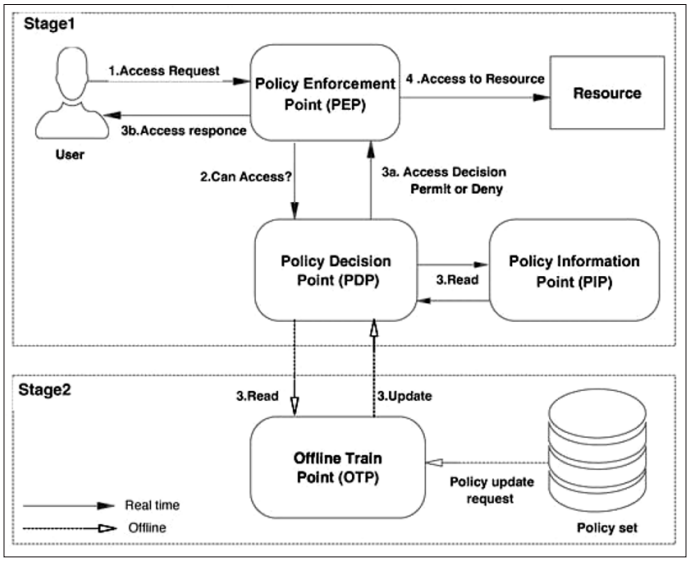

Figure 6: Flexible and Decentralised Policy Decision Point (PDP) Built on Machine Learning

Results and Discussion

Utilising the following performance evaluation criteria, we established the extent to which the suggested solution would ameliorate the policy decision point. First, the confusion matrix in Table 1 was built using the decision-making findings. The defined number of samples that were correctly refused access (DAR), the quantity of samples that were incorrectly given access (DRA), and the number of samples that were correctly rejected access (DRR) make up the total number of samples that were accessible.

Table 1: Decision Point Confusion Matrix for Policies Final Product Data from PDP

	Predicted Results	
Real Results	**Allowed Access**	**Refused Access**
Allowed Access	D_{AA}	D_{AR}
Refused Access	D_{RA}	D_{RR}

Accuracy Metric

Utilised to assess the efficacy of access control strategy policy decision points across all cases. Divide the sum of all estimates by the sum of all correct ones to obtain the percentage. To determine Accuracy (CM), apply the following formula: Incorporating the confusion matrix:

$$Accuracy = \frac{D_{AA} + D_{RR}}{D_{AA} + D_{AR} + D_{RA} + D_{RR}}$$

Precision Metric

Determines the value by dividing the total of all positive samples by the total of all adequately identified positive samples, irrespective of their accuracy. The predicted number of acceptable samples divided by the total number of accurate samples establishes this. The formula provided was used to determine this:

$$Precision = \frac{D_{AA}}{D_{AA} + D_{RA}}$$

Recall Metric

Finds its value by dividing the proportion of positive samples by the number of valid samples that are accurately categorised as actual. Using the Recall measure, we may assess the model's sample-identification capabilities. The true number of samples is raised by recall.

$$Recall = \frac{D_{AA}}{D_{AA} + D_{AR}}$$

F1-score Metric

The F1-score is calculated by averaging the two most important metrics, Precision and Recall, in a harmonic fashion. The F1-score is a measure of data imbalance. The geometric mean is different from the harmonic mean.

$$F1 - score = \frac{Precison * Recall}{Precision + Recall}$$

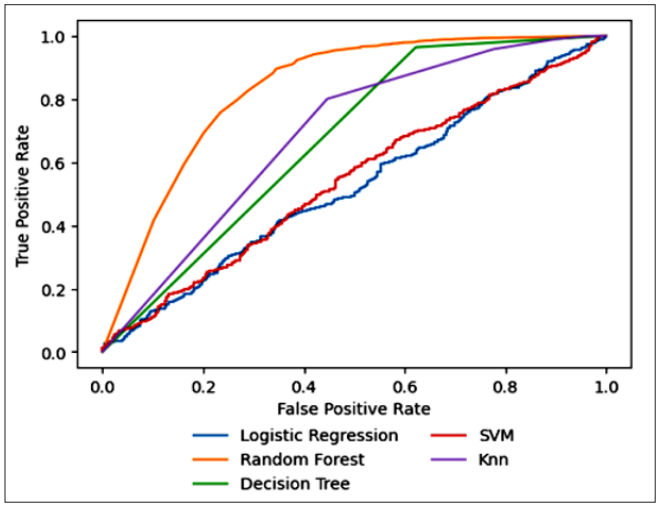

Figure 7: Examining the Outcomes of the ROC Curve Prior to Data Balancing for Every Algorithm

