Chapter 4

4. Designing of Framework

Usually organization adopts an authorization/access scheme that provides flexibility for sharing of records to a wide variety of authorized users. Variability in data ownership, custody, patient consent, established authorizations, compatibility of schemas, compliance with standards [72] [76] across hospitals, add to the complexities in sharing of data. To achieve desired but secured sharing of data, regulatory bodies such as HIPAA had proposed rules in support of patient consent, disclosure of patient’s rights and authorizations, to protect the privacy and confidentiality of EHRs. HIPAA security rules specify integrity controls, encryption [25] (164.312(e)(1)) to address transmission security and access authorization (164.308(a)(3)), modification deals with access security. Transmission security addresses the threats and attacks on the network path when the sensitive data travels from sender to receiver. Access security caters to the accessibility of data to the authorized person only. NIST SP 800-53 and 53A7 recommends security controls for federal information systems. All health agencies and professionals are legally bound to follow one or other such guidelines as prevalent in their states or countries. The complexity associated with sharing of EHR is further enhanced by global perspective as different country adopt with different by-laws.

Each healthcare organization implements an access model defining authorized access to its users. A lot of human interaction is involved in the process of securing legitimate disclosure of EHRs. It increases challenges of uncertainty and approximations in ascertaining access rights in interoperable environment. To design a framework for achieving secured data sharing, one needs to find most suitable access control model such that it enhances compatibility and portability of data between disparate healthcare organizations.

4.1 Access Control Models

Access control is a controller that permits access to the requested resource only after verifying the set rules and conditions in the access policies of an organization. The request that does not pass rules and conditions are rejected. Access control mechanism is customized as per the need and policies of every organization. Integration of inputs coming from varied access control mechanisms results in policy conflicts and breach of security usually when
organizations wish to share their resources. This problem can be resolved by writing access control policies in a manner that the organizations requesting or sending data can evaluate them on the basis of defined or set attributes, properties and conditions, both statically and dynamically. Traditional access control models (ACMs) listed as Mandatory ACMs, Discretionary ACMs and Role-based ACMs are studied to identify their suitability and limitations in resolving the above problem.

**Mandatory Access Control Model (MAC):** It is a mechanism where the control to manage the data lies with the system and users have no rights and authorities to alter the defined access. MAC enforces access control mechanisms with the help of sensitivity levels assigned to the subjects and resources in the system. The sensitivity level of subjects reflect the level of trust assigned in accessing the system, whereas sensitivity level of resources is marked on the basis of content requiring different level of sensitivity while being accessed. The level for read and write access varies between low to high levels. In other words, access policies are defined such that a subject with low sensitivity level cannot read a resource with high sensitivity level and cannot write to a resource whose sensitivity level is lower than the subject’s sensitivity level. MAC is associated and influenced with Bell-LaPadula Confidentiality Model (BLP) [142]and Biba Integrity Model[143]. The model exhibit to handle the confidentiality requirements by mandating the policy that high level data cannot flow to a lower level user. BLP model verifies and restrict the information flow from higher to lower security levels using two properties- simple security property and *-property. MAC model is mainly used by the military applications. It is a straight forward model and is highly suitable in the high risk environment and the primary objective is to ensure the confidentiality of subjects and objects. On the downside, the model is static and fails to handle dynamic alterations of given policies. It is based on the trusted assumptions that the execution code behind the concerned objects is correct which makes it unsuitable for implementing multi-level security, the need of today’s time.

**Discretionary Access Control Model (DAC):** Unlike MAC, DAC allows the users to decide access rights and privileges on the resource, by themselves. A user has the authority to share the resources under his/her ownership to other users. Access demanded for the resource is based on evaluation of the identity of the user or the group to which the user belongs. DAC makes use of access control list (ACL) [164], a structure that contains zero or more access control entries with respect to subject and objects that the subject can access. More precisely,
the access control list specifies what has to be done with an object. DAC enables fine-graining of the policies and is considered most cost-effective in its implementation. Considering the limitation of DAC, lacks of constraints and control of access rights owned by the user itself make it difficult to maintain and verify the integrity of the system.

**Role-based Access Control Model (RBAC):** A widely accepted model over DAC and MAC is RBAC, proposed by Ferraiolo[144]. RBAC exhibit centralized control in providing access to the resource and claim to reduce the complexities and cost of administering the systems. It ensures that the authorized users are provided with restricted and relevant access of specific data in a system. RBAC introduced the notions of role and privilege. Specifically, role is the user’s position in an organization and the privilege is the permissions associated with the role based on the user’s responsibilities and activities. A user can be assigned more than one role and similarly, a role can have more than one user. The model ensures system integrity and availability by explicitly controlling the access of the resources to RBAC has been rigorously explored in the past and many variations and enhancements [145] have been proposed such as role graph model by Nyanchama and Osborn [146], coalition based access control [147], RBAC96 family of models by Sandhu et al.[153]. Further, RBAC has been implemented by various researchers in the extended forms like GEO-RBAC (Geographical RBAC), ARBAC (Administrative RBAC’97), OrBAC (Organizational RBAC), with each form adding or modifying new algorithms to satisfy the need of its time. Customization of RBAC and fine-graining the policy as per the administrative requirements of the concerned domain is very difficult and requires designing of more sophisticated and specific models.

**RBAC in Healthcare Environment**

Handling conditional constraints and a separation of roles in interactive sharable healthcare environments is a limitation[145] in traditional RBAC models. A Business Process Driven – Access Control Service framework [148] chooses an appropriate access control model keeping in view the business requirements and restrictions. The framework lacks the inclusion of various constraints associated with the role and privilege assignments. A concept to categorize permissions into subspaces having meaning semantics is developed by Jaeger et al.[171], with an objective to reduce conflicts clarify authorized permission assignments. With each health care professional administering multiple roles and responsibilities, it becomes insufficient to rely on statically defined roles and permissions. Additionally, health care users experience a frequency shift in roles and permissions controlled by different
constraints. Multi-OrBAC, an extension of RBAC though reduced the management complexities using Logical Programming with Constraints (LPC), but lacks to incorporate dynamism in applying access control policies in interoperable sharing of data. Reaching to a logical decision for permitting or denying sharing of data well secured against any illegal disclosures is a big challenge.

NIST [107] proposed U.S. National Standard for RBAC through International Committee for Information Technology Standards (INCITS). The standard comprises of four components-Core-RBAC, Hierarchical-RBAC, Static and Dynamic Separation of Duty (SoD) that makes RBAC as a de facto standard in organizations with large number of users and resources. The model assigns data access rights to the users on the basis of users and resource hierarchies laid down by an organization. It confines owners or users into roles and then assigns relative permissions on the resources. Integration and verification of the rules to determine authority and constraint-based access to the health records ensure secured sharing in an interoperable healthcare environment. Satisfying constraints and restricting irrelevant access requires an access model that can distinguish roles unified with access rights bonded under contextual constraints. It requires further checking to prevent the leakage of unauthorized access to the legitimate users. Hierarchical-RBAC [144] structures the rules to reflect the authorization and accountability of the users in the most natural manner.

Access control is a framework that controls and manages how the users would access the data. IT team has a huge challenge designing a balanced and secured system suitable for entire strata of healthcare. In many cases, the newer, more complicated models arose to overcome deficiencies [145][147] in the security of existing system. Further there is often need for new models to address changes in technologies [148], organizational structures [149], organizational needs, technical capabilities, and/or organizational relationships. Many access models [70] [150] [151] namely, Attribute-based, Purpose-based (PBAC), Hierarchical Role-based (RBAC), Identity-based (IBAC), Collaboration-based etc. have been rigorously explored by the researchers over last many decades.

**Attribute-Based Access Control:** Attribute based access models [152] determines authorizations by evaluating attributes of associated subjects, objects and operations for making access decisions. Attributes in heterogeneous systems grow exponentially. Managing the access and security of data based on these attributes tends to be complex and expensive.
**Purpose-Based Access Model:** PBAC [156] is based on the notion of associating data objects with aims. PBAC provides greater privacy preservation by allocating objects with purposes [157][158]. An access control decision can be made for enabling sharing between disparate organizations by decomposing a global policy [159] without losing the sensitive information.

**Hierarchical Role-based Access Model:** RBAC[153] models use consents and rights based on the assigned roles in groups/institutions to limit access. A user has an access to the resource according to the assigned role. Role hierarchies [150] are a natural way of organizing roles to reflect authority. In healthcare domain it is common to find overlapping of roles. Hierarchical RBAC [154] establishes roles in such overlapping operations. However, RBAC cannot integrate other access parameters or related data that are significant in allowing access to the user[155].

**Identity-Based Access Model:** Identity-based access model includes Access Control Lists (ACLs) for implementing Discretionary Access Control (DAC) policies. DAC restricts access to objects based on the identity of the subjects (end-users), and/or groups to which they belong. However, in DAC, granting read access is transitive and the policies are helpless for Trojan Horse Attack [150][160]. Mandatory Access Control (MAC) policy can prevent the Trojan Horse that occurs in DAC. MAC is based on access control policy decisions, made by a central authority [150] [161]. In MAC, the individual owner of an object has no right to control the access. Thus, MAC policy fails to preserve the privacy requirement [161] for EHRs of the patients.

**Coalition Based Access Model:** CBAC model [147] is highly expressive, provides broad functionality and flexibility in collaborative environments. It overcomes the limitations of RBAC through the inclusion of teams and tasks entities. Coalitions do not include a top-level organization in which ultimate authority is vested. Hence, distributed administrative authorities must exist in such models so that distribution can be negotiated. It is possible that the collaborative organizations may follow hierarchical administrative models internally.

**Constraint-Based Access Model:** Security constraints are restrictions that influence the decision to allow or deny access to the resource while sharing them in between independent users. The Constraint-based access models [162] provide clear distinction of possible relationship between two organizations and thus enable better integration of two distinct systems.

However, administering the most suitable access control is highly difficult when applied in integrated environments. The authors [163] combine three existing access control models and
present a novel access control model for EHRs which satisfies the requirements of EHRs but the processes are more complex to implement. The choice of an access control model is governed by various factors [148] such as administrative convenience, policy-support capabilities and access control mechanisms. This section focuses on identification of most suitable access control approach in an interoperable health environment that could support secured sharing of the sensitive EHRs between health professionals.

4.2 Identifying the most suitable Access Control Model using Fuzzy TOPSIS

In order to select most apt access control model one need to consider basic security criteria like preserving confidentiality, ensuring integrity, timely availability, proper authorization, handling compatible organization policies and policy-conflict. The overall objective is to design a framework which provides clearly identified rules and authorizations to determine secured access of health records to the destined user.

Designing a framework for sharing sensitive and critical medical data is a complex process. It should include the environment in which sharing of record take place, objective with which sharing is done and level of security required. Especially in interoperable EHR environment where every patient may entail a different sensitivity level of security with respect to his/her health details. Further different data have different sensitivity, like demographic data after sanitation can be completely used for research purpose and easily shared to anyone but different medical information related to patient can be highly sensitive not to be disclosed even to subordinate. The security demand differs and depends on various factors such as occupation, designation, family background, disease etc. The data is very complex and have different attributes. Hence, we need to rank the access control models with respect to specific security demands of health data users.

