Journal of Theoretical and Applied Information Technology
15th February 2024. Vol.102. No 3
© Little Lion Scientific

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DYNAMIC ACCESS CONTROL AT THE NETWORK EDGE USING AN ADAPTIVE RISK-BASED ACCESS CONTROL SYSTEM (ad-RACs)

MUHAMMAD BELLO ALIYU[1], DR. MUHAMMAD GARBA[2], DR. DANLAMI GABI[3], DR. HASSAN U. SURU[4], DR. MUSA S. ARGUNGU[5]

[1,2,3,4]Department of Computer Science, Kebbi State University of Science and Technology, Aliero, Nigeria

E-mail: [1]mbacaspet@gmail.com, [2]garbamga@gmail.com, [3]gabsonley@gmail.com, [4]suruhassan@yahoo.com, [5]sm279arg@gmail.com

ABSTRACT

The widespread adoption of edge computing models owes to their cost-effectiveness and performance advantages for both users and service providers. However, the expanding user base and application scope raise security concerns, including potential malicious attacks due to unrestricted system resource access. Hence, this study focuses on implementing an Adaptive Risk-based Access Control System (ad-RACs) at the network edge. The ad-RACs utilizes four key inputs—user context, resource sensitivity, action severity, and risk history—to enable the CatBoost risk estimation module to evaluate security risks associated with access requests. Upon meeting the acceptable risk threshold, the Chinese wall access control policy determines access decisions. This model adapts to user behavior and patterns, updating risk history to dynamically adjust access requests. Evaluation results showed that the ad-RACs exhibited satisfactory recall and F1 score values of 100% and 98%, respectively, and a precision value of 95%, outperforming the existing system's recall of 98%, F1 score of 96%, and precision of 97%. Conclusively, the ad-RACs excelled in recall and F1 score values compared to the existing system, indicating its potential to enhance access control. Its adoption is recommended for governmental and private organizations seeking to bolster user access to sensitive resources.

Keywords: *Access Control, Adaptive Access, Adaptive Security, Network Edge, Risk-based Access*

1. INTRODUCTION

Beyond the incessant developments of ubiquitous computing, wireless sensor networks and communication, interconnected devices are increasing in number with Edge Computing playing an important role in several sectors; enabling uniquely identifiable heterogeneous physical devices to communicate over the internet [1]. On this note, registration and authorization infrastructures are critical, to register and analyze the credentials of the various entities with the intention of authorizing their requests to carry out certain actions. With no authorization infrastructure, anyone can misappropriate the infrastructure's resources; pretend to be an admin and control the infrastructure's services; attackers can gain access to any resources. As a result of the benefit inherent in edge paradigms, the deployment of an authorization infrastructure in every trust domain is critical. Thus, allows trust domain owners to distribute and implement their security policies [2]. In principle, such infrastructures can process the credentials of any entity based on an existing trust relationship. While also, considering various contextual information, such as the geographical location, resource ownership, and the description of the authentication policies. Thus, access control management system became critical. Most operations in edge networks include access request to resources, and data transmission and processing. Without an authorization mechanism, there will be unrestricted access to system resources, and hence malicious attacks can be perpetrated. To ensure authorization, it is critical for security access policy to be enforced in each trust domain and the level of resource allocation defined. Such that, for resources to be shared between any two entities, credentials and access policies are essentially needed. Sustaining and enforcing access policy may be resource consuming and thus, a well-organized and protected mechanism to sustain and enforce this policy is required. Such that, based on a predefined authorization policy, edge devices can grant access to resources.

Traditional Access control guarantees that access requests are authorized consistent with predefined rules. These rules make up an authorization policy and the way of defining and administering them institutes an access control model [3]. Various traditional access control models have been developed over time to secure access to shared resources. Such as Discretionary Access Control (DAC), Mandatory Access Control (MAC), Role-Based Access Control (RBAC), Attribute Based Access Control (ABAC) and Risk-Based Access Control. But most of the edge networks have an architecture that evolves dynamically in terms of connected users, devices, and services. Employing traditional access control technique to cater for the growing network requirements is challenging. Such as, too strict to handle unique scenarios where the policy needs to be overruled to preserve the system; lacks the requirements of a collaborative environments with dynamically secured information and permission sharing; and, lacks the flexible required to handle changing user behavior. Subsequently, attacks such as illegal leakage of information, denial of service and information altering are still dangerous.

New systems categorized by decentralization, automated reconfiguration of users and resources dynamically, poses new challenges to the traditional access control technique. In addressing these issues, previous studies had used varieties of techniques such as a blockchain- and token-based adaptive dynamic access control approach, and an adaptive risk-based access control model. The outcome of the blockchain- and token-based adaptive dynamic access control approach focused on resolving the single-point failure risk of permission control and the inability to balance dynamic fine-grained permission adjustment and real-time response. However, the computational complexity of the access control process is an inherent limitation. Contrary to this technique, the adaptive risk-based access control model focused on need to improve flexibility and scalability of the access control process. However, lack of proof of concept is an inherent limitation. Since, the current access control technique does not adequately provide efficient performance; therefore, this study aims at developing an Adaptive Risk-based Access Control System (ad-RACs) to improve dynamic access to edge resources.

1.1 Problem Statement

The increasing prevalence of edge computing models has indeed brought about cost-effectiveness and heightened performance advantages. However, the concurrent growth in user numbers, resource accessibility, and supported applications within edge networks accentuates a pressing concern – the security challenges intrinsic to these systems. The absence of a robust authorization mechanism in this context leaves system resources vulnerable to unrestricted access, creating a breeding ground for potential malicious attacks. Ensuring proper authorization becomes imperative, emphasizing the critical need for the enforcement of security access policies within individual trust domains, along with the definition of resource allocation levels. Traditional access control models, despite their efficacy in securing shared resources, face challenges in the dynamic and evolving landscape of edge network architectures. These challenges include inflexibility in handling unique scenarios, a lack of adaptability to collaborative environments with dynamic security needs, and an inability to accommodate changing user behaviours. This leaves systems susceptible to threats such as illegal information leakage, denial of service, and information tampering. Thus, traditional access control techniques prove insufficient. Past attempts at alternative approaches, such as blockchain- and token-based adaptive dynamic access control, encountered computational complexities. Given the evident inefficiencies in current access control techniques, this study seeks to address these gaps by developing an Adaptive Risk-based Access Control System (ad-RACs) aimed at enhancing dynamic access to edge resources.

1.2 Rationale for the Study

Adaptive access control is an instance of context-aware access control that seeks to equalize the trust level against risk level. This would enable a better tackling of access-related risks while refining user experience. This research offers a fine-grained access control to edge data centers by implementing an Adaptive Risk-Based Access Control System (ad-RACs). This takes into account real-time data and information when access is requested and gives dynamic response.

1.3 Contribution to Knowledge

This study makes a significant contribution to the field by introducing the Adaptive Risk-based Access Control System (ad-RACs) to address the shortcomings of traditional access control models in the dynamic environment of edge computing. The ad-RACs system is designed to enhance dynamic access to edge resources by incorporating real-time factors such as user context, resource sensitivity, action severity, and risk history. Leveraging machine learning techniques, particularly the CatBoost algorithm, the system estimates access risk and dynamically determines whether to grant or deny access. The development of this adaptive and risk-aware access control system represents a novel approach to securing interconnected devices in edge networks.

1.4 Practical Implication

The practical implications of the ad-RACs system are significant for industries and sectors relying on edge computing. With the proliferation of interconnected devices, ensuring secure access to edge resources is paramount. The ad-RACs system offers a practical solution by providing a more flexible and scalable access control mechanism. Its ability to adapt to the dynamic nature of edge networks, consider real-time contextual factors, and employ machine learning for risk assessment makes it well-suited for environments with evolving user behavior and resource configurations. The system's evaluation against existing access control solutions demonstrates promising results in terms of precision, recall, and F1-score, indicating its potential effectiveness in real-world applications. Implementing ad-RACs could lead to improved security, reduced vulnerability to attacks, and enhanced overall access control management in edge computing environments.

2. LITERATURE REVIEW

2.1 Edge Computing

Edge computing is an emerging paradigm in computing; bringing cloud computing services closer to the users and speeding up the service process and request time. Edge computing have brought about improvement in many other technological applications. Making it possible to provide high quantity of different information and services [4]. Edge Computing aims at providing a computing platform with cloud computing capacities at the network edge. The advantages of positioning cloud services at the network edge include reduced latency, increased bandwidth, reach to radio network information and location awareness. This makes it conceivable to improve existing infrastructure or implement new services. Additionally, the positioning of services is open to 3rd party service providers. Other applications area includes augmented reality, smart video acceleration, interconnected cars, and Internet of Things gateways, and so on [2]. The positioning of virtualization servers at several locations at the network edge is essential in the implementation of edge computing environment. Some deployment locations include LTE/5G base stations (eNodeB), 3G Radio Network Controllers (RNC), or multi-Radio Access Technology cell aggregation sites. This virtualization infrastructure should host both edge computing services and other related services [5].

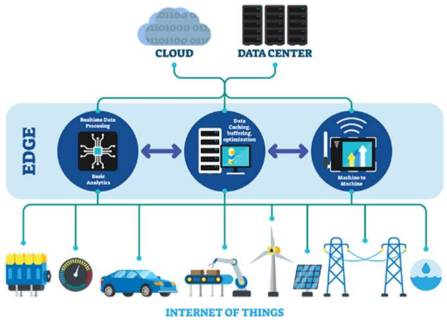


Figure 1: Real-Life Use Cases for Edge Computing [6]

2.2 Access Control Mechanism

It is noteworthy that access control is applied to constraint activities carried out by authorized users and prevents any activity that could result to a security violation. It should accomplish the security goals of confidentiality, integrity, and availability. Access control techniques imposes authorization policies, which is used to impose consent by preventing access to anything that should not be accessible to the user [7]. Presently, access control has seen application at various levels in numerous domains for resource management; allowing only authorized users access to resources in an authorized way.

