EAI Endorsed Transactions
on Pervasive Health and Technology
Research Article EAI.EU

Design of Machine Learning and Rule Based Access Control System with Respect to Adaptability and Genuineness of the Requester

Kriti Srivastava[1,*] and Narendra Shekokar[2]

[1]Research Scholar, Dwarkadas J Sanghvi College of Engineering, Mumbai, India.
[2]Professor, Computer Engineering, Dwarkadas J Sanghvi College of Engineering, Mumbai, India.

Abstract

INTRODUCTION:
Access control system (ACS) plays a major role in data security. It becomes more challenging for the system to provide accurate ACS, if data is huge and data requesters are not fixed. This is very predominant in the era of big data where new data are adding to the system very frequently. The main issue here is to justify adaptability in ACS.
OBJECTIVE:
The objective of this research is to have a comparative analysis of machine learning based access control methods with Rule based access control methods. Propose the most suitable method in detail.
METHODS:
Role based access control methods are highly robust and works effectively under known scenarios. We need additional methods to handle unknown scenarios. A decision-making method is used to identify the certainty of the rules and Mamdani fuzzy model is used to evaluate the situation based on current environmental factors. For machine learning based access control method Random Forest is used.
RESULTS:
Limitations of machine learning methods are discussed with respect to imbalanced data and bias in the algorithm. The proof of concept for rule-based access control method is tested for all the three modules involved in the framework. Certainty of the rules were accessed with the help of domain experts and accuracy of fuzzy rules were evaluated. Under critical conditions our framework was found to be accurate.
CONCLUSIONS:
Machine learning systems are not suitable for access control if they suffer with imbalance data problem. Rule based system are consistent and highly adaptable to unknown situations. Rule based systems have evaluated the genuineness of the requester based on sensitivity of information, time, location, previous history and emergency parameters.

Keywords: Modelling of Pervasive Healthcare Environments, Electronic Health Records, Identifying and addressing stakeholder needs, Security and Privacy Issues,

Received on 26 June 2020, accepted on 16 September 2020, published on 24 September 2020

doi: 10.4108/eai.24-9-2020.166359

*Corresponding author. Email: kriti.srivastava@djsce.ac.in

1. Introduction

Information security is a framework to prevent unauthorized access of data. Primary focus is to provide a balance between confidentiality, integrity and availability. Information access control mechanism can be implemented to avoid illegal access of information. This mechanism could be implemented in different ways. The first one is role-based access control, where each requester is assigned a role and each role is assigned different types of data access. Any requester will not be allowed to access any other data apart from the one assigned to its role. This mechanism is very strict and works very well with structured and known data. The second one is policy based or rule-based access control. Under this mechanism if the policy allows, a requester can access information of another role too. All of these are very effective if the data is structured and data access patterns are known to the administrator. They work very well in a static scenario. But if there is a case where suddenly the scenario changes and it is not stored in the knowledge base then role based, and rule based both will not work effectively. In such cases roles are needed to be assigned separately to the requesters and rules are explicitly required to be generated. To do this human intervention is required, which will reduce the performance and cannot be considered as completely adaptable. Ideally for a truly adaptable system, it should understand the situation and take decisions on its own. Hence there is a big need of completely Adaptable Access Control system.

Adaptability in the system can be added periodically due to change in environment or change in the existing scenario of the system [1]. Change in environment can occur due to political (suddenly government is changed so the policies can also be changed) or business (managing body changes the roles of employees), both have some kind of predictability. But if the scenario is of a completely new situation then first system will assess the risk component of changing the existing policies. The system should be able to verify the genuineness of the requester in the new situation, if satisfied, only then provide access of information to the requester. Finding genuineness of the requestor and taking appropriate decision, accordingly is the aim of this research work. This paper is a comparison of machine learning approaches with decision making and fuzzy rule generation approach.

2. Use Case

There are many use cases were the need of adaptable access control can be justified. For our research work we had considered the use case of a hospital management system (HMS). Users of the stored information in HMS were the doctors and the staff members of the hospital. Every patient's information was stored as a record. There were three types of information stored in the hospital about the patients: personal record, financial record or medical record. Every patient was assigned to one or more doctors. Only assigned doctors were allowed to access patients medical and personal records. Financial records were not accessible to the doctors under any circumstances. In a regular scenario a role-based access control could have been implemented and based on the roles, information access should have been provided.

These concept works very well under normal circumstances. There could be a possibility where a patient was critically ill at a very late hour. None of the assigned doctors were available. For quick treatment there was no time to perform all the tests and wait for the reports. Patient's record were there in the database and it could have been used. But since there was no role available in the system, who was assigned to the patient (no assigned doctor was available), his/her record was not accessible to anyone. If same situation was existing in a policy-based system and we would have tried to modify the policies based on the critical situation of the patient in this particular case, then there were chances of this new addition of rules being misused by some malicious requester in different cases. A malicious user can mislead the system and create an emergency situation to get the access of the information. Also, to make appropriate modification in policy base we need human intervention. A completely adaptable access control framework is developed which works efficiently if doctors are available, when patient is critical or in case of an epidemic or natural disaster where doctors are less, and patients are more to handle. Any genuine doctor should be allowed to access patient's information. This framework is a combination of role based, policy based and Intelligent systems.

3. Related Work

From many decades' researches had been proposing their work in the area of adaptability with existing access control methods. Role based access control is one of the oldest and quiet robust method. There are many modifications suggested on traditional Role based Access Control (RBAC). Bijol et.all had proposed a risk calculation method on traditional role-based method. As a proof of their concept they had compared their system with constraint-based risk migration and found their system to be more flexible [2]. Some researchers had integrated description logic with actions to add dynamicity to their work [3][4]. Researchers have emphasized that adaptability can be achieved by adding authorization transfer. One can explain authorization transfer is due to invoking services to the applications. This could be a reason of authorization transfer [5]. A logical representation of a five level dynamic access control model had been presented by Zhou, Ma and Wen in their work [6]. However, the properties of description logic were not illustrated. Researchers had proposed many probabilistic based risk calculations to be well informed about the kind of risk being involved in dynamically changing the access controls [7]. Yang and Liu had also proposed a two-step dynamic process to provide access control. [8] They had given a lot of emphasis on previous history and developed an adaptive algorithm. Kui Liu et. all had developed an attributed model which provides dynamicity in role-based systems with administrator's

intervention [9]. Some authors have used rewards and penalty method to finalize the risk in dynamically changing the authorities [10]. Recently authors had used markovian model for business process and established confidence level between good and harmful behavior [11].