The guideline of sharing EHR [205] differs with the change in access conditions, need and authority. Today, system required to have dynamic and concrete access control policies not only as per the need and the designation of requester, but also the delegated authority and profile of the sender. Sometime, confidentiality is given higher precedence over availability, whereas in some another case availability tend to be higher than confidentiality. Hence, it is necessary to maintain a balance between confidentiality, integrity and availability while accessing EHRs across hospitals.
Choosing the best alternative among the given options such that need of even conflicting criteria’s are considered, is always a challenging task for the decision makers and can be suitably handled using multi criteria decision making approach (MCDM). Fuzzy TOPSIS, an important MCDM technique appropriate in conditions with insufficient real data, is considered. It selects the best alternative among the set of feasible alternatives by evaluating distance between positive and negative ideal distance. It enables better decision-making, apprehending inherent vagueness prevalent in the evaluating problem. Further, sensitivity analysis is performed so that the suitable access model can be identified if the need of the environment changes.

<table>
<thead>
<tr>
<th>Alternatives</th>
<th>Features</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attribute-Based Access Control-</td>
<td>No advance knowledge of requesters for protected data/resources are</td>
<td>Highly adaptable to changing needs or operational environments Policy attributes and access decision engines can be managed centrally for large enterprises. Reduces administration efforts and duplication in managing user accounts across various environments (i.e., cross-domain) for assured information sharing. Various levels of authorizations through robust authentication mechanisms fine-grains access control model</td>
<td>Attributes in heterogeneous systems grow exponentially. Managing the access and security of data based on these attributes tends to be complex and expensive.</td>
<td>Stine et al. [152] NISTSP 800-162 ABAC [206]</td>
</tr>
<tr>
<td>A1</td>
<td>required</td>
<td>Individual attributes can be correlated from multiple authoritative sources ABAC include: Subject – An entity needing access. Object – An entity for which access is needed</td>
<td>Attributes – Information or characteristics of subjects, objects or environment conditions. Operations – Functions such as read, write, create, modify or delete. Policy – Formal rules that could be access rules, relationships, etc</td>
<td></td>
</tr>
<tr>
<td>Purpose-Based Access Model-</td>
<td>Generates the Access capability list and the Purpose capability list for both allowed or denied access to the specified data elements Capture varied requirements of users and integrate them in a common list before allowing the access to unknown but legitimate users.</td>
<td>Provides higher privacy preservation by allocating resources with purposes Decision can be made for sharing between disparate organizations by decomposing a global policy without losing the sensitive information.</td>
<td>Definition of purposes is a complicated task and requires much care and intervention of the professionals. Dependent on how fine-grained the data elements are, that is again highly complicated and time-consuming.</td>
<td>Byun &amp; Li [156] Naikuo et al. [157] Lin et al. [159] Gajanayake et al. [163]</td>
</tr>
<tr>
<td>A2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hierarchical Role-based Access</td>
<td>XACML policies are used to implement role inheritance. The policy instances allow referencing the Permission associated in one role inside the Permission associated with another role. Consents and rights are decided on the</td>
<td>Role hierarchies organize roles to reflect authority The hierarchies are highly scalable and versatile It is most conducive to interoperability.</td>
<td>Overlapping of roles RBAC is incapable to integrate other access parameters or related data that are significant in allowing access to the user Difficult to fine-grain the access control for each individual with</td>
<td>Park and Sandhu [153] Kuang et al. [154] Evered and Bögeholz [155]</td>
</tr>
<tr>
<td>Model- A3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 4.2.1 Alternatives

Table 4.1 discusses some access control models which are popularly used for sharing data through an interoperable environment. Fuzzy TOPSIS technique is applied on these models to find the most suitable alternative for designing the framework. Access controls compared in this study are primarily concerned with application-level information security rather than system-level security.

<table>
<thead>
<tr>
<th>Alternatives</th>
<th>Features</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identity-Based Access Model (IBAC) - A4</td>
<td>Includes Access Control Lists (ACLs) for implementing Discretionary Access Control (DAC) policies. IBAC looks at two factors: the actual identity of the user and the role of individual in an organization. Checks for access authorization using group structures in the organization. Decision on legitimate policy for the said user is based on various elements stating how and from where the user is connected with the system.</td>
<td>DAC restricts access to objects based on the identity of the subjects (end-users), and/or groups to which they belong. A connection profile is established for each user stating valid access policy. Improves network security by eliminating excess privilege allowing only appropriate access as per the user need.</td>
<td>Granting read access is transitive and the policies are helpless for Trojan Horse Attack. IBAC is difficult to manage, as it requires global updates with each entry or exit of users from the system or whenever there is any change in their access rights. Synchronization is equally difficult task since users are still managed in each service domain.</td>
<td>Hu et al.[160] Ferraiolo et al.[150]</td>
</tr>
<tr>
<td>Coalition Based Access Model - A5</td>
<td>Provides broad functionality and flexibility in collaborative environments. Specifies domain entities and adds conceptual elements to support semantics for collating access policies. Elements-Authorization Sets, Constraints on the Authorization Set and context state defines “who” should have access to “what” and under “which circumstances.”</td>
<td>Overcomes the limitations of RBAC through the inclusion of teams and tasks entities. Reflect our understanding of “real world” ingredients of coalitions.</td>
<td>Does not include a top-level organization in which ultimate authority is vested. Distributed administrative authorities must exist in such models so that distribution can be negotiated. Offers very little support for identifying relationships between users and their activities.</td>
<td>Cohen et al.[147]</td>
</tr>
<tr>
<td>Constraint-Based Access Model – A6</td>
<td>Security constraints influence the decision to allow/ deny access to the resource while sharing. Constraints are inherited within a role hierarchy. Constraints are associated with the static user-role assignments or dynamic user-sessions activations.</td>
<td>Provide clear distinction of possible relationship between two organizations and thus enable better integration of two distinct systems</td>
<td>It is difficult to handle dynamic Separation of Duty (SOD) as the constraints are automatically inherited down the defined hierarchy.</td>
<td>Mouratidis[162] Verhanneman et al.[172]</td>
</tr>
</tbody>
</table>
4.2.2 Criteria

The choice of an access control model is governed by various factors such as administrative convenience, policy-support capabilities and access control mechanisms, environment [168][169] it is used, how much sharing is permissible and usage. While designing the framework, security issues are discussed with respect to confidentiality, integrity, availability, and authorization, compatibility and conflict resolution. These criteria are conflicting in nature as increasing availability without proper incorporation of authorization may decrease the confidentiality and integrity. Also due to conflict or lack of compatibility between policies there may be a need to relax the authorization policy. The decision of a suitable access control depends on the given criteria.

- **Confidentiality** (C1): FISMA [44 U.S.C., Sec. 3542] [17] defines confidentiality as “preserving authorized restrictions on information access and disclosure, i.e. protecting personal privacy and proprietary information from any unauthorized access”. Different measure used to ensure confidentiality includes proper Authentication using password, single-sign-on, OTPs, two-factor authentication, biometrics etc. It further instigate the adoption of various laws HIPAA [165], HL7 [166] and guidelines [167][173] to achieve the desired security. Also, one need to establish control to identify how many individuals would be accessing the health information and also what information they are accessing. This should be dynamic enough with respect to Context of Use so that it can apply a different confidentiality impact level as per the use.

- **Integrity** (C2): FISMA [44 U.S.C., Sec. 3542][17] “Guarding against improper information modification or destruction, and ensuring information non-repudiation and authenticity”, ensures the integrity of any system. Data integrity across organizations is compromised due to non-integrated and disparate schemas [170] and languages and lack of training, conditional constraints like handling emergencies may lead to inconsistent data even within the hospital. Such problems need to be addressed within the security control framework.

- **Availability** (C3): FISMA [44 U.S.C., Sec. 3542] [17] “Ensuring timely and reliable access to and use of information” determines its availability. Stringent security policies affect the performance of the health application thereby delaying the availability of the data. Also, handling security and availability of data in emergency cases is a major challenge. The thrust is to identify the access control framework able to handle such contingencies in the best possible manner.
• **Authorizations** (C4): At various times, access control and authorizations are often mistakenly interchanged. Authorization means determining the possibilities to allow/deny the access to the resource through well-defined access controls and authentications, i.e. access control refers to a more generic concept of controlling access of data in the concerned environment. Coordination of authorizations [174] needs to be ascertained especially in collaborative environment. Location information for participating authority [175] expressed through access control systems allows a natural interoperability even in decentralized policies. With the switching of user in between departments or locations, a lot of undesired access is assigned overtime, which needs be addressed. Whether the person is allowed an access to the desired data and what actions he is permitted to perform on that data. A fine granularity can be achieved if this criterion is appropriately addressed in designing of an access control model. Achieving clearly defined authorizations would enable the administrators to provide the right level of access to the desired and authorized users.

**Organization Policies** (C5): Each hospital defines own set of rules and guidelines for its users. Integrating rules and policies of disparate hospitals divulge that similar resource may be differently accessible by the users of in their respective hospitals. Access control policies need to be adaptable and reconfigurable [172] supporting interoperable sharing of health records.

**Policy Conflicts Resolution (NIST)** (C6): Controlling and managing access control policies in integrated healthcare organizations may often lead to policy conflicts [174]. Access controls provide the limit to identify and manage the relevant set of policies. It is required to detect and remove existing and exorbitant policy conflicts [176] resulting from merging of these policies in an attempt to establish relationship between disparate EHR-systems.

### 4.2.3 MCDM Approach

MCDM approach facilitates decision making and uses numeric techniques [177] to help decision makers (DMs) choose from among the given set of alternatives. The MCDM approach works on the given ratings and the weights to evaluate each criterion. The decision makers [178] provide judgment based on collective group ideas instead of individual opinions. Popular MCDM methods[179][180] are AHP (Analytic Hierarchy Process), TOPSIS (Technique for Order Preference by Similarity to Ideal Solution), PROMETHEE
(Preference Ranking Organization Method for Enrichment Evaluations) and ELECTRE. Each of the method has some advantages and disadvantages. AHP does not handle interdependence of criteria when compared with other techniques. TOPSIS evaluates the alternatives on the basis of the shortest distance from ideal solution and the farthest distance from the negative-ideal solution. TOPSIS though easy to implement uses Euclidean distance without considering the correlation of the attributes. ELECTRE generates binary outranking relations between the alternatives and is unable to determine preferred alternative. To include uncertainty of decision maker, fuzzy set theory is used to handle data complexities. It works on imprecise input and takes into consideration the insufficient information thereby encompassing problems with high complexities into few rules.

IT heads of fifteen hospitals were appointed as the decision makers (DMs) to rank the access control models on their viability of addressing interoperability and security demands in sharing of EHRs. The choice of hospitals was done based on level of implementation and sharing of EHR in inter and intra organization. The hospitals were then categorized on the frequency of sharing as high, medium and low sharing of EHR. The IT heads were individually interviewed to record their opinions on the stated alternatives and capabilities of these alternatives to handle the specified security criteria. The major voting approach was adopted for finalizing the input of Decision Maker (DM).