2.3 Traditional Access Control Mechanism

Traditional access controls utilize static and prearranged policies to regulate access decision. Thus, the same decision is made in different scenarios by these static policies. Even though, it has been successfully applied in diverse environments to solve numerous problems, the traditional access controls are developed to provide an association between information related to an access control policy and a resource being requested access to. An access control implementation is subject to handling, ranging from an unexpected situation to numerous malevolent entities gaining access to existing accounts. Thus, traditional access control technique provides a set of drawbacks; such that unexpected circumstances cannot be handled based on the static and prearranged policies. This strict approach lacks the required strong security measures for numerous dynamic and decentralized systems, which requires increased flexibility in accessing resources. As an alternative, this unchanging approach is best suited for scenarios where there is no collection of contextual features during the access request [8]. There are numerous traditional access control techniques which includes Access Control List (ACL), Discretionary Access Control (DAC), Mandatory Access Control (MAC), Role-Based Access Control (RBAC) [9].

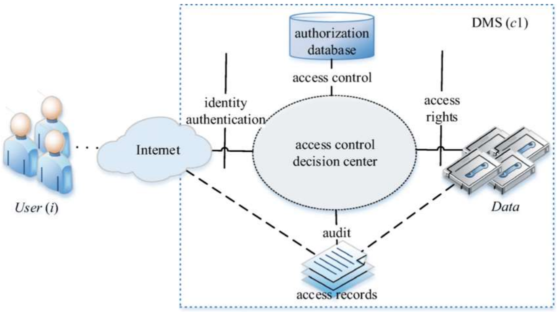


Figure 2: Traditional Access Control Framework in Distributed IoT [10]

2.4 Dynamic Access Control Mechanism

The fundamental basis of dynamic access control techniques is that contextual features are considered, along with the access policies, which are collected at the access request time to make access decisions [11]. Thus, providing increased flexibility and can be fine-tuned to various scenarios while making the access decision. The prerequisite to accept dynamic access control should be an important priority when providing a well-organized and adaptable access control technique. Nevertheless, with existing access control techniques depend on rigid access policies; they are inadequate in providing guidelines towards improving automation. With the absence of automation and the involvement of human analysis, existing access control techniques are open to errors and susceptible to several types of cyber-attacks. Moreover, traditional access controls are limited with resolving real time risks and threats when handling a formerly unknown threat. This is based on the facts that access decisions are hinged on a set of policies, that cannot determine diverse access control circumstances in a timely manner but can only resolve problems that were previously identified [12]. As opposite to static policies, dynamic access control techniques apply real-time features, such as trust, setting, risk, history, and operational need, to deliver access decisions. In addition, dynamic access control can adapt to different scenarios at the decision-making time.

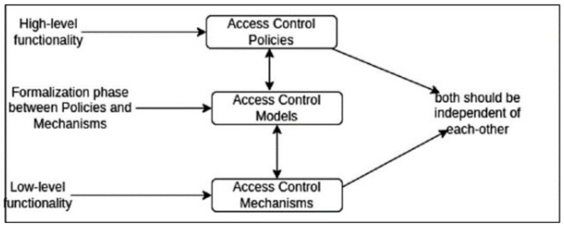

High-level functionality
Access Control Policies
Formalization phase between Policies and Mechanisms
Access Control Models
both should be independent of each-other
Low-level unctionality
Access Control Mechanisms

Figure 3: Access Control Models in Dynamic Environments [13]

2.5 Risk Estimation

Risk estimation is the process of assessing and quantifying the level of risk associated with a particular event or situation. It involves analyzing various factors and information to determine the likelihood and potential impact of a risk occurring. Risk estimation can be complex and challenging due to the diverse nature of reality and the absence of a uniform methodology for assessment and estimation. Different methods and mathematical tools are used in risk estimation, such as statistical models, Bayesian methods, and fractal theory [14]. In the financial sector, risk estimation is crucial for measuring and managing risks in areas such as stock indices and systemic risk in major banks [15]. The accuracy of risk estimation is important for decision-making and regulatory purposes, as it helps in identifying and mitigating potential risks.

Risk estimation in access control is a crucial aspect in the field of Internet of Things (IoT) security. Existing access control models are often static and cannot adapt to changing and unpredictable situations. To address this issue, researchers have proposed dynamic models such as the risk-based access control model. This model utilizes real-time and contextual features to make access decisions based on estimated risk values [16]. Several techniques have been proposed for risk estimation, including the Adaptive Neuro-Fuzzy Inference System (ANFIS) model [17], the Neuro-Fuzzy System (NFS) model, and the fuzzy inference system with expert judgment. These techniques aim to provide accurate and realistic risk values for each access request, taking into account various factors in the access control environment. The effectiveness of these techniques has been evaluated in different scenarios, such as smart homes [18], a children's hospital, and a network router. Overall, the goal is to develop efficient and reliable risk estimation techniques that can adapt to the changing conditions of the IoT environment.

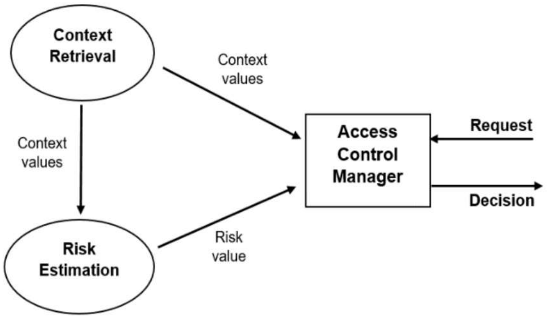


Figure 4: Risk-based Access Control Overview [19]

2.6 Security Policy

A security policy is a crucial component for organizations to define and enforce their approach to security. It serves as a central repository for intangible aspects such as corporate philosophy, mission statements, culture, and attitude to risk, which can then be translated into measurable action statements and procedures [20]. A security policy system is designed to manage and negotiate policy information across different security fields, providing suitable policy information to ensure safe communication between hosts and security gateways. It is important to introduce security controls in a controlled manner, with a formal policy in place, to ensure effective implementation of security measures [21]. Policy acts as a primary guideline for audits and defines the control framework for an organization [22]. The term "security policy" has different meanings, and a clearer definition is needed to facilitate research and standardization efforts in computer security.

Access control policies in security can be classified into different types based on their purpose and characteristics. One type is mandatory security policy, which are essential rules that must be followed to ensure information security [23]. Another type is optional security policy, which provide additional security measures that can be implemented based on specific requirements [24]. Additionally, user-defined security policy allows users to customize access control policies according to their needs [25]. In the context of dynamic policy environments, there is a need for managing policy changes and updates. This includes analyzing and classifying when an update policy occurs and providing a solution for such dynamic policies [26]. Time-dependent policies for access control are also important, where access control policies are dependent on factors such as the content of the data, the flow of information, and the time [27]. Finally, there is a focus on adapting access control policies based on conviviality recommendations, which aim to make access control mechanisms more user-friendly and less restrictive.

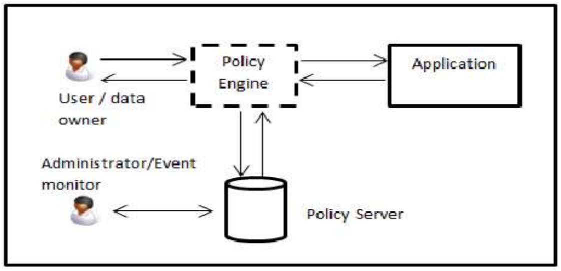


Figure 5: Common Framework of Access Control Policy [28]

2.7 Machine Learning

Machine learning is entwined with advancing algorithms that allows computer to learn. Learning is a process of discovering measurable normalities or different examples of data. The machine learning algorithms are made to characterize the human method of learning some tasks. These algorithms can likewise signify an understanding into relative strain of learning in various conditions [29]. Also, it is one of computer science quickest developing field with broad applications. Machine learning instruments are involved with creating programs with the capacity to learn and adapt [30]. With the consistently expanding quantities of data opening up, there is a valid justification to accept that smart data analysis will turn out to be extensively continuously undeniable as an indispensable component of innovative progression [31]. These days, the turn of event and improvement of new processing advancements in Big Data applications, machine learning has enhanced radically compared to the past. Today, large number of machine learning algorithms have been created, enhanced and the new enhancement in machine learning turns into the capacity to consequently apply an assortment of complicated numerical computation to a big data, which computes the outcomes a lot quicker [32]. The core of machine learning is to accumulate the data, with the experience the program learns, to create valuable information. For example, the most common way of isolating legit messages from spam messages. The input will be some documents or words comprised in the message; and the result ought to be yes or no demonstrating the message is spam or not individually, yet we do not have an algorithm to precisely recognize spam messages. Machine learning offers a solution for this task, we make available examples of the messages categorized as spam or valid and the program can inevitably figure out how to recognize them [33] [34].

2.8 Machine Learning Algorithms

- **K-Nearest Neighbour (KNN):** This is one of easiest and earliest classification algorithms. It uses instance-based reasoning. K-nearest neighbour (KNN) concept includes two objects of a comparable class having specific likenesses quantifiable utilizing distance metrics. For example, an object with an unknown class is either gathered into a comparable group as its first nearest neighbour or into the leading class via the votes of the k-nearest neighbours (*k* represents odd numbers). Through the usage of empirical cross validation, the population of the neighbours, *k*, is typically chosen. K-nearest neighbour places objects as point vectors in a multi-layered feature space [35]. K-Nearest Neighbour is nonparametric. With this flexibility, it is a trustworthy classifier in instances of datasets with high scopes in classifying malware [36].
- **Artificial Neural Networks (ANN):** It merges the reasoning power of the human brain with computational power of machine. It utilizes neurons as the determining sites and the edges between neurons to compute the involvement of each neuron in the preceding layer in the decision and result at the present neuron. It depends on pattern recognition. Its training can either be supervised or unsupervised [37].