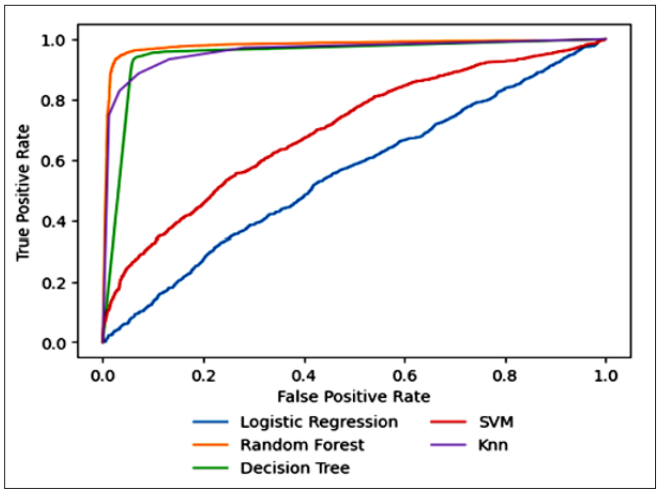

Figure 8: Examining the Outcomes of the ROC Curve after Data Balancing for Each Algorithm

Figures 7 and 8 display the experimental results, which show that the ROC curve stays the same regardless of how the sample group is distributed. This meant that the algorithms' ROC performance was low before using the data balancing technique.

Table 2: AUC Values of Different Algorithms

Methods	RF	KNN	DT	LR	SVM
AUC-Unbalanced Dataset	0.83	0.69	0.67	0.52	0.54
AUC-Balanced Dataset	0.98	0.96	0.95	0.55	0.70

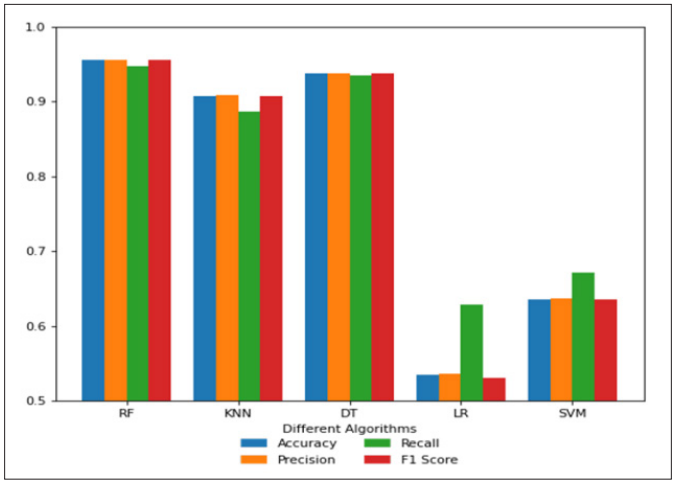

Figure 9: After Processing the Data, Compare the Approaches' Performance Metrics (Balancing Method)

As shown in Figure 9, the two algorithms with the lowest performance are logistic regression (LR) and support vector machines (SVM). Results show that RF, decision trees, and k-nearest neighbours (KNN) are all very similar.