4.2.3.1 Fuzzy Set Theory

Fuzzy TOPSIS (F-TOPSIS) [181] is a method based on measuring the creditability of each criterion among the stated alternatives. F-TOPSIS works on ranking the criteria and sub-criteria using linguistic variables represented as triangular fuzzy numbers.

Definition1: In a universe of discourse X, a fuzzy subset A of X is defined by a membership function fA(x), which maps each element x in X to a real number in the interval [0, 1]. The function value fA(x) represents the grade of membership of x in A. The nearer the values of fA(x) to unity, the higher is the grade of membership of x in A.

Definition2: A linguistic variable represented as triangular fuzzy number forming triplet A = (c, a, b). Each of these values is a fuzzy variable defined in X. A linguistic variable is characterized as the membership function fA(x) of triangular fuzzy number A and is given by:
\[ f_A(x) = \begin{cases} 
0, & x \leq a \\
\frac{x-c}{a-c}, & c \leq x \leq a \\
\frac{x-b}{a-b}, & a \leq x \leq b \\
0, & x \geq c 
\end{cases} \]  

(1)

With \(-\infty < c \leq a \leq b \leq \infty\). The value of \(x\) at \(a\) gives the maximal grade of \(f_A(x)\) i.e., \(f_A(x) = 1\); it is the highest possible value obtained on evaluation of data. The value of \(x\) at \(c\) gives the minimal grade of \(f_A(x)\), i.e., \(f_A(x) = 0\); it is the least possible value that evaluation of data can generate. Constants \(c\) and \(b\) are the lower and upper bounds of the available area for the evaluation data. These constants reflect the fuzziness of the evaluation data. The narrower the interval \([c, b]\), the lower is the fuzziness of the evaluation data.

Let \(A_1 = (c_1, a_1, b_1)\) and \(A_2 = (c_2, a_2, b_2)\) be two triangular fuzzy numbers. The distance \(d(A_1, A_2)\) between them is given by (using the vertex method):

\[ d(A_1, A_2) = \frac{1}{\sqrt{3}} \left[ (c_1 - c_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2 \right] \]  

(2)

The implementation steps of Fuzzy TOPSIS used to identify the most suitable access control model is shown in figure 4.1.
4.2.4 Implementation of Fuzzy TOPSIS in Healthcare Environment

A committee of decision makers (DMs) consisting the IT heads of health organizations was formed to select the best option. The DMs used the linguistic variables to evaluate the importance of each criterion. Table 4.3 states the potential impact (as opined by the DMs) of each alternative on achieving the security criteria during EHR sharing in interoperable environment. The impact ranges from low to catastrophic covering entire range of affected users - physicians, patients, technicians, etc.

Classifying Criteria and Alternatives: Fig. 4.2 describes the hierarchical structure followed by Fuzzy TOPSIS to evaluate and rank the given alternatives and criteria. Six alternatives (A1-A6) allowing secured sharing of EHR are chosen for evaluation. The assessment is based on six criteria (C1-C6) extracted from real-time and literature study.
Transition of real numbers into fuzzy numbers using linguistic variables: Linguistic variables provide a means of approximate characterization of problems which are too complex or imprecise to be described into precise terms. Linguistic variables are defined [181] not in terms of numbers but words or sentences as linguistic characterization are less specific than the numeric ones.

Table 2: Alternatives and criteria Ratings

<table>
<thead>
<tr>
<th>Linguistic terms</th>
<th>Linguistic terms</th>
<th>Membership Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Poor (VP)</td>
<td>Very Low (VL)</td>
<td>(1,1,3)</td>
</tr>
<tr>
<td>Poor (P)</td>
<td>Low (L)</td>
<td>(1,3,5)</td>
</tr>
<tr>
<td>Fair (F)</td>
<td>Medium (M)</td>
<td>(3,5,7)</td>
</tr>
<tr>
<td>Good (G)</td>
<td>High (H)</td>
<td>(5,7,9)</td>
</tr>
<tr>
<td>Very Good (VG)</td>
<td>Very High (VH)</td>
<td>(7,9,9)</td>
</tr>
</tbody>
</table>

The rating set (Table 4.2) for the observed alternatives is defined as \( S = \{ VP, P, F, G, VG \} \); where \( VP = \text{Very Poor}, P = \text{Poor}, F = \text{Fair}, G = \text{Good}, \) and \( VG = \text{Very Good} \). Similarly, the rating set for the selected criteria is defined as \( P = \{ VL, L, M, H, VH \} \); where \( VL = \text{Very Low}, L = \text{Low}, M = \text{Medium}, H = \text{High}, \) and \( VH = \text{Very High} \). The linguistic variables \( \{1, 3, 5, 7, 9\} \) are assigned to each rating set starting from VP or VL and reaching to VG or VH respectively.

Table 3: Linguistic Assessment of Each Alternative for each Criterion

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Alternatives</th>
<th>Decision Makers</th>
<th>Aggregate Fuzzy Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A1</td>
<td>VG, G, F</td>
<td>(3, 7, 9)</td>
</tr>
<tr>
<td>C1</td>
<td>A2</td>
<td>G, G, G</td>
<td>(5, 7, 9)</td>
</tr>
<tr>
<td></td>
<td>A3</td>
<td>G, G, G</td>
<td>(5, 7, 9)</td>
</tr>
</tbody>
</table>
DMs rated each alternative with respect to each criterion as depicted in figure 4.2. Before gathering the individual opinions a generic set of indicators to be achieved for idealistic sharable environment are determined. These indicators calculate the aggregate fuzzy ratings for the criteria. For instance, DMs opinion for criterion C2 that discusses ensuring data integrity while interoperable sharing of EHR, is: {H, H, VH}. Converting the opinion into linguistic variables generate {7, 7, 9} fuzzy values. Similar conversions are listed in table 4.3 for all stated alternatives.

Referring to table 4.3, the aggregate fuzzy weights for A2 w.r.t. C2 are: {5, 7.67, 9}. The Fuzzy Decision Matrix in table 4.4 for the alternatives (D) and the criteria (W) is obtained by dividing each weight by the maximum weight for that criterion. The calculations are shown in table 3.22, for instance, the Fuzzy decision matrix for the combination C2 (A2) is {0.33, 0.63, 1}. Further, weighted normalized matrix for the alternatives is obtained (Table
4.5) by dividing each value of fuzzy decision matrix with respective fuzzy weight rating of each criterion.

Table Error! No text of specified style in document..4: Normalized Fuzzy Decision Matrix

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Alternatives</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td></td>
<td>(0.333, 0.778,1)</td>
<td>(0.333, 0.629,1)</td>
<td>(0.556, 0.778,1)</td>
<td>(0.143, 0.524,1)</td>
<td>(0.143, 0.524,1)</td>
<td>(0.2, 0.333,1)</td>
</tr>
<tr>
<td>C2</td>
<td></td>
<td>(0.556, 0.778,1)</td>
<td>(0.556, 0.852,1)</td>
<td>(0.556, 0.778,1)</td>
<td>(0.556, 0.778,1)</td>
<td>(0.556, 0.926,1)</td>
<td>(0.556, 0.778,1)</td>
</tr>
<tr>
<td>C3</td>
<td></td>
<td>(0.556, 0.778,1)</td>
<td>(0.556, 0.852,1)</td>
<td>(0.556, 0.852,1)</td>
<td>(0.556, 0.778,1)</td>
<td>(0.556, 0.778,1)</td>
<td>(0.556, 0.778,1)</td>
</tr>
<tr>
<td>C4</td>
<td></td>
<td>(0.333, 0.737,1)</td>
<td>(0.333, 0.704,1)</td>
<td>(0.333, 0.704,1)</td>
<td>(0.556, 0.778,1)</td>
<td>(0.556, 0.704,1)</td>
<td>(0.556, 0.929,1)</td>
</tr>
<tr>
<td>C5</td>
<td></td>
<td>(0.556, 0.926,1)</td>
<td>(0.778,1,1)</td>
<td>(0.556, 0.926,1)</td>
<td>(0.556, 0.926,1)</td>
<td>(0.778,1,1)</td>
<td>(0.778,1,1)</td>
</tr>
<tr>
<td>C6</td>
<td></td>
<td>(0.778, 0.852, 1)</td>
<td>(0.333, 0.778,1)</td>
<td>(0.333, 0.629,1)</td>
<td>(0.333, 0.629,1)</td>
<td>(0.333, 0.704a,1)</td>
<td>(0.333, 0.704,1)</td>
</tr>
</tbody>
</table>

Table Error! No text of specified style in document..5: Weighted Normalized Fuzzy Decision Matrix for Alternatives

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Alternatives</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td></td>
<td>(2.333,7,9)</td>
<td>(2.333,5,667,9)</td>
<td>(3.889,7,9)</td>
<td>(1.4, 714,9)</td>
<td>(1.4,714,9)</td>
<td>(1.4,3,9)</td>
</tr>
<tr>
<td>C2</td>
<td></td>
<td>(2.778,5,969,9)</td>
<td>(2.778,6,531,9)</td>
<td>(2.778,5,963,9)</td>
<td>(2.778,7,999,9)</td>
<td>2.778, 5,963,9</td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td></td>
<td>(3.889,6,481,9)</td>
<td>(3.889,7,098,9)</td>
<td>(3.889,7,098,9)</td>
<td>(3.889, 6,481,9)</td>
<td>(3.889,6,481,9)</td>
<td>(3.889,5,864,9)</td>
</tr>
<tr>
<td>C4</td>
<td></td>
<td>(2.333,5,864,9)</td>
<td>(2.333,5,864,9)</td>
<td>(2.333,5,864,9)</td>
<td>(3.889,6,481,9)</td>
<td>(3.889,5,864,9)</td>
<td>(3.889,7,716,9)</td>
</tr>
<tr>
<td>C5</td>
<td></td>
<td>(2.778,7,099,9)</td>
<td>(3.889,7,667,9)</td>
<td>(3.889,7,667,9)</td>
<td>(2.778,7,098,9)</td>
<td>(2.778,7,099,9)</td>
<td>(3.889,7,667,9)</td>
</tr>
<tr>
<td>C6</td>
<td></td>
<td>(3.889,5,969,9)</td>
<td>(1.667,5,444,9)</td>
<td>(1.667,4,407,9)</td>
<td>(1.667,4,407,9)</td>
<td>(1.667,4,926,9)</td>
<td>(1.667,4,926,9)</td>
</tr>
</tbody>
</table>

The ideal (FPIS) and anti-ideal (FNIS) are based on the concept of relative closeness in compliance with the shorter/longer the distance of each alternative to FPIS (FNIS), the higher priority can be ranked. The distance ($d_i^+, d_i^-$) of each weighted alternative $i=1,2,\ldots,m$ from FPIS and FNIS is measured. Fuzzy positive ideal solution ($A^*$) and fuzzy negative ideal solution ($\bar{A}$) for all the alternatives is thus calculated and the distance ($d_i^+, d_i^-$) of each weighted alternative $i=1,2,\ldots,m$ from FPIS and FNIS is measured as shown in table 4.6.