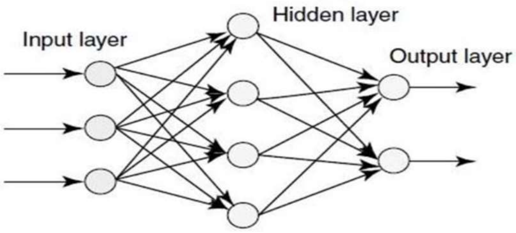


Figure 6: Structure of Artificial Neural Network [38]

- **Support Vector Machine (SVM):** It is a supervised learning algorithm wherein a given dataset is separated into various classes utilizing a hyperplane. The objective of SVM is to find this hyperplane. There could be numerous hyperplanes, however finding an optimal hyperplane is vital. The points nearest to the hyperplane in the various classes are known as support vectors and these support vectors are utilized to forecast the classes of new data points. A new incoming point is placed on the equation of the hyperplane and afterward is classified to which class it fits on the foundation of which side of hyperplane it falls on the vector space. To train our machine, supervised data is fed, that is, data with results already known. It learns the behaviour of fraudulent and authentic transactions and then it can categorize new transaction as to which class it fits [39].
- **Logistic Regression:** This is a classification operation that utilizes class for building and uses a solitary multinomial logistic regression model with a solo estimator. Logistic regression typically expresses where the limit between the classes exists, likewise expresses the class probabilities depend on distance from the limit, in a precise approach. This moves towards the extremes more quickly when data set is larger. It makes sturdier, more comprehensive forecasts, and can be fit in various ways; but those sturdy forecasts could be incorrect. Logistic regression is an approach to forecast. Nevertheless, with logistic regression, forecast results in a dichotomous result [40]. Logistic regression is one of the most generally utilized tools for applied statistics and discrete data analysis. Logistic regression is linear interpolation [31].
- **Decision Tree:** This is a computational tool for classification and forecast. A tree includes of interior nodes which signify a test on a quality, each branch signifies a result of that test and each node (terminal node) holds a class label. It recursively segments a dataset utilizing either depth first greedy approach or breadth first greedy approach and stops when all the elements have been allotted a specific class. For the segment rule to be effective, it should isolate the data into groups where a solitary class prevails in each group. All in all, the best segment will be the one in which the subsets do not intersect, that is, they are obviously disjoint to an extreme amount [41].

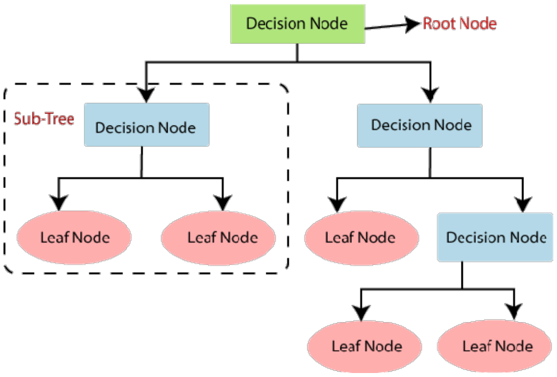


Figure 7: Example of Decision Tree [38]

- **Naïve Bayes:** The Naïve Bayes (NB) classifier is one of the essential probabilistic classifiers established on the solid autonomous assumptions that exist between features. It acknowledges that the accessibility or inaccessibility of a specific feature is sovereign of the accessibility or inaccessibility of another. Results are forecasted based on probabilities while utilizing the Bayesian classification. This classifier has a high accuracy rate, it is quick and effective. It is self-determining and based on statistics [42]. Training data using naïve bayes is very quick because of the way that it calculates the probability of the given class alone. In contrast to other Bayesian classification algorithms, naïve bayes does not manage unnecessary and unimportant features in the dataset because it will extend the discovery process and could influence the general performance of the system [43]. This algorithm is utilized in many research fields, such as text classification, spam filtering, online applications. It is the best learning algorithm for categorizing text documents [44].
- **Random Forest:** Random forest (RF) is a mix of tree prognosticators. The trees depend on the values of the random vectors tested for them. Prognosticators are arbitrarily chosen for yielding trees [45]. Random forest yields various trees and chooses the features to coordinate into each model via random selection. Be that as it may, the trees produced are not pruned [46]. The sampling of random subset features creates the random forest for each decision tree. According to [47], the accurateness of random forest is enhanced via the infusion of arbitrariness at each node of the developed tree. The relationship existing between trees is abridged via the selection of random features capable of improving the prognostication power and effectiveness. It effectively manages datasets with high dimensions encompassing missing values and can likewise manage binary data, categorical data and continuous data. Random forests overcome the issue of overfitting, they are less complex to irregularity data, trees are not pruned in light of the fact that parameters are generally set, implementation is easy, and has a high accuracy.

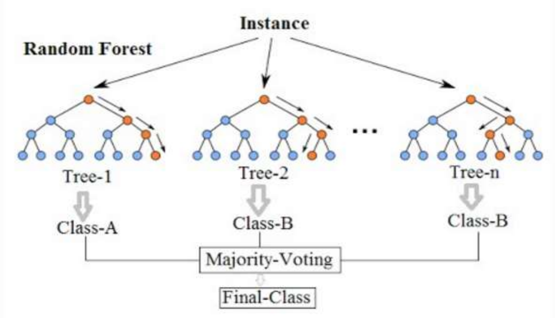


Figure 8: A pictorial representation of random forest algorithm [48]

- **Gradient Boosted Decision Trees:** This is a machine learning technique for enhancing the prognostic value of a model via moderate steps in the learning process. Each decision tree iteration includes changing the values of the factors, loads, or biases applied to all the information factors being used to predict the objective worth, determined to limit the loss function (the extent of differentiation between the expected and sincere objective characteristics). The slope is the steady change made in each progression of the interaction; boosting is a technique for rapid enhancement in predictive accuracy to an acceptable optimal value. Gradient-boosted decision trees are a popular technique for handling forecasting problems in both classification and regression domains. The technique additionally advances the learning system by dealing with the objective and reducing the quantity of iterations to get a reasonable ideal solution.

2.9 Review of Related Literature/Works

The advent of cloud storage and collaborative efforts has presented challenges in securely managing data. To address this, [49] introduced an access control mechanism using attribute-based encryption. This cryptographic approach facilitated multi-party dispersal and data security. Their proposal merged direct and indirect reversal schemes, improving flexibility and unauthorized data alteration detection.

In the realm of healthcare, the Personally Controlled Electronic Health Record (PCEHR) system in Australia faced concerns of unauthorized access. [50] proposed a "Log-in-Pair" access control model to enhance privacy and security. By utilizing real-time contextual information, [51] designed an adaptive risk-based access control model for IoT systems. This model calculated security risks associated with access requests, leveraging user attributes for access decisions.

[52] suggested novel access control techniques based on the hash tree, offering efficient user access capability reversal. [53] explored categorical quantum cryptography for cloud-based access control. [54] focused on IoT communication, proposing an access control system for web-based services, enhancing security and platform independence.

The growth of IoT led to the need for robust access control mechanisms. [55] introduced a fuzzy-extended attribute-based access control (FBAC) method. This method improved time productivity and usability while maintaining security. [56] advocated for a decentralized access control model using blockchain to ensure data owners' rights. [57] combined blockchain with IoT, introducing a non-interactive access control system.

Addressing smart home security, [58] proposed a blockchain-based access control for smart homes. Additionally, decentralization using blockchain was recommended by [59] to address IoT security concerns. To enhance LoRa terminals' security, [60] introduced a lightweight gateway architecture.

The integration of blockchain with access control garnered attention. [61] developed a revocable attribute-based access control system using blockchain. The blockchain-enabled system met essential security criteria. [62] proposed a blockchain-based adaptive dynamic access control approach to address zero-trust architecture limitations.

In conclusion, various studies have explored and proposed access control mechanisms, leveraging attributes, cryptography, and blockchain to address security concerns in diverse technology environments. These mechanisms aimed to enhance data privacy, thwart unauthorized access, and improve system performance while adapting to emerging technological challenge

2.10 Limitations of Closely Related Literature

From the literature reviewed, two research works focused on providing dynamic access control, using used a real-time user abnormal behavior detection model based on deep learning (referred to as MAN-SVMDT), which combines a neural network based on a multi-layer attention mechanism (referred to as MAN) with a support vector machine based on a decision tree (referred to as SVMDT). And an adaptive risk-based access control model which utilizes real-time contextual information associated with the requesting user to calculate the security risk regarding each access request respectively. The outcome of the research works was encouraging but are limited by the computational complexity of the algorithms. The first research introduced user trust evaluation into a role-based access control model, using a deep learning-based user abnormal behavior detection algorithm to dynamically evaluate user behavior status and update trust, establishing a short-term token-based authorization mechanism and smart contract-based decentralization authorization management framework. However, computation complexity is a downside. Also, the last research used the user’s attributes as inputs to analyze and calculate the risk value to determine the access decision. To detect abnormal and malicious actions, smart contracts are used to track and monitor user activities during the access session to detect and prevent potential security violations. In addition, the fuzzy inference system with expert judgment as an optimal approach is used to handle risk estimation process. However, implementation and evaluation are required. It can be concluded that attempt at improving access control mechanism has resulted in additional performance issues. Therefore, the primary purpose of this research is to develop an adaptive Risk-based Access Control system that is able to respond dynamically to changing access rights. Thus, adjusting user access adaptively based on the changing requirements and users’ behaviour during access sessions.

3. METHODOLOGY

This section presents a description of the methodology used to satisfy the objective of this research work. These includes several processes, procedures and architectural structures adopted within the research.

3.1 The Adaptive Risk-based Access Control System (ad-RACs)

To enable dynamic access control, real-time system attributes are required to generate the access decision. The ad-RACs has four inputs (risk factors): user context, sensitivity of resource, severity of action and history of risk. These inputs are required to evaluate the security risk related to each user's access request. Which is then compared with the access control policies to generate the necessary access decision. To enable adaptiveness of the proposed solution, analyzing user access decision can provide insights into user behavior and usage patterns, such as identifying common access patterns, frequently accessed resources, or unusual access activities, which is used to update the user's history of risk. Thus, adjusting user access adaptively based on the changing requirements and users' behaviour during access sessions. The ad-RACs proposed solution can offer a suitable security level while guaranteeing flexibility and scalability.