Lot of research work emphasized more on policy-based access control for efficiency. Later authors had discussed modifications in existing policy-based access control system to provide adaptability and dynamicity. Petracca et.all had developed a runtime tracker which access four different types of risks. Two of them were related to access control policies and other two related to access control mechanism [12]. An extension of Attribute Based Access Control (ABAC) was proposed by Fugini, Hadjichristofi and Teimourijia, where they had included environmental factors in security decision process [13]. Notion of context was introduced to provide access to physical data. Introduction to cloud and other distributed framework had major challenge in implementing efficient access control system. Malik, Anwar and Shibli had worked on a cloud-based platform and introduced an adaptive framework of risk calculation, identifying risk and reconfiguring the system again [14]. Similar work had been done by Khushali shah et.all [15]. It was observed as the data size was increasing the demand for distributed data storage was increasing the complexity of implementing consistently adaptable access control. With increased amount of unpredictable data even adaptability was a big issue in authorizing new users. So, there was a need of efficient learning mechanism. There are quite a few research works using machine learning approach for access control. Pham, Albanese and Venkatesan had used SVM for multi agent-based classification which provided efficiency in the case of high dimensional dataset [15]. ANN-GA is also very effective combination to train the input parameters and do effective predictions [17]. Neural network is a very robust algorithm and is very effective in decision making [18]. There are many applications where assessment of access grant and deny needs to be learned appropriately. One such system which is very similar to risk adaptive access control is intrusion detection system. In fact, if the system is able to identify intruders then access grant and deny can be easily handled. Machine learning approach provides effective ways to identify intruders in the system [19] [20] [21]. Recently good amount of work has been done using deep learning [22] [23]. Many authors have applied CNN, SVM and nearest neighbor to improve predictions in an unbalanced dataset [24]. Studies show that Auto encoders are very efficient in unsupervised feature extraction and practically high dimension and huge dataset can be processed on GPUs [25] [26]. There are few authors who had used auto encoders to effectively find credit card fraud and anomaly detection [27] [28]. As mentioned in the related work that machine learning approach is suitable for anomaly detection and based on this idea, we can conclude whether the requester is genuine or malicious. Taking this conclusion further we will first discuss the limitation of machine learning approaches for access control methods in the next section.

4. Implementing Access Control using Machine Learning Approach

Concept of machine learning is to consider the evidence (data) provided, derive a hypothesis, test the performance and tune the parameters till accurate model for prediction is developed. Machine learning has shown great results in various domains. The basic idea behind machine learning algorithms are to understand various patterns in the dataset, learn them well and adjust all the parameters of the hypothesis which will be able to predict future data. This can be used to implement adaptability in a system. Using various parameters and different types of observations the model can learn about the system. The learned model will be able to take correct decisions for unseen data patterns. Hence adaptive access control can be implemented using machine learning model. The kind of evidence been used for hospital management system has data access patterns which includes doctor id, patient id, type of information requested, time of request, location of request, emergency and previous history. For each of these requests the class label is either access or deny. Hence the evidence in a HMS is a typical case of supervised learning dataset.

4.1. Access Decision Based on Neural Network

Out of many supervised learning algorithms neural network is the most robust algorithm. We wanted the system to be correctly adaptable. After performing neural network with different combinations of activation functions, different loss functions and number of iterations we found the one which gave best accuracy is also not having stability. Amongst 4 cases considered, three were using SGD (stochastic gradient descent) and another used Adams optimizer. Using two different datasets, this phenomenon is tested.

Table 1. Various cases under Neural Network.

Case	Time	No of epochs	Batch size	Error	Optimizer
1	2s/ epoch	30	1024	Mean squared	Adam
2	16s/ epoch	5	128	Mean squared	Adam
3	3s/ epoch	30	1024	Binary Cross entropy	Adam
4	2s/ epoch	30	1024	Mean squared	SGD

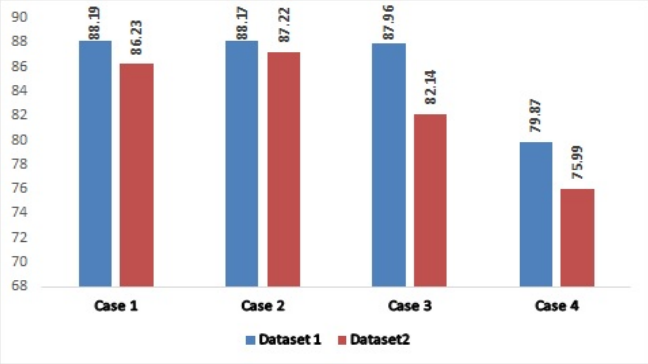

Figure 1. Accuracy measure of both the dataset for all the cases in percentage.

Dataset 1 is a hospital dataset from Scripps Mercy Hospital, San Diego, CA and dataset 2 is from kaggle. Dataset 1 has approximately 13 attributes and 22,232 observations. Dataset 2 has 19 attributes and 8, 58, 400 records. As shown in table 2, various combinations of neural network are being used to train data. Two different types of error functions mean squared error and binary cross entropy error are used to calculate the loss. SGD and ADAMS optimizer is used as activation function. Four different cases used in the research work are shown in table 1. Figure 1 is the accuracy plot for both the data sets. Lowest accuracy was found in case 4 which used SGD optimizer. Rest all used adams optimizer. In terms of accuracy case 2 provides best accuracy amongst all four cases but as shown in table 2, case 2 is leading towards over fitting. The training and validation errors are not matching.

Table 2. Over Fitting

Case	Hidden Layers	Training error	Validation error
1	2,3 nodes each	0.0929	0.0932
2	2,3 nodes each	0.0932	0.0935
3	2,3 nodes each	0.2987	0.3
4	2,3 nodes each	0.1615	0.1607

4.2. Autoencoder Based Autoencoder

Overfitting is a problem which tries to fit noisy data. In order to reduce this problem, those methods could be used which selects better attributes for prediction. Hence instead of directly applying classification first we tried using denoising auto encoder for selecting better input features. The idea behind using autoencoder is to reduce the effects of correlated attributes and use appropriate attributes for classification and prediction. Later on, the selected attributes we performed classification method such as random forest to classify and predict data.