Table Error! No text of specified style in document..6: Distance of each alternative from FPIS and FNIS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>4.02</td>
<td>4.30</td>
<td>3.17</td>
<td>5.24</td>
<td>5.24</td>
<td>5.59</td>
</tr>
<tr>
<td>C2</td>
<td>4.00</td>
<td>3.86</td>
<td>4.00</td>
<td>4.00</td>
<td>3.76</td>
<td>4.00</td>
</tr>
<tr>
<td>C3</td>
<td>3.29</td>
<td>3.15</td>
<td>3.15</td>
<td>3.29</td>
<td>3.29</td>
<td>3.31</td>
</tr>
<tr>
<td>C4</td>
<td>4.25</td>
<td>4.25</td>
<td>4.25</td>
<td>3.29</td>
<td>3.46</td>
<td>4.36</td>
</tr>
<tr>
<td>C5</td>
<td>3.76</td>
<td>3.05</td>
<td>3.05</td>
<td>3.76</td>
<td>3.05</td>
<td>4.37</td>
</tr>
<tr>
<td>C6</td>
<td>3.43</td>
<td>4.70</td>
<td>5.00</td>
<td>4.84</td>
<td>5.10</td>
<td>4.76</td>
</tr>
</tbody>
</table>

$d_i^+, d_i^-$
Using the distance of each alternative obtained in table 4.6, the closeness coefficient of each alternative is calculated (Table 4.7). The closeness coefficient (CCi) is the final ranking determining the suitability of each alternative measured on the selected criteria. The current research work observes alternative A1 i.e. Attribute-based Access Control model as the best model with CCi equal to 0.532 to ensure smooth and secured sharing of EHR in interoperable healthcare environment. The second best model marginally behind A1 is A3 i.e. Hierarchical RBAC with CCi equal to 0.530.

### Table: Closeness Coefficient for all the alternatives

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
</tr>
</thead>
<tbody>
<tr>
<td>(d_i^+)</td>
<td>22.748</td>
<td>23.325</td>
<td>22.614</td>
<td>24.427</td>
<td>24.348</td>
<td>23.813</td>
</tr>
<tr>
<td>(d_i^-)</td>
<td>25.842</td>
<td>25.713</td>
<td>25.710</td>
<td>25.192</td>
<td>25.294</td>
<td>25.522</td>
</tr>
<tr>
<td>CCi</td>
<td>0.532</td>
<td>0.524</td>
<td>0.530</td>
<td>0.508</td>
<td>0.510</td>
<td>0.517</td>
</tr>
</tbody>
</table>

### Ranking the alternatives according to CC:

On comparing the CCi values for all the alternatives, it is observed that alternative A1>A3>A2>A6>A5>A4. Thus, A1 i.e. Attribute Based Access Model (ABAC) is the most optimal solution derived through F-TOPSIS for secured exchange of EHR between healthcare organizations. The study is further extended to observe the ranking of alternatives by inducing a change in one or more criterion at a time. A sensitivity analysis is performed on the defined fuzzy ratings by varying the criteria weights.

### Table: Sensitivity Analysis Experiments and Rankings

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Criteria</th>
<th>Experiment</th>
<th>A1 (Attribute-Based)</th>
<th>A2 (Policy-Based)</th>
<th>A3 (Hierarchical Role-Based)</th>
<th>A4 (Identity-Based)</th>
<th>A5 (Coalition Based)</th>
<th>A6 (Constraint-Based)</th>
<th>Alternatives Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Very High Confidentiality</td>
<td>E1</td>
<td>0.587</td>
<td>0.348</td>
<td>0.578</td>
<td>0.514</td>
<td>0.517</td>
<td>0.507</td>
<td>A1&gt;A3&gt;A2&gt;A5&gt;A4&gt;A6</td>
</tr>
<tr>
<td>2</td>
<td>Very High Integrity</td>
<td>E2</td>
<td>0.558</td>
<td>0.541</td>
<td>0.537</td>
<td>0.519</td>
<td>0.533</td>
<td>0.520</td>
<td>A1&gt;A2&gt;A3&gt;A5&gt;A6&gt;A4</td>
</tr>
<tr>
<td>3</td>
<td>Very High Availability</td>
<td>E3</td>
<td>0.558</td>
<td>0.541</td>
<td>0.542</td>
<td>0.524</td>
<td>0.523</td>
<td>0.520</td>
<td>A1&gt;A3&gt;A2&gt;A4&gt;A6</td>
</tr>
<tr>
<td>4</td>
<td>Very High Authorization and High policy-conflict</td>
<td>E4</td>
<td>0.556</td>
<td>0.547</td>
<td>0.548</td>
<td>0.558</td>
<td>0.554</td>
<td>0.570</td>
<td>A6&gt;A1&gt;A4&gt;A5&gt;A3&gt;A2</td>
</tr>
<tr>
<td>5</td>
<td>Very High compatible Organization Policies</td>
<td>E5</td>
<td>0.567</td>
<td>0.581</td>
<td>0.582</td>
<td>0.545</td>
<td>0.547</td>
<td>0.572</td>
<td>A3&gt;A2&gt;A6&gt;A1&gt;A5&gt;A4</td>
</tr>
<tr>
<td>6</td>
<td>Very High Policy Conflict Resolution</td>
<td>E6</td>
<td>0.604</td>
<td>0.552</td>
<td>0.544</td>
<td>0.531</td>
<td>0.538</td>
<td>0.539</td>
<td>A1&gt;A2&gt;A3&gt;A6&gt;A5&gt;A4</td>
</tr>
<tr>
<td>7</td>
<td>Very High Confidentiality and High authorization</td>
<td>E7</td>
<td>0.564</td>
<td>0.541</td>
<td>0.562</td>
<td>0.527</td>
<td>0.526</td>
<td>0.529</td>
<td>A1&gt;A3&gt;A2&gt;A6&gt;A5</td>
</tr>
<tr>
<td>8</td>
<td>Very High Integrity and High authorization</td>
<td>E8</td>
<td>0.539</td>
<td>0.535</td>
<td>0.532</td>
<td>0.533</td>
<td>0.538</td>
<td>0.542</td>
<td>A6&gt;A1&gt;A5&gt;A2&gt;A3</td>
</tr>
<tr>
<td>9</td>
<td>Very High Availability and High Authorization</td>
<td>E9</td>
<td>0.539</td>
<td>0.535</td>
<td>0.536</td>
<td>0.537</td>
<td>0.531</td>
<td>0.542</td>
<td>A6&gt;A1&gt;A4&gt;A3&gt;A2&gt;A5</td>
</tr>
</tbody>
</table>
4.2.4.1 Sensitivity Analysis

Sensitivity analysis is done to identify the robustness of the ranking obtained from the above results, as it handles the inherent instability of the environment. The ranking depends on prioritizing any criterion over others to judge the suitability of the ranked alternatives and observe the change in their ranks in order to reach to a consensus. The robustness of these weights was investigated by changing criteria weights atomically and in combination both. 14 experiments (E1 - E14) each having different composition of criteria weights were considered for the sensitivity analysis. Table 4.8 show the result obtained.

Figure 4.3 shows the order of each alternative based on CCi index values with respect to different weight configurations. In experiment E1, criterion C1 is set to very high with all other criteria set to very low. Similar variations are performed for all other criteria (C2-C6) in experiments E2-E6. Experiment E7 combines two criteria C1 and C4 with C1 set to very high and C4 set to high. Experiment E8 - E10 replicates the same for (C2, C4), (C3, C4) and (C4, C6).Experiment E11 sets C2 as high with C4 as very high. Experiment E12 shows a high and medium combination of criteria C3 and C6 respectively. Similar combinations are observed between other criteria namely, (C2, C6) and (C1, C5).
The result of the sensitivity analysis identifies that from all the 14 experiments, alternative A1 has resulted as the most viable solution in 8 experiments followed by alternative A6 in 5 experiments. Relating sensitivity analysis results with varied EHR access environments, A1-attribute-based access control model and A6-Constraint based access control model seems to cover maximum security parameters as demanded by the healthcare stakeholders.

4.2.5 Observation

Viewing security demands with respect to stakeholders of EHR, it is observed that obtaining an adaptable, reliable and user-friendly access control model for secured sharing of EHR is the prime concern. Ensuring secured sharing of EHR in interoperable environment can be very well absorbed into a multi-dimension problem domain. Handling all the complexities and imprecision with respect to interoperable environment, fuzzy theory is believed to satisfy and generate the best decision. Fuzzy TOPSIS, a MCDM approach, implemented in healthcare environment enabled better decision making and improve the security of EHR sharing. Using F-TOPSIS method the alternatives were ranked and a conclusion was devised that predicted the most suitable alternative satisfying the listed security criteria to the best. The criteria were allocated fuzzy weights by the decision makers based on the demand and feasibility in the concerned environment. Access control models presume to be an efficient mechanism in ascertaining the secured access of EHR. Thus, six access control models are evaluated with respect to various security parameters taken as criteria. A preference rank
among the alternatives is obtained. Sensitivity analysis performed on the observed data set reveals that with the change in criteria weights, the ranking order among the alternatives are likely to change. This concludes that environments such as web-based and cloud-based can be built on A1 whereas mobile-based EHR sharing will be more secured with A6. Going a step further in the obtained analysis, A3 and A2 emerge as the second best alternative in certain observations. Hence, it concludes that ranking among the alternatives is quite sensitive to the changes in the weights of evaluation criteria.

4.4 Environment for the Proposed Framework

Access control models constitute access control policies that need to be mapped in a secured manner to allow sharing of sensitive health data. There exist several access control policies satisfying different genre of domains. The languages used for writing the Access Control Policies (ACPs) were explored finding Extensible Access Control Markup Language (XACML) suitable for the purpose. A successful and secured collaboration may exist if the access rules are governed by well-defined authorizations. ABAC implementation can help in predetermining access authorizations. This section describes the basic architectures of XACML and ABAC, whose combination is used to propose a secured framework for the problem identified in this research.

4.4.1 XACML Structure

XACML is an OASIS standard [151] that describes a policy language for representing authorization policies and an access control decision request/response language. The XACML model [175] employs elements such as rules, policies, rule attributes (subject, resource, action and environment conditions) and policy-combining algorithms.
Figure 4.4 demonstrates processing ACPs with respect to the syntax and semantics of XACML structure. The request is made through an application in the set format that the policy responds in the similar fashion for determining applicability of policies to requests. The request and response formats represent a standard interface, between a Policy Decision Point (PDP) and a Policy Enforcement Point (PEP). PDP presents standard behavior when processing policy whereas PEP issues requests and handle responses. The standard description of each XACML terminology is stated in Table 4.9.