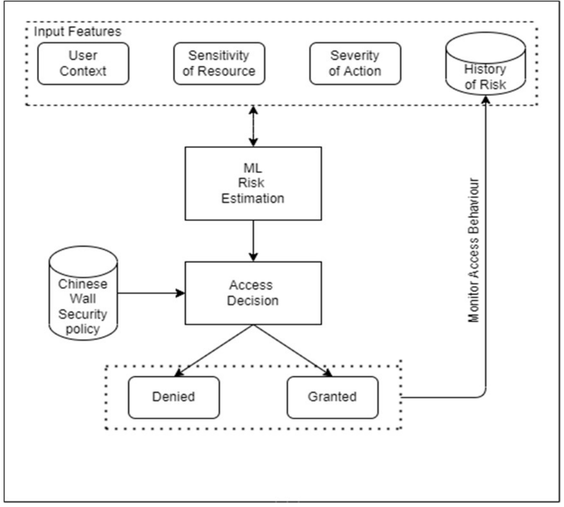


Figure 9: The Adaptive Risk-based Access Control System (ad-RACs)

Figure 9 illustrate the adaptive risk-based access control system (ad-RACs); showing how the various components connect and interact. The following subsections elucidate the various components of the proposed model.

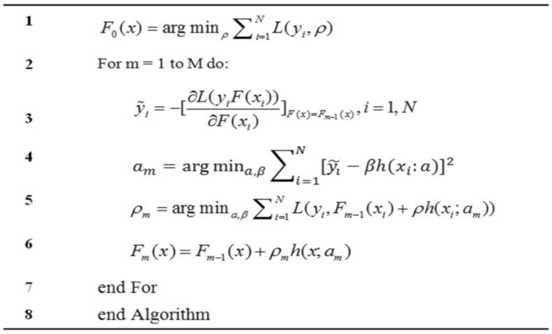
3.2 Input Features

- The **user context** characterizes the system attributes that are included at the moment the user requests access. These attributes are employed to ascertain the security risk related to the user's access request. For example, username, password, location and time.
- **Sensitivity of Resource** characterizes how valuable the resource is. Resources are allotted a sensitivity level based on who should be allowed to access the resource and how much loss is incurred if revealed. A risk metric is allotted to each resource depending on how valued the resource is. For example, the higher the sensitivity, the higher the risk metric.
- **Severity of Action** characterizes the penalties associated with action on a specific resource based on the security requirements of confidentiality, integrity, and availability. Diverse procedures have diverse influences and consequently have diverse risk values. For example, the risk associated with "read" procedure is lesser than the risk associated with "delete" or "write" procedure.
- The **history of risk** is utilized in evaluating each access request risk value, because it keeps records of previous patterns in users' access behaviour. This is useful in identifying good and bad authorized users and predicting their future behaviour.

3.3 Risk Estimation Module

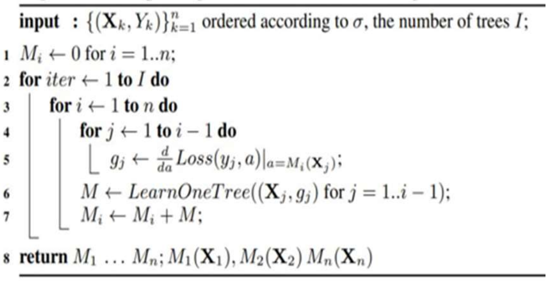
This module is in charge of utilizing the input features to evaluate the risk related to the access request. Essentially, the objective is to create an effective risk estimation procedure. The access decision ascertains whether the access is granted or denied in accordance with the access control policy. A pretrained machine learning model (CatBoost) serves as a risk estimator using the input features from the current access request to establish a risk estimate. Concretely, this estimate will be a probability of risk related to the access request, which is dependent on the complexity of the relationship between the input features and access request risk.

CatBoost is a gradient boosting decision tree algorithm introduced in 2017 by Anna Veronika, Dorogush, Vasily Ershov, and Andrey Gulin. During training, a set of "oblivious" decision trees are built consecutively. Oblivious trees are regular decision trees designed to use the same criteria at each level of the tree, successive trees are built with reduced loss compared to the previous trees. This helps reduce overfitting which is common with gradient boosted trees. CatBoost changes categorical values into numbers utilizing target statistics on mixtures of categorical features and mixtures of categorical and numerical characteristics.



```
$F_0(x) = \arg\min_\rho \sum_{i=1}^{N} L(y_i, \rho)$
For m = 1 to M do:
    $\tilde{y}_i = -[\frac{\partial L(y_i F(x_i))}{\partial F(x_i)}]_{F(x)=F_{m-1}(x)}, i = 1, N$
    $a_m = \arg\min_{a,\beta} \sum_{i=1}^{N} [\tilde{y}_i - \beta h(x_i : a)]^2$
    $\rho_m = \arg\min_{a,\beta} \sum_{i=1}^{N} L(y_i, F_{m-1}(x_i) + \rho h(x_i; a_m))$
    $F_m(x) = F_{m-1}(x) + \rho_m h(x; a_m)$
end For
end Algorithm
```

Figure 10: Classic Gradient Boosting Algorithm [63]



```
input : $\{(\mathbf{X}_k, Y_k)\}_{k=1}^{n}$ ordered according to $\sigma$, the number of trees $I$;
$M_i \leftarrow 0$ for $i = 1..n$;
for $iter \leftarrow 1$ to $I$ do
    for $i \leftarrow 1$ to $n$ do
        for $j \leftarrow 1$ to $i - 1$ do
            $g_j \leftarrow \frac{d}{da} Loss(y_j, a)|_{a=M_i(\mathbf{X}_j)}$;
        $M \leftarrow LearnOneTree((\mathbf{X}_j, g_j)$ for $j = 1..i - 1)$;
        $M_i \leftarrow M_i + M$;
return $M_1 \ldots M_n; M_1(\mathbf{X}_1), M_2(\mathbf{X}_2) M_n(\mathbf{X}_n)$
```

Figure 11: Gradient Boosting in CatBoost Algorithm [64]

CatBoost presents two crucial algorithmic advances – the implementation of **ordered boosting** and an innovative algorithm for **processing categorical features**. Both techniques utilize random variations of the training examples to match the forecast shift produced by a special kind of target leak present in all existing implementations of gradient boosting algorithms.

3.4 Access Control Policy

This policy is primarily utilized by the risk estimation module to generate access decisions. The policy is developed by the owner of the resource, to categorize the rules for responding to a user's access request. The total risk evaluated by the risk emulation module is compared with the access control policy to ascertain the access decision. For the access control policy, the Chinese wall security model is adopted. This security model was presented in 1989 by Brewer and Nash, with focus on the conflict-of-interest concept. This security model merges the fundamentals of mandatory and discretionary and is able to achieve the security goals of confidentiality and integrity. The features of Chinese wall model are objects, subjects, conflict-of-interest classes, datasets and labels. The principal governing this security model is that users are not granted access to a private information present in the domain of an organization and its competitors. No-wall users are prepared and in the event a file containing the information of the competitors is available, it is converted to unavailable. As such, this security model dynamically regulates the access control rules based on the user's behavior and access rights. In this security model, resources are clustered into diverse conflict-of-interest classes. Based on mandatory principles, all users have access rights to at most one resource in any conflict-of-interest class. The policy of this security model ascertains that a user only has access to a resource, if and only if the resource requested is part of the resources always accessed by the user, or the resource requested is not in the conflicts-of-interest classes available to the user. The Chinese Wall security policy attempts to improve the access control flexibility and adaptiveness by dynamically changing access rights based on changing requirements and user access behaviour throughout the access session.

The access control module is then created by firstly describing what is meant by a Chinese Wall and secondly, by developing a set of policies such that users (subject) can only access resources (objects) on the right side of that wall. There are three levels of significance:

- **The lowest level:** This considers specific resources concerning a service provider. These resources are referred to as objects.
- **The intermediate level:** All resources concerning a service provider are categorized and called a service provider's dataset.
- **The highest level**: All service provider's datasets who are in conflict are categorized and each of such category is called a conflict-of-interest class.

A The foundation of the Chinese Wall policy is that users only have access to resources not held in conflict with any other resource already possess. Considering the network, resource already utilized by a user are resources held on user's device, and has earlier accessed. Thus, consider the datasets for Provider A, Provider B and Provider C. And also, Provider B and Provider C are in the same conflict-of-interest class different from Provider A. A new user may easily choose to access any datasets preferred; considering no resource is possessed hence no conflict exist. But such conflict may later exist. Furthermore, suppose the user accesses the dataset of Provider B first; that means the user now possesses resource relating to Provider B dataset. Consequently, the user then requests access to the Provider A dataset. This is allowable considering Provider A and Provider B datasets belong to dissimilar conflict-of-interest classes and thus no conflict exists. Nevertheless, if the user requests access to the Provider C dataset, the request must be denied due to the existence of a conflict-of-interest in the requested dataset (Provider C) and one already possessed (Provider B). Thus, the user possesses {"Provider B", "Provider A''} datasets. Note that, it does not matter whether the Provider B dataset was accessed before or after the Provider A dataset. However, if Provider C dataset was accessed before Provider B dataset, the restrictions would change. In this case, access rights to the Provider B dataset would be denied and the user would possess {"Provider C", "Provider A"} datasets. This describes the Chinese Wall. In the first instance, users have complete freedom to access any resource. But once the decision has been made, a Chinese Wall is set up around that dataset and "the wrong side of this Wall" is any dataset in the same conflict-of-interest class as the dataset within the Wall. However, users have freedom to access any other dataset in a different conflict-of-interest class, but immediately that decision is made, the Wall shape changes to accommodate the new dataset. Thus, the Chinese Wall policy is a refined blend of discretionary and mandatory control.