Table 3. Various Combinations of Auto Encoder

Case	Hidden layers	Batch size	Iterations	Optimizer
1	1 layer, 19 nodes	1	2000	Adam
2	1 layer, 19 nodes	1024	2000	Adam
3	1 layer, 19 nodes	1024	20000	Adam
4	1 layer, 19 nodes	128	2000	Adam
5	3 layers, 9/3/9 nodes	128	2000	Adam
6	3 layers, 9/3/9 nodes	1	2000	Adam
7	3 layers, 9/3/9 nodes	1024	2000	Adam

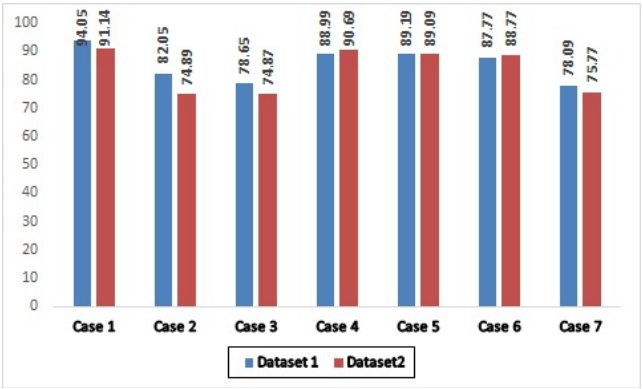

Figure 2. Accuracy graph for autoencoder based classification in percentage

Table 3 shows various combinations of autoencoders. Case 1 has shown better accuracy for both the datasets. Figure 2 shows corresponding accuracies for all the various cases analyzed. But when only random forest was applied for prediction, it showed better results than previous ones as shown in figure 3. To understand this phenomenon a small study on datasets were done and it is found that both the datasets have class imbalance problem. Dataset 1 has 60: 40 and dataset 2 has 80:20 ratios of each class. It had been proven that class imbalance problem can be cured by various

sampling techniques [29] [30]. Ensemble models had also shown great results with imbalance dataset [31]. Hence Random Forest which is one of the ensemble models gives best results for both the datasets.

Machine Learning approach is completely data driven. Prediction quality is affected by the completeness of data, type of attributes selected, noise in data, bias and many other factors. One needs to take care of all these aspects in order to develop a robust model to predict whether the requester is genuine or not. These models may be complex as well as time consuming. Machine learning approaches are completely data dependent. In order to do correct predictions, the models should be trained with correct and well-prepared data. If we are not sure about the correctness of data, then we may not develop machine learning models which will generalize well for predictions. In the next section we will discuss our approach which is not dependent on available data, but it utilizes the benefits of decision making and improvement on regular rule-based system.

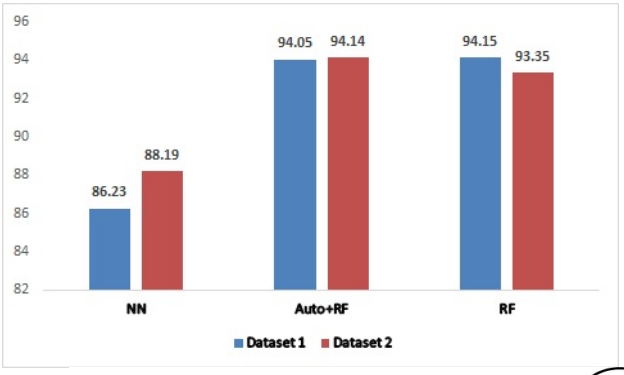

Figure 3. Comparison of Neural Network, Auto encoder with Random forest for classification and simple random forest for classification. (in %)

making methods which considers various parameters and ranks all the candidates. Decision making systems can be used for rule-based system to identify which rules are certain. One of the very popular decision-making method is Topsis. Parida and Sahoo [32] are first few authors who have discussed Topsis method and its usefulness in multi criteria decision making.

It is used in various fields to find certainty of a candidate. In [33] [34] authors have presented a simplified approach of topsis using triangular fuzzy membership for each candidate. In [35] authors have compared normal topsis and fuzzy topsis model. Büyüközkan et all [36] have discussed decision making methods using Topsis for various candidates of renewable energy. Fuzzy Topsis is also discussed in detail with applications in [37] [38]. They have applied fuzzy AHP method for weight calculation. This has shown reliable results

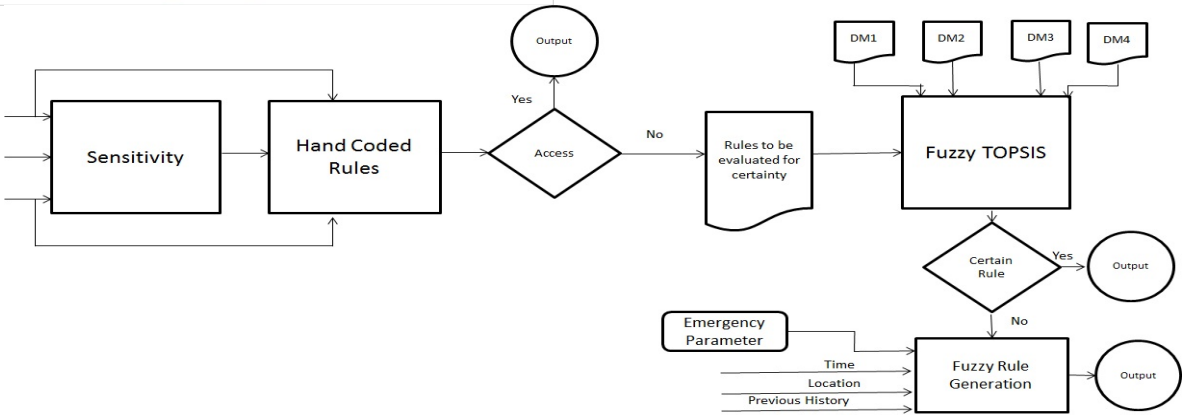

Figure 4. System Architecture

5. Implementing Access Control using Decision Making and Fuzzy Rule based

In the previous section we observed that machine learning approach is completely dependent on the kind of data we use for model training. This data ideally should be a true representation of the distribution but in reality, this is not possible due to error in data collection sources. Before training the model, lot of thoughts are given to preprocess the dataset. Many imputation techniques are also applied to it.