### Table 4.9: XACML Policy Structure

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>The entity (user) requesting to access the resource. The subject is sometimes referred to as a requestor.</td>
</tr>
<tr>
<td>Subject Attributes</td>
<td>Characteristics or attributes of a subject such as name, designation, authority that distinguishes that person from all others.</td>
</tr>
<tr>
<td>Object/Resource</td>
<td>Entity (data) to be protected from unauthorized use.</td>
</tr>
<tr>
<td>Privilege/Action</td>
<td>The authorized behavior of a subject as defined by an authority and embodied in policy or rules.</td>
</tr>
<tr>
<td>Environment Condition</td>
<td>Constraints set on the access rules and used as attributes at decision time to influence an access decision.</td>
</tr>
<tr>
<td>Policy and rules</td>
<td>A repository of privileges available to subjects from the perspective of the resource that needs protection.</td>
</tr>
<tr>
<td>Access control mechanism</td>
<td>Access control mechanism is the process of assembling authorization information, which includes information about the resource to be protected, the subject requesting access, the policies governing access to the resource, and any contextual information needed to make a decision.</td>
</tr>
<tr>
<td>Policy Combining Point (PCP)</td>
<td>To combine disparate policies using rule-combining algorithm</td>
</tr>
<tr>
<td>Policy Evaluation Point (PEvP)</td>
<td>To verify the rules against set properties and environmental conditions</td>
</tr>
<tr>
<td>Policy Enforcement Point (PEP)</td>
<td>To enforce the decision using SLs and authorizations</td>
</tr>
<tr>
<td>Policy Decision Point (PDP)</td>
<td>To render a decision and generate matching rules</td>
</tr>
</tbody>
</table>
An XACML policy [175][182] structure shown in figure 4.5 consist of three major components- a Target, a Rule Set and a rule combining algorithm. An access control policy P is a tuple (T, t, S, R, A, C, E), that consist of attributes: - a Policy Target (T) consisting of Rule target (t) defined for a finite set of Subject (S), Resource (R) and Action (A) along with Condition (C), if any. Each rule defines the Effect (E) that decides whether the access should be granted or denied. The Target (T) component refers to the department or group to which the policy (P) is applicable. The subject may be any user like a family doctor, intern, patient, etc. requiring access to the resource. Resource is various types of patient’s health data categorized as demographic data, clinical data and financial data. Access to these resources by the subjects may have multiple actions. Actions are the permissions like read, write, view, print etc. on the resources. Actions may be bounded with conditional constraints enabling restricted access to the users. The Condition element (C) is optional and is used to impose a constraint on the rules, if required. Whenever access request is made in XACML, it is passed to PDP that evaluates the request against the given access control policies and provides permit or deny decisions accordingly.

---

**Figure** Error! No text of specified style in document..5: XACML Policy Structure

Rule defined in policy P represents a particular domain attribute (Dom) i.e. Subject, Resource, Action or Condition, comprising of one or more attribute values (Val). Each policy comprise of multiple rules forming unique tuples comprising of XACML attributes. Each rule is governed by the effect (E) i.e. permit rules and deny decisions, such that:
Subject (S) = \{s_1, s_2, s_3, \ldots \ldots , s_n\} where, s_i \subset S or s_i \subseteq S ranging 1 \leq k \leq n where k is a subset of n that defines the total number of subjects in the healthcare unit such that 1 \leq S \leq n.

Resource (R) = \{r_1, r_2, r_3, \ldots \ldots , r_n\} where, r_i \subset R or r_i \subseteq R ranging 1 \leq k \leq n where k is a subset of n that defines the total number of resources in the healthcare unit such that 1 \leq R \leq n.

Action (A) = \{a_1, a_2, a_3, \ldots \ldots , a_n\} where, a_i \subset A or a_i \subseteq A ranging 1 \leq k \leq n wherek is a subset of n that defines the possible number of actions a Subject can perform on the resource such that 1 \leq A \leq n.

The effect (E) can be set to either Permit or Deny. It determines decision that allows or denies access to the resource (R) i.e. EHR to the user (S) specified in the rule attribute. Action (A) attribute withholds what the user can do with the data. A rule is a combination of a subset of Subject(S), Resource(R) and Action (A).

4.4.2 Attribute based Access Control Model (ABAC)

ABAC is a logical access control methodology where attributes are evaluated to determine authorization. Authorization allows the user to perform a set of operations determined by evaluating attributes associated with the subject, object, requested operations, and, in some cases, environment conditions. An access control framework that is consistent with ABAC is XACML. ABAC designed using XACML defines the following possible states of allowing access of EHR to the legitimate users:

Definition 1: An ABAC rule r states: if C then E, i.e. where the condition C is met the Effect is permit. For ex. if the hospital administrator allows the specialist to have write permission on the patient’s diagnostic records and he property defined permits Specialist to write to the diagnostic data, the decision would be permit.

Definition 2: An ABAC rule r states: if S and R then E, i.e. where the combination of subject S and resource R is met, the Effect is permit. For ex. if the policy rule allows the Family doctor unconditional access to the Clinical records of the patient, the decision would be permit.
The above definitions allow clear and simple integration of various policies but encompass policy conflicts and rule redundancies. The policies need to be fine-grained especially while collaborating in interoperable environment so as to restrict undue disclosure of sensitive health records.

Attribute-based access control (ABAC) implementation can reduce the risk of unauthorized access to the data due to its ability of implementing consistent policies and also update them easily to address dynamic change in attribute values. ABAC system is partially dependent on the organization’s authorities for allocation of resources (EHR) to the respective subjects (Users). As defined by NIST, “An access control method (ACM) where subject requests to perform operations on objects are granted or denied based on assigned attributes of the subject, assigned attributes of the object, environment conditions, and a set of policies that are specified in terms of those attributes and conditions”.

A successful and secured collaboration may exist if the access rules are governed by well-defined authorizations. Authorization, which specifies who is able to access what EHR, demand clearly defined data ownership and consent management [184]. The problem can be sought through tightly coupled access control policies [148][185] capable enough to proactively inspect and respond optimally. The rule mappings between two collaborative access control policies can be done either manually or automatically. Manual mapping is not possible in a large domain as that of health care, however, even for automatic mapping, we need certain algorithms and principles that can generate a feasible rule-set free of any conflicts and disparities. The permutations and combination of rules of two different policies generate multiple rules. Maximal rules can achieve interoperability and allow access to the data, but also lead to security breaches. It requires the designing of framework that follow the principles of least privilege ensuring generation of relevant rules only.

ABAC verifies and takes decision on the basis of attribute values assigned under each XACML element in the given rules or policies. Under this arrangement, policies can be created and managed without direct reference to users and resources, and users and resources can be established without reference to policy. In the context of sharing sensitive health information, of the patient, across hospitals, an access control framework (Figure 6.1) designed using ABAC approach is proposed. The health providers (doctors, nurses and other staff) exhibit a vertical hierarchy according to their roles and responsibilities in the organization. The hierarchies may significantly differ from each other. Hence, these
hierarchies need to be standardized in order to generate commonalities between them. A filter-based approach can then be applied to identify the relevant matching roles based on the similarities of hierarchical positions of the users in their respective hospitals. Identifying similarities between user attributes would ensure robust authorization, thereby, enhancing the availability complemented with confidentiality and privacy of health data.

4.5 Similarity Score Algorithm

Lin et al. [183] developed a filter to pre-compile the policies and reach to the most similar policies for further evaluation in a grid computing system. It is a lightweight approach that safely reduces the policies by pruning design rules on the basis of hierarchical distances between attributes and elements of the policies. The similarity score is evaluated based on the cross-product of each rule attribute having one or many attribute values. The attributes may contain values of type categorical or numerical in nature.

The notations defined by Lin et al. to express the working of proposed algorithm are listed in table 4.12. The algorithm to compute policy similarity scores (P1, P2) is illustrated using a flowchart in Fig. 4.6. P1 and P2 are two policies having n, m rules respectively. Exclusive match is obtained on each element of the policy sets. Each rule is matched with all the rules of another policy and a similarity index is obtained. The obtained scores are then averaged to calculate the overall similarity score stating the level of similarities identified between both policies. The algorithm handles matched and unmatched rules in both the policies.
4.5.1 Computational Steps

- The rules are categorized on the basis of their effects, i.e. Permit (PR) or Deny (DR). Each rule is checked against the effect and stored in an array of permit rules (PR1, PR2) and deny rules (DR1, DR2) of policies P1 and P2 respectively.
The rule similarity function $S_{\text{rule}}$ between two rules $R_{1i}$ and $R_{2j}$ in policies $P_1$ and $P_2$ respectively, is computed as follows:

$$S_{\text{rule}} = w_t S_t + w_c S_c$$  \hspace{1cm} (1) \hspace{1cm} (\text{Rule Target (T) and Condition(C)})

$$S_t = w_s S_s + w_r S_r + w_a S_a$$  \hspace{1cm} (2) \hspace{1cm} (\text{Subject(S), Resource(R), Action (A)})

$w_t, w_c, w_s, w_r$ and $w_a$ are weights assigned to the XACML attributes such that $w_t + w_c = 1$, and $w_s + w_r + w_a = 1$ showing the relative importance of its contribution in overall similarity estimation.

Attributes and their values

The values $(v_{11}, v_{12}, \ldots)$ of the attributes \{a$_1$, a$_2$, a$_3$, \ldots\} in each element \{S, R, A, C\} may be of two types - categorical or numerical. The categorical values are different values of the same element that satisfies a particular rule. For instance, an attribute ‘Designation’ of element Subject contains values like Family_Doc, Matron, and Patient that belong to categorical type. All integers or date/time type values belong to the numeric type like time from 8.00 to 16.00.