3.5 The Process Flow of the Adaptive Risk-based Access Control System (ad-RACs)

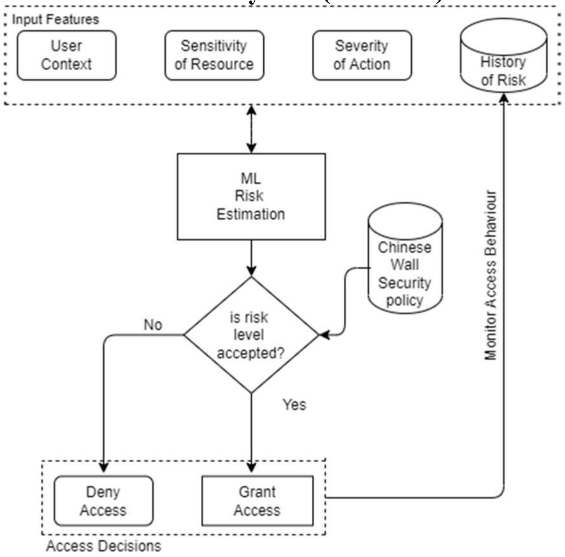


Figure 12: The Process flow of the Adaptive Risk-based Access Control System (ad-RACs)

Figure 12 illustrates how the process flow starts when the ad-RACs receives a user's access request. The risk estimation module takes the required input features (user, resource, action and risk history) of the user, and uses a pretrained CatBoost model to evaluate the risk concerned with a user access request. If it is within the acceptable risk threshold, the access control policy is then used to ascertain the access decision. This access decision is used to update the user's history of risk. Thus, we have the following decisions:

- A new user has access to any preferred resource, as no conflict-of-interest exists.
- If the user accesses a resource in B and request access to resource in A, then access is granted if B and A belong to different conflict-of-interest classes.
- If the user accesses a resource in B and request access to resource in C, then access is denied if B and C belong to the same conflict-of-interest classes.

3.6 Algorithm in achieving the Adaptive Risk-based Access Control System (ad-RACs)

This procedure is divided into two stages which includes risk estimation and access control. The method determines that no malevolent action was shown given a clear concise path.

```
INPUT: Input Features
OUTPUT: Access Decisions
    Begin
        a. Input request features
        b. Estimate access risk
        c. Is the risk accepted?
        d. IF yes, apply access policy
        e. Grant access
        f. ELSE, deny access
        g. Update history of risk
    End.
```

3.7 Risk Estimator Evaluation Parameter

The performance of the risk estimator is evaluated taking into consideration the following metrics.

- **Response Time:** This poses the question; will the access control system be able to process user access requests in a timely fashion in line with the operational needs? Access control execution requires a number of operations to allow a user's access request, and to check for the risk related to the access request. The metric can be attained by the computational intricacies calculation conferring to the system model.
- **F1 Score:** In practical use, it is expected for any system controlling access that the number of granted access will significantly outnumber denied access, because for regular requests access is usually granted and access is only denied in cases of malicious or unclear activity/requests. For this reason, the f1 score is used to evaluate the model performance correctly given the imbalanced that will emerge

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

in the access control system. F1 score is the harmonic mean of the Precision and Recall of the model. It is a value that represents the trade-off made by the model, whilst dealing with Precision and Recall, it ranges from 0 to 1, where 1 represents a perfect model, and 0 represents a no skill model. A middle score like 0.5 indicates, a good Recall, bad Precision or vice versa.

- **Area Under the Characteristic Receiver Operating Curve:** The Area Under the Characteristic Receiver Operating Curve (AUCROC) measures how well the model generalizes on the data it is trained on. It measures the true positive rate versus the false positive rate at dissimilar probability limits to determine if there are certain distributions of the data that cause the model to produce more false positives.
- **Brier Score Loss:** The Brier Score Loss is described as the mean squared error of the predicted probabilities of a model. The Brier score loss enables us to estimate the confidence a model has in its predictions of probabilities. It takes the prediction made by the model and retrieves the error or difference compared to the actual prediction and squares this average before averaging the result across all the classes.

3.8 Benchmarking Parameter

To compare the performance of the Adaptive Risk-based Access Control System (ad-RACs) against existing access control techniques the following metrics would be used.

- **Precision:** Often referred to as positive predictive value, gauges how accurately the adaptive access control system makes accurate predictions. It determines the proportion of accurate positive instances (also known as true positive instances) to all of the system's positive predictions. Precision in an adaptive access control system reveals the proportion of allowed accesses that were truly approved and legitimate. Low false positives and potential security breaches are indicated by a high precision, demonstrating that the system is correctly detecting authorized users.
- **Recall:** Also referred to as sensitivity or true positive rate, measures how well a system can recognize every instance of a given class. It determines the proportion of accurate positive forecasts to all positive cases. Recall evaluates how effectively the system recognizes and permits access to authorized users in the context of adaptive access control. A high recall means the system successfully detects the majority of authorized users, lowering the possibility of false negatives where legitimate people are denied access.
- **F1 score:** The harmonic mean of recall and precision is the F1 score. In order to evaluate the overall effectiveness of the adaptive access control system, it provides a balance between these two indicators. Due to the F1 score's consideration of both false positives and false negatives, it is especially helpful when the class distribution is unbalanced. A system that successfully detects authorized users while minimizing false positives and false negatives has a higher F1 score in adaptive access control.

3.9 Requirement Specification

This section presents the essential settings to be met in other to guarantee the success of this research. The ad-RACs was built to satisfy the functional requirements.

- The system has a data entry page to receive requests for access.
- The system has a data record output available if access is granted.
- The system assigns a risk score to all granted access requests.

4. IMPLEMENTATION AND EVALUATION

This section presents the results, findings, and inferences from the study.

4.1 Implementation

This research makes use of Python as the programming Language. It is implemented in a Jupyter Notebook, which allows for iterative experimentation, visualization and recording of results. A synthetic dataset using the input features (severity of action, history of risk, sensitivity of resource, user context. There will be no pre-processing since the dataset is synthetically generated. The output of tests and evaluation is a graphical presentation of the outcomes and the performance of the ad-RACs. It is also available via the Jupyter Notebook, but can be viewed without any extra requirements. An environment which simulates the behaviour of input features supplied by edge devices is also simulated with StreamLit. StreamLit is an open-source python package for creating dataApps, this enables us to simulate from end-to-end the flow of logic and information in the ad-RACs.

4.2 Data Preparation and Feature Distribution

The data for this study is the amazon access control dataset, which contains 32769 entries and 5 features including the target feature [65]. The feature names are. User context, resource sensitivity, action severity, risk history and the target (decision). The dataset is highly imbalanced with access granted decisions accounting for 93.75% of all instances.

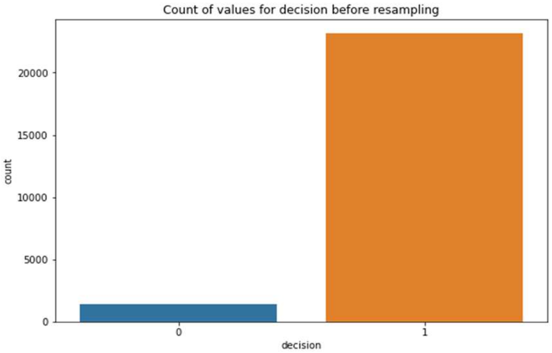


Figure 13: The Imbalanced Dataset with High access granted decision

To deal with this imbalance, a strategy of oversampling and undersampling is employed. This strategy combines SMOTEN (oversampling) and ENN (Undersampling) to resample and balance the data. SMOTEN (Synthetic Minority Oversampling Technique for Nominal data), works by oversample representing entries in a feature space and grouping similar entries together, similar to the K-Nearest Neighbors algorithm. In these groups a line is drawn between members and a synthetic (new) entry is added on this line. ENN (Edited Nearest Neighbors) is an undersampling technique which works by removing samples or entries which are not representative of its neighbors. For example, if a sample “yes” has two neighbors which are “no” then it is removed from the dataset. The point of this is to ensure that there is more separation between entries in the two different classes.

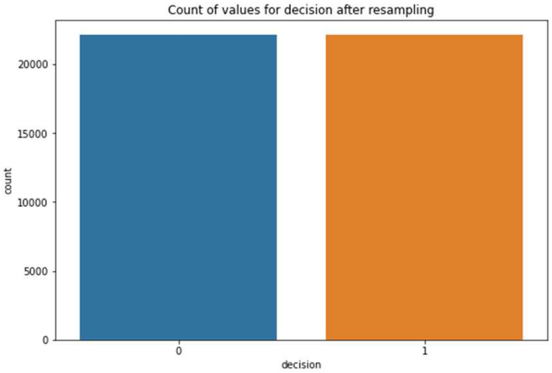


Figure 14: The distribution of the target (decision) after Resample

In the Figures 15 to Figure 18, we explore whether certain features may have predictive quality in terms of their distribution in relation to the target decision.

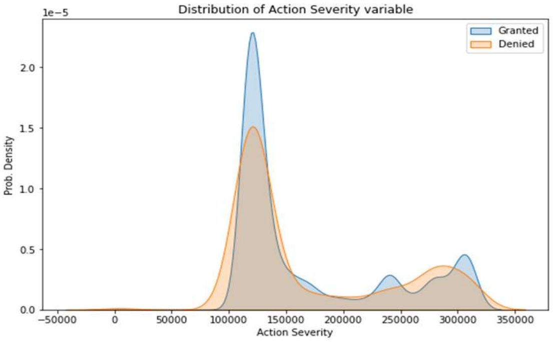


Figure 15: Distribution of Action Severity Variables

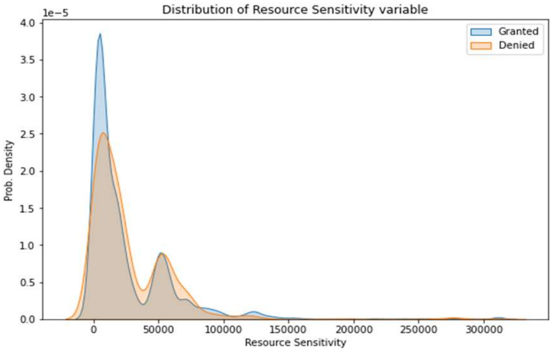


Figure 16: Distribution of Resource Sensitivity Variables

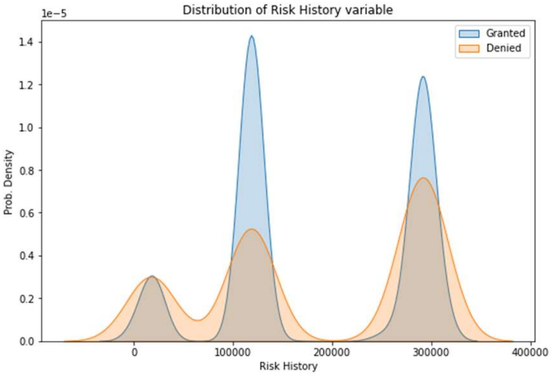


Figure 17: Distribution of Risk History Variable

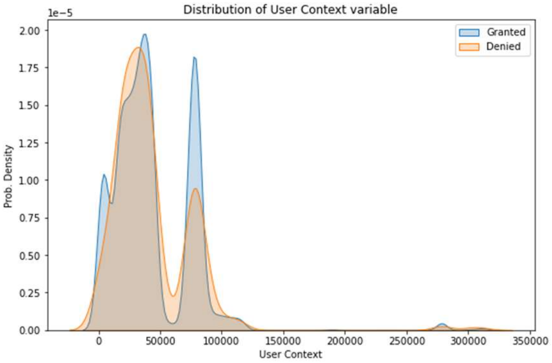


Figure 18: Distribution of User Context Variable

From the Figures, it is seen that both granted and denied access requests have the same distribution with peaks and troughs in the same regions.