But learning in machine learning algorithm mostly happens through bias. So, there are high chances that machine learning models may not generalize well. There are many decision-making methods which considers various parameters and ranks all the candidates. Decision making systems can be Based on above mentioned findings we decided to apply a modified fuzzy rule-based access control instead of machine learning based access control.

The novelty of this work is to develop two phase rule analysis for providing adaptive access control. First set of rules are based on requesters role, patients' group and Sensitivity of the information. These rules are hand coded but their effectiveness is checked through a fuzzy decision-making system. All the rules which fall above a threshold are uncertain hence only those rules qualify for adaptability (possibility of change). Rules which prove to be certain cannot be changed by the system. System diagram is shown in figure 4. The request which falls under uncertain criteria go to the next phase. Since these are uncertain, we can verify with current parameters if making change in the rule is justified. Here system generates its own rules, based on the current scenario to decide whether to provide access or to deny.

5.1. Sensitivity of Information

As shown in figure 4, our use case, the hospital management system will take three types of inputs, doctor id, patient id and the information which the doctor wants to access. Based on the doctor id and patient id the system will understand the relationship of the doctor with the patient. This relationship can be of either of the three types: Assigned doctor to the patient, related doctor to the patient (anesthetic or junior doctor) or unrelated to a specific patient id.

In any access control system, we need to categories the assets based on their sensitivity. In this case also we needed to categorize the information. There are four main categories of information's used in HMS. First is Personal Information 1, which includes name, address, occupation, phone number etc. Second is Personal Information 2, which includes age, blood group etc. Third is financial Information such as PAN, adhar, payment mode used etc. Fourth is medical information such as diagnostic category, type of service etc. For the first three categories the administrator can decide the sensitivity level but for the fourth category we needed doctor's perspective also. So we conducted a survey on the medical information and their requirements with respect to a specific types of doctor (such as pediatric, gynecologist, nurse practitioner etc). The survey included pediatrician, gynecologist, urologist, surgeon, anesthetic, neurologist, general practitioner, ophthalmologist, gastro, nurse and ENT. Based on the survey results, a mapping of doctor specialty, diagnostic category and type of service was done. A sample is shown in the below table 4.

Table 4. Sample sensitivity calculation for medical data

Doctor Specialty	Diagnostic Category	Type of _Service	Sensitivity
Pediatrics	Newborns and neonates' conditions	Any	low
Pediatrics	Ear nose throat	Any except surgery	medium
Pediatrics	Ear nose throat	surgery	high

Under normal circumstances pediatric takes care of child below 12, they are also present during childbirth and give medicine to children for their illness. If any child needs any specific surgery such as a surgery for ear or nose, then he/she should be taken to an ENT. So, if a pediatrician wants to access information of a newborn then any type of service is of low sensitivity. If he wants to access information of ENT which excludes surgery, then it's of medium sensitivity. Else it will be of high sensitivity. This way mapping of the entire doctor specialty with the sensitivity level is done. Once this look up table 5 is created then based on the doctor id, patient id and Information, the system will set the sensitivity of the information.

For the medical information we needed to do a mapping of sensitivity assigned to the doctor category and sensitivity given by the doctors. As shown in the table low sensitivity to assigned doctor and low sensitivity to the information given by the doctor results in low sensitivity level for medical information category in the table. If low sensitivity and medium, then it results in medium sensitivity. Various combinations are shown in the table. This completed the first module of our research. With the output of this module hand coded rules are generated by the system administrator. According to these rules if the doctor id is assigned to patient id only then the information should be accessed by the doctor else deny the access. Our objective was to find under certain circumstances considering various factors can the deny rules be converted into access so that other genuine doctors can access patient's information and provide treatment? Most of the work in this area is actually done on finding if the doctor is genuine or not. Our work is on rule adaptation based on current circumstances, the need of treatment and access permission given to the doctor. In regular access control system even if the doctor was associated to the hospital, he/she may not be given access to any kind of information on regular basis. If there was a need and circumstances were adverse, then the hand coded rules would have been changed. Next two sections will discuss these topics in detail.

Table 5. Information Sensitivity

Doctor Category wrt D.id and P.id	Personal Information 1	Personal Information 2	Financial Information	Medical Information

Assigned doctor	L	L	H	L x L = L L x M=M L x H = H
Related doctor	H	M	H	M x L = M M x M = M M x H = H
Unrelated doctor	H	H	H	H x L = H H x M = M H x H = H

5.2. Need of Fuzzy Inference

In this research work one of the objectives was to decide if the rules created by the administrator can be modified under certain circumstances or not. There were many factors which influence this decision such as time, location, emergency, sensitivity of data and previous history. For rules each of these factors should have labels. Now if we consider time as 6pm, it can be a good time for a doctor whose OPD time is scheduled in this hour. If there was a doctor whose OPD time was till 5pm then 6pm time could have two views here. Some can say 6pm was a bad time for a doctor to request for a patient's information as OPD time is over. Another view was that there were chances that doctor, or patient came late for OPD hence doctor was checking the patient till 6pm. In second view 6 pm would have been labeled as good or ok. Similarly, for all the other factors there cannot be a fixed threshold value which can find a justifiable label. Decisions were based on relations and different people will have different opinion. Here use of fuzzy makes a lot of sense. In this research we had worked with triangular fuzzy membership values for all the parameters used for decision making. A number is represented as a triplet (a, b, c). Triangular fuzzy membership (equation 1) is given as:

$$\text{Triangle}(x; a, b, c) = \begin{cases} 0 & \text{If } x \le a \\ \dfrac{x-a}{b-a} & \text{If } a < x \le b \\ \dfrac{c-x}{c-b} & \text{If } b < x < c \\ 0 & \text{If } x \ge c \end{cases} \quad (1)$$

5.3. Decision Making for Uncertainty Identification

Inputs to the system were doctor id, patient id and information. Patient belongs to any of the three categories: Elite (celebrities and politicians), Regular or Donor (those who have signed up for organ donation). Based on patient id each doctor could be labeled as Assigned, Related (junior doctor, anesthetic etc) and Unrelated. The system will take the inputs and map the sensitivity level of each request made. This sensitivity level is based on the inputs given by various doctors and role mapping to the asset as discussed in earlier section.