- **To calculate the similarity score for attributes with categorical values:** The hierarchical distance between two categorical attributes is calculated for each value of attributes (a$_1$, a$_2$) to find the shortest path ($SPath(v_1,v_2)$, Eq. 4) in the hierarchy. $S_{\text{cat}}(a_1,a_2)$ is the similarity score (Eq. 3) for all matched attributes and compensating score $\delta$ is calculated (Eq. 5) as an average similarity score for all unmatched attributes $M_v$ (i.e. when v$_{ik}$ = v$_{2l}$) is the set of pairs matching attribute values and $N_v$ is the total number of values in an attribute. The attribute values (v$_{i1}$, v$_{i2}$, \ldots, v$_{in}$) of attributes \{a$_1$, a$_2$, a$_3$, \ldots\} are defined in the policy sets of two or more healthcare units. Each attributes \{a$_1$, a$_2$, a$_3$, \ldots\} consist of unique set of attribute values in the given policies.

$$S_{\text{cat}}(a_1,a_2) = \frac{1}{2} \left[ 1 + \frac{\sum_{v_{ik},v_{2l}} \text{scat}(v_{1k},v_{2l}) + \delta}{\max(N_v_{1},N_v_{2})} \right]$$  \hspace{1cm} (3), where

$\text{scat}(v_1,v_2) = 1 - \frac{SPath(v_1,v_2)}{2H}$ \hspace{1cm} (4) \hspace{1cm} $SPath(v_1,v_2)$ denotes the length of the shortest path between two values $v_1, v_2$ and H is the height of the hierarchy.
\[ \delta = \begin{cases} \frac{\sum_{(V_{1k}, V_{2l}) \in E^{p}_{1}} S_{\text{cat}}(V_{1k}, V_{2l})}{N_{v_{2}}}, & N_{v_{1}} > N_{v_{2}} \\ \frac{\sum_{(V_{1k}, V_{2l}) \in E^{n}_{1}} S_{\text{cat}}(V_{1k}, V_{2l})}{N_{v_{1}}}, & N_{v_{1}} < N_{v_{2}} \end{cases} \]  

To calculate the similarity score for attributes with numerical values: The similarity of two numerical attribute values \((v_{1}, v_{2})\) is based on the difference between two values.

\[ S_{\text{num}}(a_{1}, a_{2}) = \frac{1}{2} \left[ 1 + \frac{\sum_{(V_{1k}, V_{2l}) \in E^{p}_{1}} S_{\text{num}}(V_{1k}, V_{2l})}{\max(N_{v_{1}}, N_{v_{2}})} \right] \]  

where

\[ S_{\text{num}}(v_{1}, v_{2}) = \frac{|v_{1} - v_{2}|}{\max(v_{1}, v_{2})} \]

Determine for each PR (DR) rules in P1 (P2) which PR (DR) rules are similar using one-to-many \(\Phi\) mappings for each \(S_{\text{rule}}(r_{i}, r_{j})\). \(\varepsilon\) is the threshold value to set the acceptable level of similarity approximation between two policies.

\[ \Phi(r_{i}) = \{ r_{j} \mid S_{\text{rule}}(r_{i}, r_{j}) \geq \varepsilon \} \]

Compute the rule set similarity score for each (PR1, DR1) and (PR2, DR2) rules of policies P1 and P2 respectively that conform to \(\Phi\) mapping. Find the average of all the computed rule similarities of permit \((S_{\text{rule-set}}^{P})\) and deny \((S_{\text{rule-set}}^{D})\) rule sets.

\[ r_{S_{1i}} = \begin{cases} \frac{\sum_{r_{j} \in \varnothing P_{1}(r_{i})} S_{\text{rule}}(r_{i}, r_{j})}{|\varnothing P_{1}(r_{i})|}, & r_{i} \in PR_{1} \\ \frac{\sum_{r_{j} \in \varnothing D_{1}(r_{i})} S_{\text{rule}}(r_{i}, r_{j})}{|\varnothing D_{1}(r_{i})|}, & r_{i} \in DR_{1} \end{cases} \]  

\[ r_{S_{2j}} = \begin{cases} \frac{\sum_{r_{j} \in \varnothing P_{2}(r_{j})} S_{\text{rule}}(r_{j}, r_{i})}{|\varnothing P_{2}(r_{j})|}, & r_{j} \in PR_{2} \\ \frac{\sum_{r_{j} \in \varnothing D_{2}(r_{j})} S_{\text{rule}}(r_{j}, r_{i})}{|\varnothing D_{2}(r_{j})|}, & r_{j} \in DR_{2} \end{cases} \]
\[ S_{\text{rule-set}}^P = \frac{\sum_{i=1}^{N_{PR1}} r_{si} + \sum_{j=1}^{N_{PR2}} r_{sj}}{N_{PR1} + N_{PR2}} \]  

(12)

\[ S_{\text{rule-set}}^D = \frac{\sum_{i=1}^{N_{DR1}} r_{si} + \sum_{j=1}^{N_{DR2}} r_{sj}}{N_{DR1} + N_{DR2}} \]  

(13)

- Finally, compute the overall similarity score \((P_1, P_2)\) by summing the obtained average and overall policy target \((T)\). \(S_{\text{policy}}\) is the summation of a function \(S_T(P_1, P_2)\) that computes similarity between policy targets \((T)\) of \(P_1\) and \(P_2\) and similarity scores of matched permit \((S_{\text{rule-set}}^P)\) and deny \((S_{\text{rule-set}}^D)\) rule sets in \(P_1\) and \(P_2\). \(W_T, W_p\) and \(W_d\) are weights or values such that, \(W_T + W_p + W_d = 1\), to be associated with each attribute showing relative importance of its contribution in overall similarity estimation.

\[ S_{\text{policy}}(P_1, P_2) = W_T S_T(P_1, P_2) + W_p S_{\text{rule-set}}^P + W_d S_{\text{rule-set}}^D \]  

(14)

### 4.5.2 Case Study - Calculating Similarity Score between policies of two healthcare units

XACML access control policies exhibit relationships and dependencies on various attributes. These dependencies and relationships are mathematically exhibited using set theory.

Policy sets of hospital A and B: Hospitals comprise of Policy set (Pset) that is a collection of policies defined for each user accessing the resource in the hospital. Each policy in Pset constitutes rule-sets (Rset) defining subject (users of EHR), resource (EHR) and action (permission) attributes. Multiple actions may be defined for each combination of subject and resource attribute value.

\[ P_{\text{set}}(A) = \{AP_1, AP_2, AP_3, \ldots, AP_n\} \text{ such that } \forall i: (AP_i, i) \subset \text{Rule Set} (S, R, A, C) \]

\[ P_{\text{set}}(B) = \{BP_1, BP_2, BP_3, \ldots, BP_m\} \text{ such that } \forall i: (BP_i, i) \subset \text{Rule Set} (S, R, A, C, E) \]

Rule \((S, R, A, C)\) attributes in the Rule set state the logical combinations of subject accessing the resource for specified actions under stated conditions. The subject is a subset of all the users in the hospital. Similarly, resource, action and condition are the subsets of their respective domain values as defined by the hospital. Subject, Resource and Action Attributes in policy set:

\[ S(A) \supseteq Uset(A) \text{ and } S(B) \supseteq Uset(B) \]

\[ R(A) \supseteq EHRset(A) \text{ and } R(B) \supseteq EHRset(B) \]
A(A) \supseteq \text{Actset}(A), \text{ where } \text{Actset}(A) \Rightarrow \forall i \exists j: (\text{Actset}(A, a) =\{a_{ai1}, a_{ai2}, a_{ai3}, \ldots, a_{arij}\})

A(B) \supseteq \text{Actset}(B), \text{ where } \text{Actset}(B) \Rightarrow \forall i \exists j: (\text{Actset}(B, a) =\{b_{ai1}, b_{ai2}, b_{ai3}, \ldots, b_{rij}\})

**Conditional constraint (if defined) in each rule set of the stated policies:** The rules may impose conditional constraint(s) on allowing or denying access to the resource. C(A) is a subset of all possible conditions specified in hospital A and the same is applied in hospital B as well.

C(A) \supseteq \text{Condset}(A), \text{ where } \text{Condset}(A) \Rightarrow \forall i \exists j: (\text{Condset}(A, c) =\{aci_1, aci_2, aci_3, \ldots, aci_j\})

C(B) \supseteq \text{Condset}(B), \text{ where } \text{Condset}(B) \Rightarrow \forall i \exists j: (\text{Condset}(B, c) =\{bci_1, bci_2, bci_3, \ldots, bci_j\})

**Composition of rule-sets comprising of XACML attributes:** Pset is a superset of Rset where each Rset comprise of various rules stating clear access decisions under controlled conditions.

\( \forall i : (P_i, p) \supseteq \text{Rset}(A) =\{AR_1, AR_2, AR_3, \ldots, AR_i\}, \text{ where } \text{Rset}(A, AR_i) \Rightarrow \forall i \exists j: (AR_i) \subseteq (S_{ij}, R_{ij}, A_{ij}, C_{ij}) \)

\( \forall i : (P_i, p) \supseteq \text{Rset}(B) =\{BR_1, BR_2, BR_3, \ldots, BR_i\}, \text{ where } \text{Rset}(B, BR_i) \Rightarrow \forall i \exists j: (BR_i) \subseteq (S_{ij}, R_{ij}, A_{ij}, C_{ij}) \)

**Categorization of Rules**

Let \( sp =\{PR_1, PR_2, PR_3, \ldots, PR_n\} \) be all the permit rules and \( sd =\{DR_1, DR_2, DR_3, \ldots, DR_n\} \) be all the deny rules that collectively represent policies P1 and P2, such that:

P1 = \{sp_1 + sd_1\} and P2 = \{sp_2 + sd_2\}

**Categorization of Attributes**

Let S’, R’, A’ and C’ be the subset of S, R, A, and C attributes each having multiple attribute values \{a_{i1}, a_{i2}, a_{i3} \ldots, a_{ik}\} defined for any PR or DR rule.
For each rule $R_1i$ in $sp_1$ and $sd_1$, $R_1i \in S_{P_1} \subseteq S$ where $S_{P_1} = \{v11, v12, v13, \ldots, v1k\}$ such that $1 \leq k \leq n$

For each rule $R_1i$ in $sp_1$ and $sd_1$, $R_1i \in R_{P_1} \subseteq R$ where $R_{P_1} = \{v11, v12, v13, \ldots, v1k\}$ such that $1 \leq k \leq n$

For each rule $R_1i$ in $sp_1$ and $sd_1$, $R_1i \in A_{P_1} \subseteq A$ where $A_{P_1} = \{v11, v12, v13, \ldots, v1k\}$ such that $1 \leq k \leq n$

For each rule $R_1i$ in $sp_1$ and $sd_1$, $R_1i \in C_{P_1} \subseteq C$ where $C_{P_1} = \{v11, v12, v13, \ldots, v1k\}$ such that $1 \leq k \leq n$

**Calculating Similarity Score**

Using equations 1-13, evaluate Similarity Rule ($S_{rule}$) for each PR and DR rules of $P1$ in $P2$ and $P2$ in $P1$ respectively, and generate a set of all the permit rule set ($S_{rule-set}^P$) and the deny rule set ($S_{rule-set}^D$) according to the set range($\varepsilon$):

$S_{rule-set}^P(P1, P2)$, and $S_{rule-set}^D(P1, P2)$

Evaluate the Similarity Score of $P1$ and $P2$ using equation 14 and obtain $S_{Policy}(P1, P2)$.

**4.5.2.1 Access Control Policies**

**Policy 1 – Hospital A**

**Facility:** OPD

**Subject:** Family_Doc, Intern, Matron, Technical_staff, Patient

**Resource:** Clinical_data, Demographic_data, Financial_data

**Action:** Read, Write

**Condition:** Family_doc, Intern and Matron can access clinical data any time and can perform any action on it. Administrative staff can read and write the demographic details of the patient between fixed hours. The nomination and financial details of the patient can only be viewed by the HOD or Family_Doc only through patient authority. Patient has read permission on clinical data under the authorization of Family_doc. Access to read the Patient's clinical and nomination data is denied to its family if patient’s consent is not true. Patient is denied an access to clinical data while it is updated by health provider.
Environment: Controlling_authority- Patient, Family_Doc
Patient_Consent- True or False
Update - True or False
Time – A time range specifying the authorized access

Policy 2 – Hospital B
Facility: IPD
Subject: HOD, Family_Doc, Specialists, Staff Nurse, patient, Technical staff
Resource: Clinical_data, Demographic_data, Financial_data
Action: Read, write, print
Condition: The patient allows HOD, Family_Doc and Specialists to access its clinical details. The physical details can be accessed by Staff Nurse and Lab Technician only between 8.00 am to 4.00 pm. The patient has complete access rights on the demographic details and the financial details. Staff can access these details if the patient permits. Technical staff cannot access the physical details if the doctors or authorized personnel are updating the records. Patient is denied reading his clinical details while the doctors are updating his/her records.