4.3 Performance Evaluation of the Risk Estimation Model

In this section we compare the performance of various risk estimator models given the data and compare them to the proposed model. The metrics used to make comparisons are AUCROC, F1 score, Brier Score loss, Response time.

From Table 1 the performances of the models tested, are shown. The metrics used are F1 score, Area Under the Curve of the Receiver Operating Characteristic, Brier score, and Response Time. The results show a general improvement across all metrics between single classifiers and ensemble models, the results also suggest that Tree based algorithms such as Decision Tree, Random Forest, Bagging perform better across all metrics, when compared to other types of machine learning algorithms.

Table 1: Performance of Risk Estimator models.

Algorithms	***F1 Score***	***AUCROC***	***Brier Score***	***Response Time***
Logistic Regression	0.789078	0.491645	0.342243	0.001547
K-Nearest Neighbors	0.933254	0.554128	0.123764	0.352373
Decision Tree	0.958544	0.554139	0.078970	0.002740
Support Vector Machines	0.755871	0.545839	0.381911	50.810147
Naïve Bayes	0.643486	0.511876	0.509337	0.002695
Artificial Neural Network	0.648626	0.539330	0.502380	00.16960
Ensemble Models				
Random Forest	0.964249	0.545784	0.068595	0.271960
AdaBoost	0.846969	0.536895	0.260710	0.101529
Bagging	0.958976	0.548587	0.078238	0.025989
Proposed Model	**0.971872**	**0.543397**	**0.054437**	**0.033512**

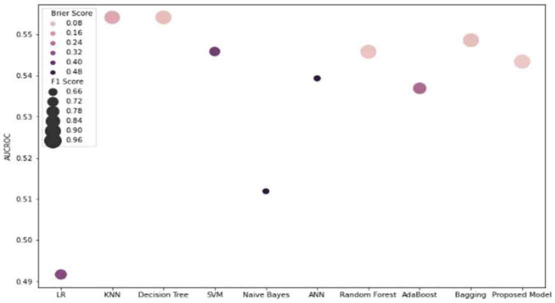


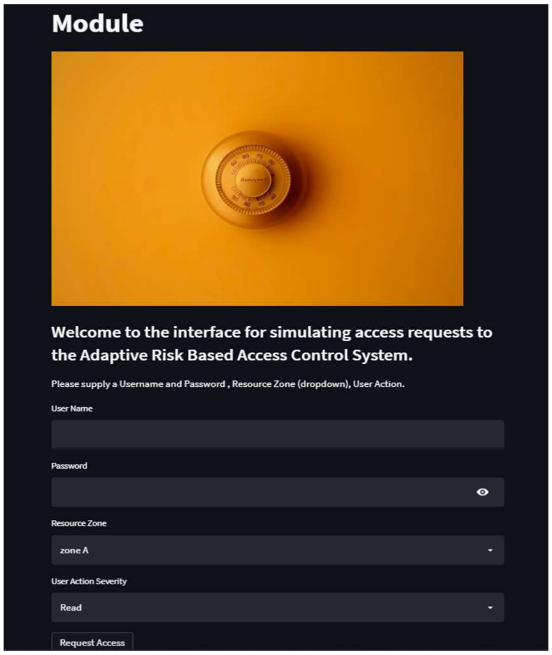
Figure 19: A Visualization of the Models Performance in relation to all three metrics

Figure 19 presents a visualization of how the models perform in relation to all three metrics described above. In the plot, the higher the blob, the better the AUCROC, the darker the blob, the worse the Brier score, and the larger the blob the better the F1 score. The better models therefore, have a large light-colored blob, high on the graph. The models having the best performance across these metrics as identified from Figure 19 as K-Nearest Neighbors, Decision Tree, Random Forest, Bagging and the proposed Model. With the proposed model having the lightest color (best Brier Score). While the Artificial Neural Net, Logistic Regression and Naïve Bayes model perform comparatively worse. AUCROC as a metric is a useful comparison across datasets and models that try to model problems in the same domain, and is general, better for describing performance than other metrics like accuracy. However, for implementing the model the F1 score and Brier score are of more important because they inform us of the particular performance of the model, and are often monitored in a deployment environment.

4.4 Evaluation of the Adaptive Risk-Based Access Control System (ad-RACs)

In this section, the simulated ad-RACs is evaluated. The system was built to meet the requirement specification using python and consists of two units; the risk estimation module and the access policy module. The risk estimator is based on the CatBoost gradient boosting algorithm in python. It is first trained on a subset of available data, then its performance is validated using another subset of data. If the performance of the model is not desirable its hyperparameters (learning rate, leaf regularization, and tree depth) are changed using a search space algorithm called grid search. When the optimal set of hyperparameters are achieved, the model is saved as binary file. This file will later be read into the simulation system. The access control policy is based on the Chinese wall security policy.

- **Login Page:** In the login page, a new or returning user may supply their username and password, this is recorded by the database for further processing, then the user is expected to supply a zone for data access and the user action. The username, password, number of login attempts are recorded and used to generate a user content context variable which is also stored in the database. The stored user context is later used as an input feature to the machine learning algorithm in order to determine whether to grant access or not.


Module

Welcome to the interface for simulating access requests to the Adaptive Risk Based Access Control System.

Please supply a Username and Password , Resource Zone (dropdown), User Action.

User Name

Password

Resource Zone

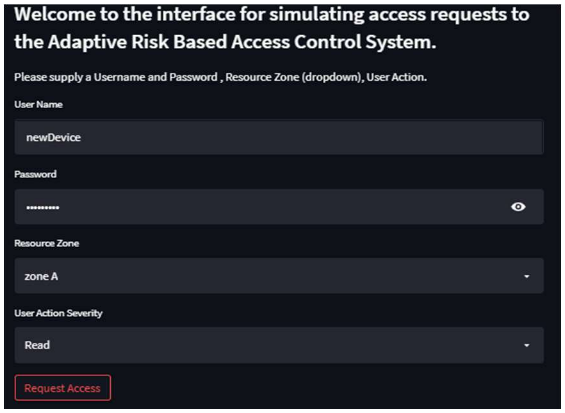
zone A

User Action Severity

Read

Figure 20: The Login page for the simulated adaptive risk-based access control module

In the simulated environment the risk estimation model takes four inputs, user context, resource sensitivity, history of risk, and severity of action to estimate the risk of granting access to the user. If access is granted to a user, the Chinese wall policy module determines whether to give access to data depending on the policy's rules. Generally speaking, the rule is that no user from one zone may access (read/write/create/delete) data from a zone that it is in conflict with. If there is no conflict access is granted to the user. If the user is new, then the selection is stored by the system to ensure that when the user returns, these details can be used to determine whether to give access and to what zone.



Welcome to the interface for simulating access requests to the Adaptive Risk Based Access Control System.

Please supply a Username and Password , Resource Zone (dropdown), User Action.

User Name: newDevice

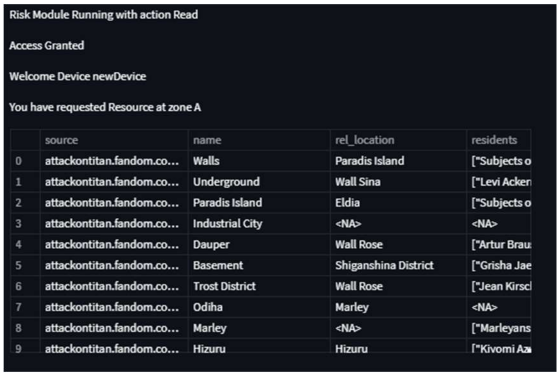
Password: ••••••••••

Resource Zone: zone A

User Action Severity: Read

Figure 21: Login page with Read action selected for Zone A

- **Reading Data:** On the login page, a new user is able to select which action to perform and what zone the resource is located. The zone information is used to cleanly separate different collections of data.



Risk Module Running with action Read

Access Granted

Welcome Device newDevice

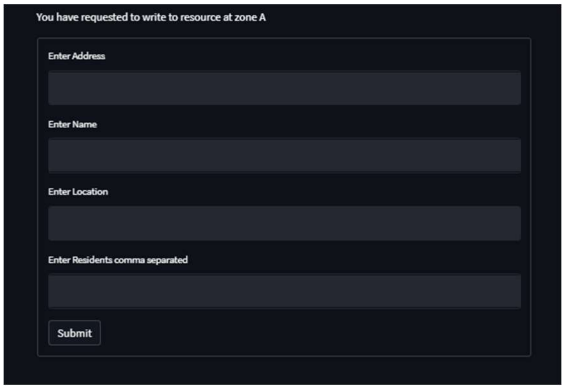
You have requested Resource at zone A

	source	name	rel_location	residents
0	attackontitan.fandom.co...	Walls	Paradis Island	["Subjects o
1	attackontitan.fandom.co...	Underground	Wall Sina	["Levi Acker
2	attackontitan.fandom.co...	Paradis Island	Eldia	["Subjects o
3	attackontitan.fandom.co...	Industrial City	<NA>	<NA>
4	attackontitan.fandom.co...	Dauper	Wall Rose	["Artur Brau
5	attackontitan.fandom.co...	Basement	Shiganshina District	["Grisha Jae
6	attackontitan.fandom.co...	Trost District	Wall Rose	["Jean Kirsc
7	attackontitan.fandom.co...	Odiha	Marley	<NA>
8	attackontitan.fandom.co...	Marley	<NA>	["Marleyans
9	attackontitan.fandom.co...	Hizuru	Hizuru	["Kivomi Az

Figure 22: Data Requested for Read action for Zone A

Upon clicking the request access button, the user is directed to the data available for this resource if they are granted access by the running risk estimator module and also if they do not violate any of the Chinese wall policy rules. The data is then released as a non-editable table of data as shown in Figure 22.