We have discussed machine learning systems can be used for adaptable access policies but since it is highly dependent on data, we cannot be certain about the effectiveness of these models in adverse situations. We may not have enough variety of data samples to provide effective learning. In such case we need to switch to a model which is not data dependent and still adaptable. For this reason, we were back to original role based and rule-based system, as they had proven their effectiveness for static and regular situations. Using these traditional systems and convert them to adaptable systems without any previous data was the challenge of this work. Traditional rule-based system gave importance to roles and certain known conditions. These rules provide access to only those requesters who were assigned to the asset. For example, in HMS the doctors will only be provided access to his/her own patient's data. In this case no matter what the situation was, system will not allow accessing any other patient's records if the doctor was not assigned. This was the limitation of policy-based system also but in our system, it will be the first step of filtering the strict and correct rules for providing access. Based on the literature survey it is proven that existing rule-based systems highly depend on roles and assigned assets. Hence, they will allow only assigned requester to access information. There are 27 rules possible if relations of doctor, patient category and sensitivity are given. Sample is shown in table 6.

Table 6. Sample of Rules

Patient	**Doctor**	**Sensitivity**	**Access/ Deny**
Elite	Assigned	Medium	Access
Regular	Unrelated	Low	Deny
Elite	Related	High	Deny
Donor	Related	Low	Deny
Donor	Assigned	Medium	Access

Out of 27 rules there are 9 rules which have access as a decision label, so they are filtered out in the first step. For the rest 18 we have to identify adaptability. Our logic of applying adaptability will be to first identify, out of 18 rules which ones were more uncertain. Certainty means rules were well observed by the experts and modifying them under any circumstance shall not be allowed. Uncertain rules mean, even the experts were not sure about the decision of these

rules hence they can be modified. We have used Fuzzy Topsis Model for finding out uncertain rules.

Fuzzy topsis model uses the opinion of various experts for evaluation of uncertainty. We have considered four different categories of experts. First category of expert was from security domain. Second category were experts who had worked to enhance data availability in a hybrid structure with huge data. Third category is people who were biased toward hand coded rules (believe in traditional security system). Fourth category were patients view. Survey results were collected from all these categories. Category wise results were combined in four different decision matrixes. Using triangular fuzzy membership function these entire decision matrices were fuzzified. Seven varieties of linguistic scales were used.

Very poor (0,0,0.1),
Poor (0,1,3),
Medium poor (1,3,5),
Fair (3,5,7),
Medium Good (5, 7, 9),
Good (7, 9, 10)
Very Good (9, 10, 10).

Following the steps of Topsis model all the fuzzified decision matrixes were combined together using following logic.

$$x_{ij} = (a_{ij}, b_{ij}, c_{ij}) \text{ where} \quad (2)$$

$$a_{ij} = \min^k \{ a_{ij}^k \}; \quad b_{ij} = 1/k^k \quad c_{ij} = \max^k \{ c_{ij}^k \}$$

Next we have to find cost and beneficial attributes. Cost attributes means those parameters whose values should be minimized for optimization. Out of four parameters in decision matrix previous history and emergency are cost attributes because we want these linguistic values to be minimum. Time and Location parameters should be high so they are beneficial parameters. Next a normalized fuzzy decision matrix is created based on following equation

For beneficial

$$r_{ij} = \frac{a_{ij}}{c^*_i}, \frac{b_{ij}}{c^*_i}, \frac{c_{ij}}{c^*_i} \quad (3)$$

where $c_j^* = \max_i \{ c_{ij} \}$

For cost:

$$r'_{ij} = \frac{a'_{ij}}{c^*_i}, \frac{b'_{ij}}{c^*_i}, \frac{c'_{ij}}{c^*_i} \quad (4)$$

where $a'_j = \min^i \{ a_{ij} \}$

Next we created combined weight matrix. There are many methods which can be used for weight calculation, but we used experts view for weight calculation. We followed seven varieties for linguistic scale.
Very Low (0,0,0.1),
Low (0, 0.1, 0.3),
Medium Low (0.1, 0.3, 0.5),
Medium (0.3, 0.5, 0.7),
Medium High (0.5, 0.7, 0.9),
High (0.7, 0.9, 1.0),
Very High (0.9, 1.0, 1.0).

For each attribute,

$$\text{New weightage} = (w_{1j}, w_{2j}, w_{3j}) \quad (5)$$

Where $w_{1j} = \min^k \{w_{1j}^k\}$;
$w_{2j} = w_{2j}^k$;
$w_{3j} = \max^k \{w_{3j}^k\}$

After finding the weight matric we computed A*: Fuzzy Positive Ideal Solution (FPIS), which is the highest values in the matrix. A-: Fuzzy Negative Ideal Solution (FNIS), which is lowest values in the matrix. Next we computed the distance between FPIS and FNIS using fuzzy vertex theory.

$$d(x', y') = \sqrt{\tfrac{1}{3}[(a_1 - a_2)^2 + (b_1 - b_2)^2 + (c_1 - c_2)^2]} \quad (6)$$

We get two tables one representing FPIS and other showing FNIS. Calculate d* values by adding all the values in each row of FPIS table and d- values by adding all the values in each row of FNIS table. Finally, we calculate the closeness coefficient which decides the rank of each candidate.

$$CC_i = (d / (d + d^*) \quad (7)$$

All the 18 rules will have their CCj values. Low CCj value means certain and high CCi value means uncertain. Top few certain rules will be excluded so we excluded all the candidates whose CCj values are less than a threshold. Details are discussed in proof of concept. These certain rules should not be modified under any circumstances. Hence modifying these rules are out of scope of this work. Administrator can only decide about these. Rest all the rules whose CCj values are above threshold will be qualified for nest level of evaluation. Threshold values are decided by the experts in the system.