Environment: Controlling_authority- Patient, Family_Doc
Patient_Consent- True or False
Update - True or False
Time – A time range specifying the authorized access
Considering the policies P1 and P2 describing the workflow of hospital A and B respectively, the similarity algorithm explained in section 4.8 was applied to find the similarity between P1 and P2. The hierarchies shown in figure 4.7 and figure 4.8 with respect to the users and resources accessed by the users are taken as common for both the hospitals. The following set of PR and DR rules were formed using XACML-based policy structure. The ACPs generated in XACML are listed in figures 4.9 and 4.10.

**XACML- Access Control Policies for Hospital A and Hospital B**

```
PolicyId=P1
 <PolicyTarget GroupName=HospitalA_OPD>
 <RuleId=R11 Effect=Permit>
 <Target>
 <Subject Designation belong_to{Family_Doc, Intern, Matron}>
 <Resource FileType belong_to{Clinical_Data}>
 <Action AccessType belong_to{Read, Write}>
 </Target>
 </Rule>
 <RuleId=R12 Effect=Permit>
 <Target>
 <Subject Designation belong_to {Med_Record_Officer, Executive} <Resource FileType belong_to{Demographic_data}>
 <Action AccessType belong_to {Read, Write}>
 <Condition 8.00 < time <17.00>
 </Target>
 </Rule>
 <RuleId=13 Effect=Permit>
 <Target>
 <Subject Designation belong_to {HOD, Family_doc}>
 <Resource FileType belong_to{Nomination, Financial_Data}>
 <Action AccessType belong_to {Read}>
```
Figure Error! No text of specified style in document.

XACML-based ACP for Hospital A

PolicyId=P2
<PolicyTarget GroupName=HospitalB_IPD>
  <RuleId=R21 Effect=Permit>
    <Target>
      <Subject Designation belong_to {HOD, Family_Doc, Specialist}>
      <Resource FileType belong_to {Clinical_Data}>
      <Action AccessType belong_to {Read, Write}>
      <Condition Controlling_authority=Patient>
    </Target>
  </Rule>
  <RuleId=R22 Effect=Permit>
    <Target>
      <Subject Designation belong_to {Staff, Self}>
      <Resource FileType belong_to {Physical}>
      <Action AccessType belong_to {Read, Write}>
      <Condition 8.00 < time < 16.00>
    </Target>
  </Rule>
  <RuleId=23 Effect=Permit>
    <Target>
      <Subject Designation belong_to {Staff, Self}>
      <Resource FileType belong_to {Demographic_data, Financial_Data}>
      <Action AccessType belong_to {Read, Write, Print}>
      <Condition controlling_authority=Patient>
    </Target>
  </Rule>
  <RuleId=R24 Effect=Deny>
    <Target>
      <Subject Designation belong_to {Technical_staff}>
      <Resource FileType belong_to {Physical_data}>
      <Action AccessType={read, write}>
      <Condition update=true>
    </Target>
  </Rule>
</PolicyTarget>
XACML policy sets (P1, P2) for A and B

P1 = \{ R11, R12, R13, R14, R15, R16 \}
sp1 = \{ R11, R12, R13, R14 \} and sd1 = \{ R15, R16 \}
P2 = \{ R21, R22, R23, R24, R25 \}
sp2 = \{ R21, R22, R23 \} and sd2 = \{ R24, R25 \}

4.5.2.2 Calculations

Similarity between the policies of P1 and P2 is calculated in a generalized manner. For instance, to find the \( S_{\text{rule}} \) for R11 and R21 in Policy P1 and P2, respectively:

\( S_{P_1} = \{ v11, v12, v13 \} \)
\( S_{P_2} = \{ v21, v22, v23 \} \)

On the basis of the hierarchical distance (if Subject (P1, P2) is a categorical predicate) between each value of \( S_{P_1} \) and \( S_{P_2} \), the matrix in table 4.10 is obtained.

<table>
<thead>
<tr>
<th>P2/P1</th>
<th>v11(Family_Doc)</th>
<th>v12(Intern)</th>
<th>V13(Matron)</th>
</tr>
</thead>
<tbody>
<tr>
<td>v21(HOD)</td>
<td>0.67</td>
<td>0.67</td>
<td>0.33</td>
</tr>
<tr>
<td>v22(Family_Doc)</td>
<td>1</td>
<td>0.67</td>
<td>0.33</td>
</tr>
<tr>
<td>v23(Specialist)</td>
<td>0.67</td>
<td>0.67</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Matching Pairs: (v11, v22)

Calculate compensating score
From Table 4.10, highest similarity score is identified as, v12, v21 (0.67), v12, v23 (0.67), v13, v21 (0.33) and v13, v23 (0.33). Select any one that is highest: v12, v21 (0.67). For other lower values or non-matching scores, calculate δ, such that:

δ for (v13, v23) = 0.33

Calculate Ss (Similarity between Subjects (P1, P2))

\[
\frac{1}{2} \left[ 1 + \frac{1 + 0.67 + 0.33}{3} \right] = \frac{1}{2} (1 + \frac{2}{3}) = 0.83
\]

Calculate Sr (Similarity between Resource (P1, P2))

\[ R_{P1} = \{ v11 \} - \text{Clinical_data} \]
\[ R_{P2} = \{ v21 \} - \text{Clinical_data} \]

As both are matching pairs depicting same values, the average similarity score for this rule is 1.

Calculate Sa- (Similarity between Action (P1, P2))

Table 4.11 shows the similarities between Action attributes of P1 and P2 of hospital A and B respectively.

\[ A_{P1} = \{ v11, v12 \} \]
\[ A_{P2} = \{ v21, v22 \} \]

<table>
<thead>
<tr>
<th>P2/P1</th>
<th>v11 (read)</th>
<th>v12 (write)</th>
</tr>
</thead>
<tbody>
<tr>
<td>v21(read)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>v22(write)</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

\[
\frac{1}{2} \left[ 1 + \frac{1 + 1}{2} \right] = 1
\]

Keeping the weights constant, calculate St ((Similarity between Target (P1, P2)):

\[
S_t = \frac{S_s + S_r + S_a}{3} = \frac{0.83 + 1 + 1}{3} = 0.94
\]

S_c- (Similarity between Condition (P1, P2))
The condition element is Controlling_authority which is Nil in P1. Hence, the similarity score is 0.5

**Calculating Srule**

\[
S_{\text{rule}}(R11, R21) = \frac{(S_t + S_c)}{2} = \frac{(0.94 + 0.5)}{2} = .72
\]

\( S_{\text{rule}} \) were obtained through similarly calculating other permit and deny rules of P1 and P2.

**Combining Rules using \( \Phi \) mappings**

Determine similarity between each pair of \( S_{\text{rule}}(r1_i, r2_j) \) using \( \Phi \) mappings where \( \epsilon \), the threshold value ascertaining the acceptable level of similarity approximation between two policies is set to 0.7. The pairs that satisfied the approximation are listed below.

**\( \Phi \) mappings for P1**

\[
\Phi^P_1 = \{R11 \rightarrow \{R21\}, R12 \rightarrow \{\}, R13 \rightarrow \{R21, R23\}, R14 \rightarrow \{R23\}\}
\]

\[
\Phi^D_1 = \{R15 \rightarrow \{R25\}, R16 \rightarrow \{R24, R25\}\}
\]

**\( \Phi \) mappings for P2**

\[
\Phi^P_2 = \{R21 \rightarrow \{R11, R13\}, R22 \rightarrow \{\}, R23 \rightarrow \{R13, R14\}\}
\]

\[
\Phi^D_2 = \{R24 \rightarrow \{R16\}, R25 \rightarrow \{R15, R16\}\}
\]

**Average all the matching pairs**

For each rule \( R1_i \) in P1, the corresponding rule similarity score \( RS_{1_i} \) is computed

\[
\begin{align*}
R11 \rightarrow R21 &= 0.72 & R14 \rightarrow \{R23\} &= .86 \\
R12 &= 0 & R15 \rightarrow \{R25\} &= .71 \\
R13 \rightarrow \{R21, R23\} &= 1/2(0.92+.91) = .91 & R16 \rightarrow \{R24, R25\} &= 1/2(.93+1) = .96
\end{align*}
\]

For each rule \( R2_j \) in P2, the corresponding rule similarity score \( RS_{2_j} \) is computed

\[
\begin{align*}
R21 \rightarrow \{R11, R13\} &= 1/2(0.72+.92) = .82 & R24 \rightarrow \{R16\} &= 0.93 \\
R22 &= 0 & R25 \rightarrow \{R15, R16\} &= 1/2(0.71+1) = .85
\end{align*}
\]
R23 \( \rightarrow \) \{R13, R14\} = 1/2(.91+.86) = .88

**Generate Permit Rule set and Deny Rule Set**

\[ S^{P}_{\text{rule-set}} = (.72+.91+.86+.82+.88)/7 = 0.60 \]

\[ S^{D}_{\text{rule-set}} = (.71+.93+.96+.85)/4 = 0.86 \]

**Calculating the overall Policy Similarity Score**

The similarity rule between Policy Targets (ST) is calculated in the same way as that of rule target: \( ST = 0.7 \)

\[ S_{\text{Policy}} = 1/3(0.7) + 1/3(0.6) + 1/3(0.86) = .68 \]

**4.5.2.3 Observation**

Access control decisions can be expressed by identifying an authorization through the combinations of matched and unmatched attribute values of both rules. The similarity algorithm applied to multi-level user and resource hierarchy reveals certain patterns that can determine authorization to permit or deny access to the data. 4x3 Permit Rule Pairs and 2x2 Deny Rule Pairs in P1 and P2 accumulated to 12 permit rules and 4 deny rules respectively. Applying similarity score algorithm, the rules were reduced to 5 permit and 3 deny rule pairs. Pruning of rules permits to further examine the similar role-pairs for authorizations controlling access to EHRs while sharing between different hospitals. Based on the observations and the derived results, properties are formalized that can be implemented to verify the rule-similarity of two distinct policies and reflect the authorization decisions for both permit and deny effects.

The domain (Dom) is generic way is defined as a pool of values belonging to the same category of data. In this context, the domains considered are Subject (S), Resource (R) and Action (A) associated with each rule (R1, R2, ……., Rn) defined in the access control policy.

**Property 1: Complete Domain and Complete Conditional Equality:**

\[ \{((\text{Dom} (a1i) \in R1i = \text{Dom} (a2j) \in R2j) \land (\text{Val} (a1i) \in R1i = \text{Val} (a2j) \in R2j)) \land ((C_{P1} \in R1i) = (C_{P2} \in R2j))\} \]
If the attributes belong to same domain and represent complete similarity between the values, the existence of matching conditions in both policies is 96%. It is predicted that wherever the data access is controlled by a defined authority, the policies tend to be similar, thus achieving secured sharing of data.