- **Creating Data:** Using the login page, the user can fill in their credentials, with the create action, they are granted access to a form if they do not violate the access control policy. The form is shown in Figure 23. New data can be filled in using the form, when the data is filled, it is stored in the database, and the pop up with the message Database updated is displayed. When the next "Read" user action is used in the database, the new entries can be seen as the last rows of the dataset.



You have requested to write to resource at zone A

Enter Address

Enter Name

Enter Location

Enter Residents comma separated

Figure 23: Entering new data into zone A

- **Updating and Deleting Data:** The logical flow for deleting and updating data is similar to reading and creating, as from the login page those actions can be specified and depending on the zone selected the name of the data entry to be updated or deleted is requested from the user. If access is granted, the entry is either deleted/updated from the database, with a pop message indication the success of the action.

4.5 Benchmarking

To evaluate the performance of the Adaptive Risk-based Access Control System (ad-RACs), we compared the precision, recall and F1-Score values of the MAN-SVMDT algorithm [62] against that of the was compared with the ad-RACs using the same 5-fold cross-validation method. The comparison results are shown in Table 2.

Table 2: Performance Comparison

Models	Precision	Recall	F1 Score
MAN-SVMDT	97%	98%	96%
Ad-RACs	95%	100%	98%

This study's evaluation revealed that, in comparison to other approaches, the performance of the ad-RACs is quite good. The adaptive Risk-based Access Control System has a better recall and f1 score values of 100% and 98% respectively. This means, the ad-RACs correctly identifies all positive cases (legitimate access) without any false negatives. This means that the system doesn't miss any legitimate access attempts. And also, indicates a high overall balance between precision (correctly granted access) and recall (minimizing false negatives) in the system's performance. However, in terms of precision the existing solution performed better at 97%. This means that, the existing system is better at correctly identifying and allowing legitimate access requests while minimizing false positives. Table 2 show that, the proposed adaptive Risk-based Access Control System performance is satisfactory.

4.6 Evaluation of the Adaptive Risk-Based Access Control System (ad-RACs)

The evaluation of the research highlights several key security implications as follows:

- **Access Control:** The research focuses on an adaptive risk-based access control system. This implies that the security of the system heavily relies on the accuracy and effectiveness of the risk estimation model. If the model fails to accurately assess the risk associated with granting access, it can result in unauthorized access or false positives/negatives, compromising the security of the system.
- **User Authentication:** The login page plays a crucial role in user authentication. If the login process is not adequately secured, it can lead to unauthorized individuals gaining access to the system. Therefore, the implementation of robust authentication mechanisms, such as strong passwords and secure login protocols, is essential to prevent unauthorized access.
- **Chinese Wall Policy:** The Chinese wall policy module is responsible for enforcing access restrictions based on conflict rules. If there are any vulnerabilities or weaknesses in the implementation of this module, it may allow users to access sensitive data from conflicting zones, violating the security policies in place.
- **Machine Learning Security:** The use of machine learning algorithms introduces security considerations. It is important to ensure the integrity and confidentiality of the trained models and their associated parameters. Unauthorized access to the machine learning models or manipulation of their inputs can lead to malicious activities or biased decision-making, compromising the security of the system.
- **Data Privacy:** As the system involves collecting and storing user data, data privacy is a significant concern. It is essential to implement appropriate measures to protect the confidentiality and integrity of the collected user data, such as encryption, access controls, and secure storage practices. Failure to adequately address data privacy can result in unauthorized access or data breaches.
- **System Monitoring:** System monitoring plays a crucial role in maintaining a secure access control environment by actively monitoring user behavior and access activities. Monitoring user behavior and maintaining a user risk history can help detect anomalies, suspicious activities, or potential security breaches. By focusing on the monitoring of user behavior, particularly the access decision, organizations can enhance their security measures and effectively detect and respond to potential threats.

4.7 Limitations of the Study

During the conducting of this research the following limitations were encountered. Firstly, there is a paucity of data in this area of research, which does not encourage the use of predictive models on the problems that currently exist. This could form a consequent research direction; for the data gathering for access control systems. Furthermore, this research was limited in the data used for user context. The user context information which forms one of the form inputs to the adaptive model only contains information about the user's password and login time, other useful data like the login location, device, login lag are not captured.

5. CONCLUSION

The research developed an Adaptive Risk-based Access Control System (ad-RACs) for improved access control at the network edge. The ad-RACs was developed, implemented and evaluated. The findings of the evaluation suggested that the ad-RACs can be implemented and used in situations where secure access to resources and performance are necessary.

The results showed that the ad-RACs was able to improve security by reducing the risk of unauthorized access and security breaches, increase flexibility by allowing access controls to be adjusted in real-time based on changing risk levels, and improve compliance by providing a way to control access to sensitive information and resources. Different machine learning models was evaluated for the purpose of risk assessment. The model used achieved a 0.06 (6%) for Brier Score, 0.98 (98%) for F1 score and 0.55 (55%) for AUCROC which cumulatively, outperformed all other models tested. Simulating an adaptive risk-based access control system using the evaluated model and a Chinese Wall policy showed it is able to take user input generate input features determining access status. Additionally, benchmarking against existing adaptive access control techniques showed that the ad-RACs had a better recall value at 100% and f1 score value at 98% as against existing system having a better precision value at 97%.

This research concluded that the developed ad-RACs performs better than existing access control system in terms of correctly identifies all legitimate access without any false negatives. This means that the system doesn't miss any legitimate access attempts. And also, indicates a high overall balance between correctly granted access and minimizing false negatives in the system's performance. It is therefore recommended that the model can be used as a reference for future research in the field of risk-based access control, and can help organizations improve the security of their resources while balancing business requirements and regulatory compliance.

Some recommendations are:

- A large part of the performance of a machine learning model is the quality and quantity of data available to it. As such, more data should be generated to bridge the gap of quality and quantity of data available in this field.
- In this study, the research scope did not cover user behavior monitoring as a way of getting better data concerning the user behavior while using the resource. This research can be improved by the development of a system that is able to collect the user behavior data for the purpose of updating the user risk history input feature.
- A further research direction not explored in this study is the factor interpretability plays in the use of machine learning models. In further research machine learning models can be used to evaluate risk and be ranked based on their interpretability. White box models; models which are easily interpretable will be given the higher ranking, while Black box models; models with low interpretability will be given a low rank. Better interpretability will improve AI governance and enable researchers and other interested parties easily identify bias with a machine learning model which may be disproportionately affecting certain kinds of users.

REFERENCES

[1] B. N. Silva, M. Khan and K. Han, "Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities.," *Sustain,* p. 697–713, 2018.

[2] R. Roman, J. Lopez and M. Mambo, "Mobile Edge Computing, Fog et al.: A Survey and Analysis of Security Threats and Challenges," *Future Generation Computer Systems,* pp. 1-31, 2016.

[3] M. Benantar, "Access Control Systems: Security, Identity Management and Trust Models," 2006.

[4] O. M. Al-Mendah and S. M. Alzahrani, "Cloud and Edge Computing Security Challenges, Demands, Known Threats, and Vulnerabilities," *Academic Journal of Research and Scientific Publishing,* vol. 2, no. 21, pp. 156-175, 2021.

[5] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher and V. Young, "Mobile Edge Computing: A key technology towards 5G," 2015. [Online]. Available: http://www.etsi.org/technologies-clusters/technologies/mobile-edge-computing.

[6] IEEE, "Real-Life Use Cases for Edge Computing," 2023. [Online]. Available: https://innovationatwork.ieee.org/real-life-edge-computing-use-cases/.

[7] V. Suhendra, "A Survey on Access Control Deployment," in *In Communications in Computer and Information Science*, vol. 259, Berlin/Heidelberg, Springer, 2011, pp. 11-20.

[8] N. Metoui, "Privacy-Aware Risk-Based Access Control Systems," *Ph.D. Thesis, University of Trento,* 2018.

[9] H. F. Atlam, M. A. Azad, M. O. Alassafi, A. A. Alshdadi and A. Alenezi, "Risk-Based Access Control Model: A Systematic Literature Review," *future internet,* pp. 1-23, 2020.

[10] N. Shi, T. Liang, C. Yang, C. He, J. Xu, Y. Lu and X. Hao, "BacS: A blockchain-based access control scheme in distributed internet of things.," *Peer-to-Peer Networking and Applications.,* vol. 14, no. 6, p. 2585–2599, 2021.

[11] Q. Wang and H. Jin, "Quantified risk-adaptive access control for patient privacy protection in health information systems," in *the 6th ACM*

Symposium on Information, Computer and Communications Security—ASIACCS '11, Hong Kong, China, 2011.

[12] T. Brooks, C. Caicedo and J. Park, "Security Vulnerability Analysis in Virtualized Computing Environments.," *International Journal of Intell.igent Computer Resources,* p. 263–277, 2012.

[13] U. P. Rao, P. Choksy and A. Chaurasia, "A Motive Towards Enforcement of Attribute-Based Access Control Models in Dynamic Environments," in *International Conference on Security, Privacy and Data Analytics ISPDA 2022*, Singapore, 2023.

[14] T. Galanc, W. Kołwzan, J. Pieronek and A. Skowronek-Grądziel, "Risk estimation and decision making in management (in selected areas of science)," *Operations Research and Decisions,* vol. 30, no. 1, pp. 46-66, 2020.

[15] M. Bräutigam, M. Bräutigam, M. M. Dacorogna and M. Kratz, "Predicting Risk with Risk Measures: An Empirical Study," *Social Science Research Network,* pp. 1-45, 2018.

[16] H. F. Atlam and G. Wills, "ANFIS for risk estimation in risk-based access control model for smart homes," *Multimedia Tools and Applications,* vol. 82, pp. 18269-18298, 2022.