5.4. Rule Generation using Current Parameters

All the hand coded rules which qualify for next level of evaluation are tested with the current parameters. In this module system finds how critical is the situation based on current time, current location, emergency level and Previous History. In this module, rules will be generated on fuzzified parameters. With rule generation our system actually has a good blend of hand coded rules and system generated rules for final decision making. The output given by this module will be the final output. Current emergency parameters are calculated by TRIAGE criteria. If a patient's information is

asked then his/her SpO2, RR, GCS, BP (SYS), HR and temperature values are needed to the emergency module.

Table 7. Decision Matrix Given by Security Experts

Decision matrix 1	Security							
Rules	Patient	Doctor	Sensitivity	Output	Time	Location	Emergency	Previous history
4	Elite	Related	Low	Deny	h	h	h	h
5	Elite	Related	Medium	Deny	m	m	h	h
6	Elite	Related	High	Deny	h	h	h	h
7	Elite	Unrelated	Low	Deny	h	h	H	h
8	Elite	Unrelated	Medium	Deny	h	h	h	h
9	Elite	Unrelated	High	Deny	h	h	h	h
13	Regular	Related	Low	Deny	m	m	h	h
14	Regular	Related	Medium	Deny	h	h	h	h
15	Regular	Related	High	Deny	h	h	h	h
16	Regular	Unrelated	Low	Deny	h	h	h	h
17	Regular	Unrelated	Medium	Deny	h	h	H	h
18	Regular	Unrelated	High	Deny	h	h	h	h
22	Donor	Related	Low	Access	m	m	m	m
23	Donor	Related	Medium	Deny	h	h	h	h
24	Donor	Related	High	Deny	h	h	h	m
25	Donor	Unrelated	Low	Deny	h	h	h	H
26	Donor	Unrelated	Medium	Deny	h	h	H	H
27	Donor	Unrelated	High	Deny	h	h	h	h

Table 8. Decision Matrix Given by Big Data Experts

Decision Matrix 2	HPC							
Rules	Patient	Doctor	Sensitivity	Output	Time	Location	Emergency	Previous History
4	Elite	Related	Low	Deny	m	m	m	m
5	Elite	Related	Medium	Deny	m	m	h	h
6	Elite	Related	High	Deny	h	h	h	h
7	Elite	Unrelated	Low	Deny	h	h	H	h
8	Elite	Unrelated	Medium	Deny	h	h	h	h
9	Elite	Unrelated	High	Deny	h	h	h	h
13	Regular	Related	Low	Deny	l	l	m	m
14	Regular	Related	Medium	Deny	m	m	h	h
15	Regular	Related	High	Deny	m	m	h	h
16	Regular	Unrelated	Low	Deny	M	m	h	h
17	Regular	Unrelated	Medium	Deny	h	h	H	h
18	Regular	Unrelated	High	Deny	h	h	h	h
22	Donor	Related	Low	Access	l	l	l	l
23	Donor	Related	Medium	Deny	l	l	m	m
24	Donor	Related	High	Deny	m	m	h	h
25	Donor	Unrelated	Low	Deny	h	h	h	h
26	Donor	Unrelated	Medium	Deny	h	h	H	h
27	Donor	Unrelated	High	Deny	h	h	h	h

This module will then categorize the patient's condition as emergency or non-emergency. If the case is an emergency case then there are three categories, whether 5mins or 15mins or 1 hour is left for the patient to survive. 5mins and 15 mins are really critical hence in such cases, other parameters (time and location) should have low importance. For 1 hour emergency we give equal importance to all the parameters as assigned doctor can be contacted during this period. For non-emergency it follows role-based model. All the four parameters are fuzzyfied using the same seven values linguistic scale used above for experts. These values are fed to mamdani model. For all possible combinations of parameters, fuzzy values are calculated.

6. Proof of Concept

There are two main modules in this work. The fuzzy topsis module takes all the hand coded rules which have Deny in the decision as inputs along with four different decision matrices by the experts. Samples of decision matrix given by the experts. As shown in the table 7 and 8 all the survey results were aggregated based on various category. Experts have

given their views to all the parameters influencing the access rules. In these modules all these expert views were used to find certain and uncertain rules. Fuzzy Topsis calculates closeness coefficients for each of the rule. More the value of closeness coefficient more uncertain is the rules.

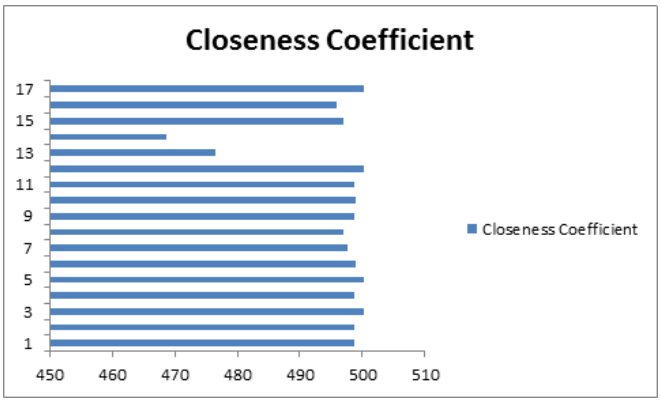

Figure 5. Closeness Coefficient

As shown in the figure 5, rule number 13, 14 and 16 are having less value of closeness coefficient. Here we have to make a wise decision on the threshold value as we will be going to filter out hand coded certain rules. Analyzing these three rules we see that rule 13 and 14 are for related doctor asking regular patients low and medium sensitivity data. Rule 16 is for unrelated doctor asking regular patients low sensitivity data. If we increase the closeness coefficient rule number 15 is selected. If we select this rule, then there will be no rule selected for the combination of related doctor and regular patient. Hence, we decided to keep the threshold value as 0.495932. All the rules with closeness coefficient above this value will be selected for next level of evaluation. This is a one-time activity so expert intervention will not delay the decisions.