**Property 2: Complete Domain and Partial Conditional Equality:**

\[\left(\text{Dom}(a_{1i}) \in R_{1i} = \text{Dom}(a_{2j}) \in R_{2j}) \land (\text{Val}(a_{1i}) \in R_{1i} = \text{Val}(a_{2j}) \in R_{2j})) \land (\text{C}_P \cap \text{R}_{1i} \neq \{\emptyset\})\right]\]

If the attributes of same domain depict total similarity between their values and a conditional constraint is defined in at least one policy, there exists a probability of 72% that the two policies can be integrated.

**Property 3: Partial Domain and Complete Conditional Equality:**

\[\left(\text{Dom}(a_{1i}) \in R_{1i} \neq \text{Dom}(a_{2j}) \in R_{2j}) \lor (\text{Dom}(a_{1i}) \in R_{1i} = \text{Dom}(a_{2j}) \in R_{2j})) \land (\text{Val}(a_{11} \cup a_{12} \cup \ldots \cup a_{1k}) \in R_{1i}) \cap (\text{Val}(a_{21} \cup a_{22} \cup \ldots \cup a_{2j}) \in R_{2j}) \neq \{\emptyset\}) \land (\text{C}_P \cap \text{R}_{1i} = \text{C}_P \cap \text{R}_{2j})\right]\]

If exist equality in only some domains with similarity in some attribute values and exhibit complete conditional equality in both the policies, it is observed that the major contributor to the permissible collaboration of the two policies is the conditional constraints. The existence of conditions in both policies guarantees 91% possibility that the two policies would integrate under well-defined authorizations and enable controlled and viable access to the data.

**Property 4: Partial Domain and Partial Conditional Equality:**

\[\left(\text{Dom}(a_{1i}) \in R_{1i}) \cap (\text{Dom}(a_{2j}) \in R_{2j}) \neq \{\emptyset\}) \land (\text{Val}(a_{1i}) \in R_{1i} \neq \text{Val}(a_{2j}) \in R_{2j})] \land [((\text{C}_P \in \text{R}_{1i}) \neq (\text{C}_P \in \text{R}_{2j})) \land ((\text{C}_P \in \text{R}_{1i}) \cap (\text{C}_P \in \text{R}_{2j}) \neq \{\emptyset\})]\]

In spite of partial domain equality resulting in partial similarity of attribute values and existing but dissimilar conditional constraints, there exists a clear indication of 86% possibility that the two hospitals can collaborate and gain access to the data.

**Property 5: No Domain and No Conditional Equality:**
In cases where there is a dissimilar domain, the similarities achieved show that the attribute values completely or partially match with each other. No clear authorization is visible, as no conditions match. This indicates (0.82) facilitating decision-making for allowing the health professionals to share the data.

The research provides a solution for a distributed environment where two hospitals are collaborating and sharing healthcare data under proper control over its access. The access control policies are compared to provide a permit and deny access to the required data. For all permits and deny rules in two different policies, similarity score is calculated and a reduced set of permit and deny rules is obtained. The obtained rules identify the generic properties to determine authorizations in the course of sharing of healthcare data between legitimate and authorized stakeholders. Partial similarity between some values of similar domain attributes and non-matching conditional constraints with no similarity between attributes belonging to different domains indicate the probability of determining authorizations in such disparate environment. The rule similarities obtained signify a possibility of secured sharing between two hospitals through well-defined authorizations. Such a model provides a viable solution in an interoperable environment.

### 4.5.2.4 Interpretation

A final rule set is the combination of rules that is obtained on merging of access control policies of two hospitals. Given two sets of attributes, \( \{(S_{p1}, R_{p1}, A_{p1}, C_{p1}) \in R1i\} \) of P1 and \( \{(S_{p2}, R_{p2}, A_{p2}, C_{p2}) \in R2j\} \) of P2, the obtained rule pairs satisfying the authorizations is stated below.

\[ R11 \rightarrow R21 \]

The resource and action attributes in both rules have similar values but differ in subject attributes. Authorization to access data is explicitly stated in P2, whereas no such restriction is imposed in P1. The two policies are similar on the basis of domain equality. Hence,
policies P1 and P2 are similar, iff there exist PR1 and PR2 rules exhibiting same domain with similar or different attribute values.

R13 $\rightarrow \{R21, R23\}$

The subject, resource and action attribute values are partially similar in P1 and P2. No similarity is observed between attributes belonging to different attribute domains. The authorizations controlling access are clearly stated in both the policies and exhibit complete similarity. The policies are similar on the basis of Conditional Equality. All other attribute values are dissimilar indicating defined authorization as the only factor allowing the two hospitals to collaborate and share the resources. Hence, policies P1 and P2 are similar, iff there exist PR1 and PR2 rules exhibiting same conditional authority with similar attribute values.

R21 $\rightarrow \{R11, R13\}$

Another very interesting result is obtained by merging of these two rules. The condition attributes are contributing parameter enabling to take a justified and logical decision in such a scenario. As compared with R21, with the domain as same, partial subject similarity exist in R11 and a proper subset in R13. The resource similarity is obtained in R11 where the subject domain is same and resource is different in R13 due to resource values belonging to different domains. The action attribute show a partial similarity between the pairs. Condition is not defined in R11 and is same in R13. Hence, policies P1 and P2 are similar, iff there exist PR1 and PR2 rules exhibiting same domain with no or partially similar conditional authority.

R23 $\rightarrow \{R13, R14\}$

The subject attributes belong to different domains and only partial similarity exist in the resource attributes. The action has one value in common whereas the condition is same in one R13 but exhibit same domain. A much complex relationship is revealed between the resulting rule pairs. There exist no single attribute that can act as a deciding parameter for determining authorization. On analyzing each attribute pair, it is found that a combination of the subject and the conditional constraint plays an important role in determining the rule-similarity. Hence, policies P1 and P2 are similar, iff there exist PR1 and PR2 rules exhibiting different domain with partially similar conditional authority.
R14 $\rightarrow$ R23

The subject and action attribute values are partially similar in P1 and P2. No similarity is observed for resource and conditional attributes. A unique finding of these matching pairs highlights conditional constraint as a vital contributor in deciding the relevant authorization. Hence, policies P1 and P2 are similar, iff there exist PR1 and PR2 rules exhibiting partial similarity in subject and action domain.

R15 $\rightarrow$ R25

The subject attributes belong to different domains and only partial similarity exist in the resource attributes. The action has one value in common and no similarity exists in the conditional constraints. It identifies that if exists any two attributes showing partial similarity within the same domain, there exist a possibility of obtaining logically devised authorization for taking a logical decision for accessing the data by the users.

R16 $\rightarrow$ \{R24, R25\}

All the subject, resource, action and condition values are exactly similar in R25 but show the difference of subject, resource and action domain. The action has one value in common. The conditional constraints show a complete similarity between the two rules of P1 and P2. Hence, policies P1 and P2 are similar, iff there exist PR1 and PR2 rules exhibiting complete or no similarity in subject and resource and action domain with complete similarity between all three rules.

4.6 Gaps in the existing Algorithm

The algorithm [183] iterates through entire policy set to calculate the similarities between each rule attribute irrespective of the data under query. It determines the closeness of two policies concluding the probability that the two policies can securely integrate with each other. Identification of this similarity is not an assurance of having similar or acceptable authorization to allow integration of selected policies. The heterogeneity in organizational hierarchy is another challenge that may turn the similarity score upside down. In healthcare domain, one user works in different role-capacity at any given point of time. For ex., a doctor may be a primary doctor for one patient and act as specialist for another patient. The access rights must differ in both the situations. Moreover, the hierarchy and authority is not static as
for ex., a doctor can work as a member of team controlled by a consultant or can work independently in the same hospital. The accountability and responsibility differ in both the situations. Hence, it is challenging to determine authorized and consolidated access to EHRs.

Integrating access control policies of different hospitals may give rise to policy conflicts due to disparity in access privileges of health data users. Another problem while integrating the policies is manifold increase in the number of rules. Handling each rule and preventing security breach is highly difficult and computationally expensive. It further exhibit rule redundancy, thereby resulting in waste of storage space and other computational resources. Policy conflicts and redundancies needs to be resolved before integrating the policies to achieve availability and confidentiality in policy integration process.

Table Error! No text of specified style in document.12: Notations (Source: Lin et al. [183])

<table>
<thead>
<tr>
<th>Notation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>Policy</td>
</tr>
<tr>
<td>PR</td>
<td>Permit rule set</td>
</tr>
<tr>
<td>DR</td>
<td>Deny rule set</td>
</tr>
<tr>
<td>R</td>
<td>Rule</td>
</tr>
<tr>
<td>A</td>
<td>Attribute</td>
</tr>
<tr>
<td>V</td>
<td>Attribute Value</td>
</tr>
<tr>
<td>H</td>
<td>Height of the Hierarchy</td>
</tr>
<tr>
<td>$S_{policy}$</td>
<td>Similarity Score of two policies</td>
</tr>
<tr>
<td>$S_{rule}$</td>
<td>Similarity Score of two rules</td>
</tr>
<tr>
<td>$S^p_{rule-set}$</td>
<td>Similarity Score of two sets of permit rules</td>
</tr>
<tr>
<td>$S^d_{rule-set}$</td>
<td>Similarity Score of two sets of deny rules</td>
</tr>
<tr>
<td>$S_{(Element)}$</td>
<td>Similarity Score of Elements (Element)$\in {‘T’,’t’,’c’,’s’,’r’,’a’}$</td>
</tr>
<tr>
<td>$S_{cat}$</td>
<td>Similarity Score of two categorical values</td>
</tr>
<tr>
<td>$S_{pred}$</td>
<td>Similarity Score of two categorical predicates</td>
</tr>
<tr>
<td>$R_s$</td>
<td>Similarity Score between a rule and a policy</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>Rule Mapping</td>
</tr>
<tr>
<td>$M_a$</td>
<td>Set of pairs matching attribute names</td>
</tr>
<tr>
<td>$M_v$</td>
<td>Set of pairs matching attribute values</td>
</tr>
<tr>
<td>$N_{PR}$</td>
<td>Number of permit rules in a policy</td>
</tr>
<tr>
<td>$N_{DR}$</td>
<td>Number of deny rules in a policy</td>
</tr>
<tr>
<td>$N_a$</td>
<td>Number of attribute in an element</td>
</tr>
<tr>
<td>$N_v$</td>
<td>Number of values of an attribute</td>
</tr>
<tr>
<td>$S_{Path}$</td>
<td>Length of shortest path of two categorical values</td>
</tr>
<tr>
<td>$w_{(Element)}$</td>
<td>Weight of similarity scores of elements,Element$\in {‘T’,’t’,’c’,’s’,’r’,’a’}$</td>
</tr>
<tr>
<td>$E$</td>
<td>Rule similarity threshold</td>
</tr>
<tr>
<td>$\Delta$</td>
<td>Compensating score for unmatched values</td>
</tr>
</tbody>
</table>