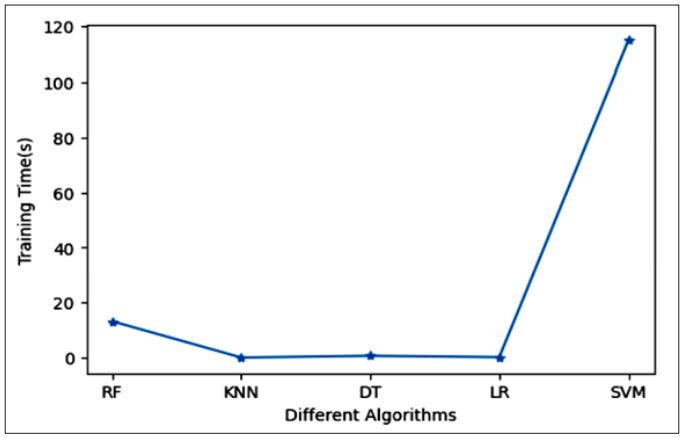

Figure 10: After Data Balance, Compare the Training Time Outcomes of Various Approaches

A key component of access control security is the ability to respond quickly to requests for access and to update the system's access control policy dynamically; this keeps the system from becoming a bottleneck, which is why time is so important. So, as shown in Figure 10, we looked at how long it took to train the model and how long it took for a policy decision point (PDP) to update its decision on whether to enable or disable access. When compared to the other methods, SVM training took the longest, while LR, KNN, and DT all took around the same.

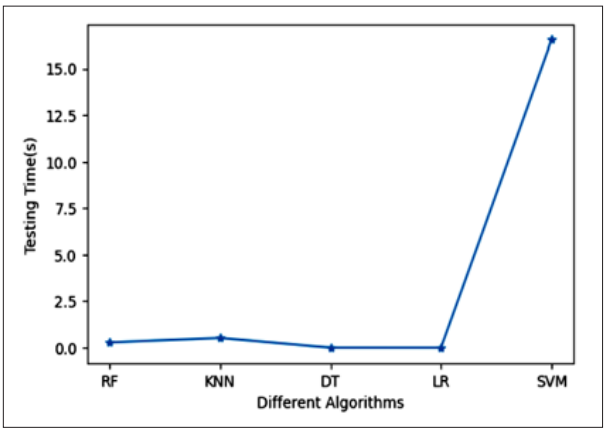

Figure 11: After Data Balance, Compare the Outcomes of Various Testing Methods

In addition, as shown in Figure 11, data balancing approaches were used to analyse the testing time outcomes of different methodologies, with an emphasis on security access decision-making. A number of methodologies were taken into account for the analysis, including RF, KNN, DT, LR, and SVM (support vector machine). The goal of the study was to compare the approaches' efficiency by looking at how long it took to test them following data balancing.

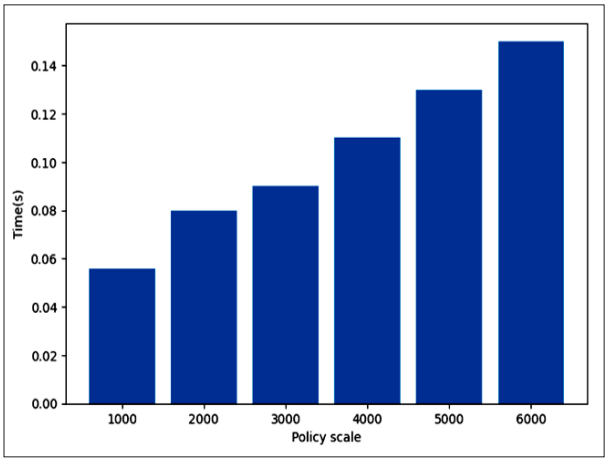

Figure 12: Determine the Impact of Various Policy Scales on Time

Experiment results showing the flexibility of ML-based PDP with respect to policy size over time are shown in Figure 12. A measure of adaptability was the number of seconds it took for access control to react to complicated policies. It was decided to send out random requests with policy sizes ranging from 1000 to 6000 for the experiment. At a policy scale of 6000, the time cost was around 0.15 s. The model seemed to have a favourable reaction to the policy's magnitude.