[17] H. F. Atlam, M. A. Azad and N. F. Fadhel, "Efficient NFS Model for Risk Estimation in a Risk-Based Access Control Model," *Sensors,* vol. 22, no. 5, pp. 2005-2005, 2022.

[18] H. F. Atlam, A. Alenezi, R. J. Walters and G. Wills, "An Overview of Risk Estimation Techniques in Risk-based Access Control for the Internet of Things," in *2nd International Conference on Internet of Things, Big Data and Security. INSTICC*, 2017.

[19] N. N. Diep, L. X. Hung, Y. Zhung, S. Lee, Y. k. Lee and H. Lee, "Enforcing Access Control Using Risk Assessment.," in *Fourth European Conference on Universal Multiservice Networks (ECUMN'07)*, Toulouse, France, 2007.

[20] M. Loots, "Importance of a security policy," *South African Journal of Information management,* vol. 3, no. 2, pp. 1-22, 2001.

[21] B. Simon, "Security policy.," *Computers & Security,* vol. 9, no. 7, pp. 605-610, 1990.

[22] C. Wright, "Chapter 6 - Security Policy Overview," in *The IT Regulatory and Standards Compliance Handbook*, Syngress, 2008, pp. 115-147.

[23] A. Liu, X. Du, N. Wang, X. Wang, X. Wu and J. Zhou, "Implement Security Analysis of Access Control Policy Based on Constraint by SMT," in *2022 IEEE 5th International Conference on Electronics Technology (ICET)*, Chengdu, China, 2022.

[24] J. Dong and Q. Zhao, "Security access control policy of information system under multi-domain mode," *International Journal of Internet Protocol Technology,* vol. 11, no. 1, pp. 44-50, 2018.

[25] T. K. Dang, H. X. Son and L. K. Tran, "XACs-DyPol: Towards an XACML-based Access Control Model for Dynamic Security Policy.," *arXiv: Cryptography and Security,* pp. 1-12, 2020.

[26] P. Vasilikos, F. Nielson and H. R. Nielson, "Time dependent policy-based access control," *Informatics,* pp. 1-19, 2017.

[27] D. E. Kateb, N. Zannone, A. Moawad, P. Caire, G. Nain, T. Mouelhi and Y. L. Traon, "Conviviality-driven access control policy," *Requirements Engineering,* vol. 20, no. 4, pp. 363-382, 2015.

[28] S. Ahmad, S. Z. Z. Abidin, N. Omar and S. Reiff-Marganiec, "Managing access control policy from end user perspective in collaborative environment," in *COS 2014 - 2014 IEEE Conference on Open Systems*, 2015.

[29] T. Anish and K. Yogesh, "Machine Learning: An artificial intelligence methodology," *International Journal of Engineering and Computer Science,* 2013.

[30] S. Shalev-Shwartz and S. Ben-David, "Understanding Machine Learning From Theory to Algorithms," 2014.

[31] F. Osisanwo, J. Akinsola, O. Awodele, J. O. Hinmikaiye, O. Olakanmi and J. Akinjobi, "Supervised Machine Learning Algorithms: Classification and Comparison," *International Journal of Computer Trends and Technology (IJCTT),* vol. 48, no. 3, pp. 128-138, June 2017.

[32] C. Rich and N.-M. Alexandru, "An Empirical Comparison of Supervised Learning Algorithms," in *23rd international conference on Machine learning*, Pittsburgh, Pennsylvania, 2006.

[33] E. Alpaydin, "Introduction to machine learning," 2014.

[34] A. E. Mohamed, "Comparative Study of Four Supervised Machine Learning Techniques for Classification," *International Journal of Applied Science and Technology,* vol. 7, no. 2, pp. 5-18, 2017.

[35] Y. Ye, T. Li, D. Adjeroh and S. S. Iyengar, "A survey on malware detection using data mining techniques," *Association for Computing Machinery (ACM),*, pp. 1-40, 2017.

[36] I. Firdausi, A. Erwin and A. S. Nugroho, "Analysis of machine learning techniques used in behavior based malware detection," in *2nd International Conference on Advances in Computing, Control and Telecommunication Technologies (ACT).*, 2010.

[37] m. d. l. g. s. chandrahas, "credit card fraud detection using neural networks," *international journal of comoputer science,* vol. 4, no. 7, 2017.

[38] M. Krishna and D. Reshma, "Review On Fraud Detection Methods in Credit Card Transactions," in *2017 International Conference on Intelligent Computing and Control (I2C2'17)*, 2017.

[39] d. a. a. Nancy, "credit card fraud detection using svm and reduction of false alarms," *inyternation journal of innovations in engineering and technology,* 2016.

[40] I. Newsom, "Data Analysis II: Logistic Regression.," 2015. [Online]. Available: http://web.pdx.edu/~newsomj/da2/ho_logistic.pdf.

[41] J. Yashvi, T. Namrata, D. Shripriya and J. Sarika, "A Comparative Analysis of Various Credit Card Fraud Detection Techniques," *International Journal of Recent Technology and Engineering (IJRTE),* pp. 402-407, 2019.

[42] K. Goeschel, "Reducing false positives in intrusion detection systems using data-mining techniques utilizing support vector machines, decision trees and naive bayes for off-line analysis," *Institute of Electrical and Electronics Engineers (IEEE),* 2016.

[43] G. Meena and R. R. Choudhary, "A review paper on IDS classification using kdd 99 and nsl-kdd datasets in weka," in *International Conference on Computer, Communications and Electronics (Comptelix)*, Jaipur, 2017.

[44] L. Vanitha and D. M. Mary, "A comparative study of classification algorithms used in network intrusion detection systems (NIDS)," *ARS - Journal of Applied Research and Social Sciences,* vol. 3, no. 23, pp. 7-14, 2016.

[45] A. Riyad and M. I. Ahmed, "An ensemble classification approach for intrusion detection," *International Journal of Computer Applications,* vol. 80, no. 2, pp. 37-42, 2013.

[46] B. A. Tama and K.-H. Rhee, "A detailed analysis of classifier ensembles for intrusion detection in wireless network.," *Journal of Information Processing Systems,* vol. 13, no. 5, pp. 1203-1212, 2017.

[47] S. Revathi and A. Malathi, "Optimization of kddcup99 dataset for intrusion detection using hybrid swarm intelligence with random forest classifier," *International Journal of Advanced Research in Computer Science and Software Engineering,* vol. 3, no. 7, pp. 1382-1387, 2013.

[48] W. Koehrsen, "Random forest simple explanation," 2017. [Online]. Available: https://medium.com/@williamkoehrsen/random-forest-simple-explanation-377895a60d2d.

[49] J. Kim and S. Nepal, "A Cryptographically Enforced Access Control with a Flexible User Revocation on Untrusted Cloud Storage," *Data Science and Engineering,* vol. 1, no. 3, p. 149–160, 2016.

[50] P. Vimalachandran, H. Wang, Y. Zhang and G. Zhuo, "The Australian PCEHR System: Ensuring Privacy and Security through an Improved Access Control Mechanism," *EAI Endorsed Transactions on Scalable Information Systems,* vol. 3, no. 8, pp. 1-8, 2016.

[51] H. F. Atlam, R. J. Walters, G. B. Wills and J. Daniel, "Fuzzy Logic with Expert Judgment to Implement an Adaptive Risk-Based Access Control Model for IoT," *Mobile Networks and Applications,* p. 2545–2557, 2019.

[52] X. Ding, X. Jiang, H. Bi and J. Fang, "On the Access Control Mechanism of Wireless Sensor Network," pp. 52-62, 2017.

[53] L. Qiu, X. Sun and J. Xu, "Categorical quantum cryptography for access control in cloud computing," *Soft Computing,* vol. 22, p. 6363–6370, 2018.

[54] L. Cruz-Piris, D. Rivera, I. Marsa-Maestre, E. d. l. Hoz and J. R. Velasco, "Access Control

Mechanism for IoT Environments Based on Modelling Communication Procedures as Resources," *Sensors,* pp. 1-21, 2018.

[55] Y. Xu, W. Gao, Q. Zeng, G. Wang, J. Ren and Y. Zhang, "A Feasible Fuzzy-Extended Attribute-Based Access Control Technique," *Security and Communication Networks,* pp. 1-12, 2018.

[56] J. Ma, H. Xue, F. Wang, Y. An, D. Han, D. Wang, M. Zhao and S. Bi, "A Data Access Control Method Based on Blockchain," in *ISAIC 2020*, 2021.

[57] Q. Yang, M. Zhang, Y. Zhou, T. Wang, Z. Xia and B. Yang, "A Non-Interactive Attribute-Based Access Control Scheme by Blockchain for IoT," *Electronics,* pp. 1-11, 2021.

[58] W. Zhang and H. Yan, "A blockchain-based access control scheme for smart home," in *EEI 2021*, 2021.

[59] A. I. Abdi, F. E. Eassa, K. Jambi, K. Almarhabi and A. S. A.-M. AL-Ghamdi, "Blockchain Platforms and Access Control Classification for IoT Systems," *Symmetry,* pp. 1-17, 2020.

[60] Y. Jiang, H. Fu, A. Hu and W. Sun, "LoRa-Based Lightweight Secure Access Enhancement System," *Security and Communication Networks,* pp. 1-16, 2021.

[61] X. Liu, Y.-g. Zheng and X.-z. Li, "A revocable attribute-based access control system using blockchain," in *Journal of Physics: Conference Series: EEI 2021*, 2021.

[62] K. Yang, D. Li, L. Zhou and K. Cheng, "Research on Adaptive Dynamic Access Control Model Based on Blockchain and Token," *Journal of Physics: Conference Series,* pp. 1-8, 2022.

[63] T. Peretz, "Mastering The New Generation of Gradient Boosting," 2023. [Online]. Available: https://www.kdnuggets.com/2018/11/mastering-new-generation-gradient-boosting.html.

[64] Y. Lesley, "Tree Series 2: GBDT, Lightgbm, XGBoost, Catboost," 19 May 2018. [Online]. Available: https://yanpuli.github.io/posts/2018/05/blog-post-13/.

[65] L. Massaron, «Amazon Employee Access Challenge,» 04 August 2021. [Online]. Available: https://www.kaggle.com/datasets/lucamassaron/amazon-employee-access-challenge.