The next module takes all the current parameters such as time, location, emergency and previous history. Triangular fuzzy membership function is applied on all these four parameters and using Mamdani model rules will be generated. There are four parameters and each parameter have three different values so rules will be possible. Output of Mamdani model is a fuzzy value so we need to defuzzify them. Output ranges between 0 to 1and an appropriate alpha value will decide whether the answer is Access or Deny. For this work alpha value is 0.3. So, all the values greater than 0.3, will be Deny. Out of 81 generated rules based on the alpha value 21 were Access and 57 Deny shown in figure 6

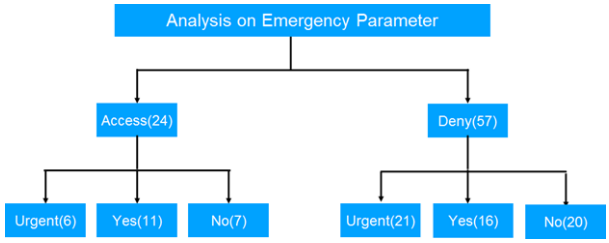

Figure 6. Analysis of Emergency Parameter

Out of four parameters Emergency is the most important parameter, shown in the figure. Performing analysis on emergency parameter we found that out of 24 Access, 6 were Urgent cases, and 11 were Yes (emergency case can wait till 1 hour). Both these cases are genuine, so Access is justified. But there are 7 rules which give Access to a no emergency case. This cannot be justified under any circumstances. There are 57 generated rules that are labeled as Deny. Here 20 cases of no emergency are denied which is completely justified. But 21 Urgent and 16 Yes cases are also denied. We need to analyze each rule under these categories. Table 9 shows Time, Location and Previous History parameter values for 21 Urgent cases which were denied. There are five cases where all the three parameters were good or ok, still under urgent emergency circumstance access were denied. Our model is actually not giving correct results for these cases.

Table 9. Analysis on Urgent cases for Deny

Time	Location	Previous History
Good	Good	Good
Good	Good	OK
Good	OK	Good
Good	OK	OK
Good	OK	Bad
Good	Bad	OK
Good	Bad	Bad
OK	Good	Good
Good	OK	Bad
OK	OK	Good
OK	OK	OK
OK	OK	Bad
OK	Bad	OK
Bad	Good	Good
Bad	Good	OK
Bad	OK	Good
Bad	OK	OK
Bad	OK	Bad
Bad	OK	Bad
Bad	Bad	OK
OK	Bad	Bad

Similarly, for 16 emergency cases which were Yes (1hr emergency) were also denied. Analyzing other parameters in the following table 10, we see that there is one case where all the parameters values are either good or ok. Hence aggregating all the wrong results, we get 16% error rate. When we compare our model's accuracy (84%) with machine learning approaches (approximately 90%), it appears to be less by approximately 6%. But machine learning approaches have just relied on data. They assume that data is accurate and the learning about the system will be correct. Also, no

consideration is given in the machine learning models about the change in the system. In our system we have included expert opinion, traditional rule-based system and fuzzy inference for access control system. Combining all these gives more confidence to our decision as compared to machine learning approaches.

Table 10. Analysis for Yes Cases for Deny

Time	Location	Previous History
Good	Good	OK
Good	Good	Bad
Good	OK	OK
Good	OK	OK
Good	Bad	OK
Good	Bad	Bad
OK	Good	Bad
OK	OK	Bad
OK	Bad	Good
OK	Bad	Good
Bad	Good	Bad
Bad	OK	OK
Bad	Good	Bad
Bad	Bad	Good
Bad	Bad	OK
Bad	Bad	Bad

7. Conclusion and Future Scope

Implementing adaptable access control needs a good balance of security as well as data availability. As we have discussed in the use case there may be an urgent need to provide access to an unauthorized user. It becomes challenging for the system to understand the genuinity of the requester. In this paper we had discussed a two-step methodology to decide whether to provide access to a requester. Our method takes the accuracy of existing Rule based system and adds environmental factors, analyze current risk and then takes final decision. Our method emphasizes on considering fuzzy parameters as in every expert have different views on various parameters. Fuzzy parameters increase the horizon of decision making. Our method gives more confidence to the decision. On evaluating the proof of concept, the error was measured to be 16%. Currently we are working on adaptable access control on a distributed framework. Finding a robust as well as adaptable access control in hybrid data layer is our current work.

References

[1] Bassam Farroha and Deborah Farroha, "Challenges of "operationalizing," dynamic system access control: Transitioning from ABAC to RAdAC, IEEE International Systems Conference SysCon, 2012, pp 1-7.

[2] Khalid Zaman Bijon, Ram Krishnan and Ravi Sandhu, "A Framework for Risk- Aware Role Based Access Control," 6th Symposium on Security Analytics and Automation, 2013, pp 462- 469.

[3] L. Ma, S. Ma, J. Lv and Y. Sui," A Dynamic Description Logic-Based Formalism for RBAC" in proceedings of Computer Science and Convergence of Information Technology, 2009, pp 970- 975

[4] M. Uddin, S. Islam and A. Al-Nemrat, "A Dynamic Access Control Model Using Authorising Workflow and Task-Role-Based Access Control," in IEEE Access, vol. 7, pp. 166676-166689, 2019.

[5] Zaharn M, "Non Inclusion Property in Multilevel Caches Revisited, "International Journal of Computers and Their Applications, 2007, 14(2): 99-108.

[6] Zhou Y, Li M and Min W, "A Multi-level Dynamic Access Control Model and Its Formalization" 2nd International Conference on Information Science and Control Engineering, 2015, pp: 23-27

[7] Qihua W and Hongxia J, " Quantified risk-adaptive access control for patient privacy protection in health information systems," ASIACCS 6th ACM Symposium on Information, Computer and Communications Security, 2011, PP 406-410

[8] Y. Yang and S. Liu, " Research on the qualification method of the operational need based on access purpose and exponential smoothing," IEEE 7th Joint International Information Technology and Artificial Intelligence Conference, 2014 pp 516-522

[9] Kui Lui, Zhurong Zhou, Quianguo Chen and Xiaoli Yang, "Towards an attributed authorization model with task role based control for WFMS," IEEE 2nd International Conference on Communication Technology (ICCT), 2015, pp: 361 – 371.

[10] Riaz Ahmed Shaikh ; Kamel Adi ; Luigi Logrippo ; Serge Mankovski, "Risk-based decision method for access control systems," IEEE 9th International Conference on Privacy Security and Trust, 2011, pp; 584-588

[11] Carlos Eduardo da Silva, Jos´e Diego Saraiva da Silva, Colin Patersonand Radu Calinescu,"Self-Adaptive Role-Based Access Control for Business Processes," 2017 IEEE/ACM 12th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS) pp: 193-203

[12] Giuseppe Petracca, Frank Capobianco, Christian Skalka and Trent Jaeger, "On Risk in Access Control Enforcement," Proceedings of the 22nd ACM on Symposium on Access Control Models and Technologies, SACMAT 2017, pp: 31-42

[13] Mariagrazia Fugini, George Hadjichristofi and Mahsa Teimourijia, " Dynamic Security Modelling in Risk Management Using Environmental Knowledge," 23rd International WETICE Conference, 2014, pp 429-434.