Conclusion

Access control decision-making methods were examined and assessed in this study. At the usual policy decision point for access control, predefined rules are employed to make access decisions. Consequently, there are a number of issues with decision-making and access control, such as competing policies, insufficient entities, subpar PDP performance, response times to access requests, and financial concerns. However, the proposed method enhanced the access control policy decision point (PDP) based on machine learning by translating PDP concerns into a straightforward yes/no request classification. To begin, there was a balanced dataset. Attributes and characteristics have their dimensions subsequently shrunk. As part of the enhanced PDP-based ML method, two models were developed: one for testing training models and another for decision-making.

To improve upon the current access control, the proposed solution proposed a lightweight, distributed, and dynamic policy decision point (PDP) in place of an access decision based on predetermined rules. In spite of the scenario's complexity, the amount of entities, and the number of concurrent big requests, it handled them all with ease. The policies were also made more secretive and private without requiring direct contact or connection to the PAP policy, which provided security and privacy for access control.

References

1. Mohammed IA (2021) Identity Management Capability Powered by Artificial Intelligence to Transform the Way User Access Privileges Are Managed, Monitored and Controlled. International Journal of Creative Research Thoughts (IJCRT) 4719-4723.
2. Schrimpf A, Drechsler A, Dagianis K (2021) Assessing Identity and Access Management Process Maturity: First Insights from the German Financial Sector. Information Systems Management 94-115.
3. Damon F (2019) A framework for identity and access assurance, Doctoral Thesis. Johannesburg: Creative Commons.
4. Sudarsan SV, Schelén O, Bodin U (2021) Survey on Delegated and Self-Contained Authorization Techniques in CPS and IoT. IEEE Access 9: 98169-98184.
5. Naik N, Jenkins P (2020) uPort Open-Source Identity Management System: An Assessment of Self-Sovereign Identity and User-Centric Data Platform Built on Blockchain. 2020 IEEE International Symposium on Systems Engineering (ISSE) 1-7.
6. Faruk MJ, Shahriar H, Valero M, Barsha FL, Sobhan S, et al. (2021) Malware Detection and Prevention using Artificial Intelligence Techniques. 2021 IEEE International Conference on Big Data (Big Data) 5369-5377.
7. Umurerwa J, Lesjak M (2021) AI implementation and usage: A qualitative study of managerial challenges in implementation and use of AI solutions from the researchers' perspective. Umea University, Faculty of Social Sciences, Department of Informatics 1-34.
8. Dignum V (2021) The role and challenges of education for responsible AI. London Review of Education 19: 1-11.
9. Zhu G, Al-Qaraghuli Y (2022) AI-Assisted Authentication: State of the Art, Taxonomy and Future Roadmap. Cryptography and Security 1-25.
10. Versola L (2021) Machine Learning in Identity and Access Management. Zero Trust Edge https://www.zerotrustedge.com/blog/machinelearning-in-identity-and-access-management/.
11. Dasgupta D, Akhtar Z, Sen S (2020) Machine learning in cybersecurity: a comprehensive survey. Journal of DefenseModeling and Simulation: Applications, Methodology, Technology 1-50.
12. Korolov M (2022) Top Three Use Cases for AI in Cybersecurity. Data Center Knowledge https://www.datacenterknowledge.com/security/top-three-use-cases-ai-cybersecurity.
13. ServReality (2022) Artificial intelligence in cybersecurity: pros and cons. ServReality: https://servreality.com/blog/artificialintelligence-in-cybersecurity-pros-and-cons/.

14. Bosri R, Rahman MS, Bhuiyan MZ, Omar AA (2021) Integrating Blockchain with Artificial Intelligence for Privacy-Preserving Recommender Systems. IEEE Transactions on Network Science and Engineering 8: 1009-1018.
15. Guo K, Ren S, Bhuiyan MZ, Li T, Liu D, et al. (2020) MDMaaS: Medical-Assisted Diagnosis Model as a Service with Artificial Intelligence and Trust. IEEE Transactions on Industrial Informatics 16: 2102-2114.
16. Zhang G, Li J, Bamisile O, Cai D (2021) Spatio-Temporal Correlation-Based False Data Injection Attack Detection Using Deep Convolutional Neural Network. IEEE Transactions on Smart Grid 1-1.