[14] A. A. Malik, H Anwar and M. A. Shibli,"Self Adaptive Access Control and Delegation in Cloud Computing,"17thb IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, China 2016, pp 1-8

[15] Khushali Shah, Priya Shah, Kriti Srivastava and Narendra Shekokar, "Dynamic Access Control in a Document Data Store," International Journal in Advance Research in Computer Science and Software Engineering (IJARCSSE), Vol 7, Issue 5, May 2017, ISSN: 2277 128X, pp: 518-522

[16] luan Huy Pham, Massimiliano Albanese and Sridhar Venkatesan, "A quantitative risk assessment framework for adaptive intrusion detection i n cloud," security and privacy SPC 2016, pp. 489-497

[17] J Ponni and K. L. Shunmuganathan, "Multi-agent system for data classification from data mining using SVM," 5th Green Computing Communication Conservation Energy IEEE ICGCE 2013, pp. 828-832

[18] Baris Yuce and Yacine Rezgui, "An ANN-GA semantic rule based system to reduce the gap betweem predicted and actual energy consumption in buildings," IEEE Trans on Automation Science and Engineering, vol 14, no 3, pp. 1351- 1363, july 2017.
[19] R Seteono, B Baesens and C Mues, "Recurssive neural network rule extraction for data with mixed attributes," IEEE Trans in Neural Network, vil 19, no 2, pp. 299-307, february 2008.
[20] Longjie Li, Yang Yu, Shenshen Bai, Yinj Hou, Xiaoyun Chen, "An effective two step intrusion detection approach based on binary classification and kNN," IEEE Access vol 6 2018, pp. 12060-12073.
[21] Chie Hong Lee, Yann Yean Su, Yu-Chun Lin, Shie –Jue Lee, "Machine Learning based network intrusion detection," 2nd IEEE International Conference on Computational Intelligence and Application ICCIA 2017, pp. 79-83.
[22] Gunupudi Rajesh Kumar, Nimmala Mangathayaru, Gugulothu Narasimha, Gali Suresh Reddy," Evolutionary approach for Intrusion detection," International Conference on Engineering and MIS (ICEMIS), 2017, pp. 1-6.
[23] Y LeCun, Y. Bengio and G. Hinton, " Deep Learning," Nature, vol 521, no 7553, pp. 436-444, May 2015 [Online] Available: http://www. Nature.com/doifiner/10.1038/nature14539
[24] I. Goodfellow, Y Bengio and A Courville, "Deep Learning," 2016 [Online]. Available: http://www.deeplearningbook.org
[25] [28] M.M.U. Chawdhury, F Hammond, G Konowiz, C Xin and H. W. Li,"A few shot deep learning approach for improved intrusion detection," 8th Annual Ubiquitous Computing Electronic and Mobile Communication Conference (UEMCON) 2017, pp. 456-462.
[26] N Shone, T. N. Ngoc, V. D. Phai, Q Shi," A deep learning approach to network intrusion detection," IEEE Trans on Emerging Topic in Computational Intelligence, 2018, vol 2, issue 1, pp. 41-50.
[27] F Farahnakian and J Heikkonen, " A deep autoencoder based approach for intrusion detection system," 20th International conference on Advance Communication technology ICACT, 2018, pp. 178-183.
[28] G Rushin, C Stancil, M Sun, S Adams and P Beling," Horse Race Analysis in Credit Card fraud –deep learning, logistic regression and gradient boosted tree," IEEE SIEDS 2017, pp. 117-121.
[29] Sukara Barua, Md. Monirul Islam, Xin Yao and Kazuyuki Murase," MWMOTE- Majority Weighted Minority Oversampling Techniques for Imbalanced Data Set Learning," Vol 26, No 2, Frb 2014, pp 405-425.
[30] Sachin Subhash Patil and Shefali Pratap Sonavane," Enriched oversampling Techniques for Improving Classification of Imbalanced Big Data," IEEE 3rd International Conference on Big Data Computer Science and application, 2017.
[31] Bashir Elkarami, Abed Alkhateeb and Luis Rueda, "Cost-Sensitive Classification on Class- Imbalanced Ensembles for Imbalanced Non-coding RNA Data," IEEE EMBA International Student Conference, 28th -31st May 2016.
[32] P. K. Parida and S. K. Sahoo, " Multiple attribute Decision Making Approach by Topsis Technique," International Journal of Engineering Research and Technology (IJERT), vol 2, Issue 11, Nov 2013, pp 907-912.
[33] Balwinder Sodhi and Prabhakar T. V, " A Simplified Description of Fuzzy TOPSIS," Cornell University, arXiv: 1205.5098v2, jun 2017
[34] Sorin Nadaban, Simona Dzitacb and Ioan Dzitaca, "Fuzzy TOPSIS: A General View," Information Technology and Quantitative Management (ITQM), 2016, pp: 823-832
[35] R. Dharamraja and C. Sharmila Mary, "The Evaluation of Topsis and Fuzzy Topsis Method for Decision Making in Data Minig," International Research Journal of Engineering and Technology (IRJET), Vol 3, Issue 9, Sep 2016
[36] G. Büyüközkan and S. Güleryaz, "Fuzzy Multi Criteria Decision Making Approach for Evaluating Sustainable Energy Technology Alternatives," International Journal of Renewable Energy Sources, vol 1, 2016
[37] S. Chen, S. Cheng and T. Lan, "A new multicriteria decision making method based on the topsis method and similarity measures between intuitionistic fuzzy sets," 2016 International Conference on Machine Learning and Cybernetics (ICMLC), Jeju, 2016, pp. 692-696.
[38] S. Zhou, W. Chang, S. Zhou and W. Liu, "The method of risk evaluation for equipment development based on triangular fuzzy number and TOPSIS," The 26th Chinese Control and Decision Conference (2014 CCDC), Changsha, 2014, pp. 2272-2276.