International Journal of Advanced and Applied Sciences, 4(1) 2017, Pages: 74-83

Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Role-based efficient information extraction using rule-based decision tree

Imran Khan [1,*], M. Sher [1], Syed M. Saqlain [1], Husnain A. Naqvi [1], Anwar Ghani [1], M. Usman Ashraf [2], Javed I. Khan [3]

[1]Department of Computer Science and Software Engineering, International Islamic University, Islamabad, Pakistan
[2]IBMS, University of Agriculture, Faisalabad, Pakistan
[3]Department of Computer Science, Kent State University, Kent Ohio, USA

A R T I C L E I N F O

Article history:
Received 27 October 2016
Received in revised form
27 December 2016
Accepted 7 January 2017

Keywords:
Privacy rules
Logical rules set
Formalization
Health care
EHR
Rule base decision tree

A B S T R A C T

This article presents a framework to reduce the comparison complexity required to evaluate requests for releasing Protected Health Information (PHI). A new methodology is introduced to divide HIPAA (Health Information Portability Accountability Act) into small independent integrate-able modules to facilitate the implementation process. The HIPAA World Rule Model is used for decision using formalized legal text. In order to reduce the time complexity of logical rule set comparison process for Role/Actor based approach, RBDT and Rules Filtering Algorithm are used. It reduces the time complexity of rule generation process from $O(n^5)$ to $O(n)$ for producing quick responses to access requests. The achieved results show significant improvement even for huge data.

1. Introduction

Individual's health information is personal information that cannot be disclosed without consent of the concerned individual, except in certain cases. As a result, the U.S. department of Health and Human Services (HHS) proposed privacy rules to be implemented in HIPAA. These rules govern the use and disclosure of individual's health information which is also referred as PHI. This information is managed by covered entities which are health plan, health care provider, or a clearinghouse, and enforced by the Office for Civil Rights (OCR). Flow of PHI within one or among multiple covered entities and others is one of the prime goals of HIPAA privacy rules. These rules allow the use and disclosure of PHI to further advance quality health care or maintain public health while protecting privacy of individuals (Breaux and Antón, 2008). HIPAA privacy rules are composed of legal text and this makes these rules complex in nature. Enforcing these rules would require substantial efforts if no form of automated information system is used.

The current implementation of HIPAA privacy rules by covered entities doesn't fully comply with these rules. For example, individuals do not have the opportunity to disclose only certain parts of their PHI like lab results. They can only disclose or deny the use of their PHI as an identical object. This might impose some violations of HIPAA privacy rules and could result in a fine of up to $25,000 per year for civil rights, criminal penalties that could range from $50,000 to $250,000 and prison between 1 to 10 years in jail. However, using a systematic way to formalize HIPAA privacy rules to logic forms in order to conform to them has been considered as a solution by several research papers (Maxwell and Antón, 2009; Lam et al., 2009). But the solutions to the problems of integrating these rules and tying them together haven't reduced the complexity of getting the right decision to whether allow or deny access to PHI.

It is difficult for individual to understand security and privacy policies and too complex for software engineers to design and implement systems. Analysts have to define mechanisms to disambiguate the regulations so those can be clearly specified as software requirements, to ensure that these systems comply with the policy (Breaux et al., 2006).

In a developer's point of view, the main challenge is to understand and define system requirements, given HIPAA's legal language (Maxwell and Antón, 2010). The need for such understanding is not unique to HIPAA, but it is particularly difficult in this case because organizations are performing their own

* Corresponding Author.
Email Address: iman.khan@iiu.edu.pk (I. Khan)
https://doi.org/10.21833/ijaas.2017.01.011
2313-626X/© 2017 The Authors. Published by IASE.

HIPAA interpretations. It is expected that healthcare firms will pay out $17.6 billion in the next few years to bring their systems and procedures in compliance with HIPAA (Massey et al., 2010). The Fig. 1 shows flow of information in the usual healthcare system.

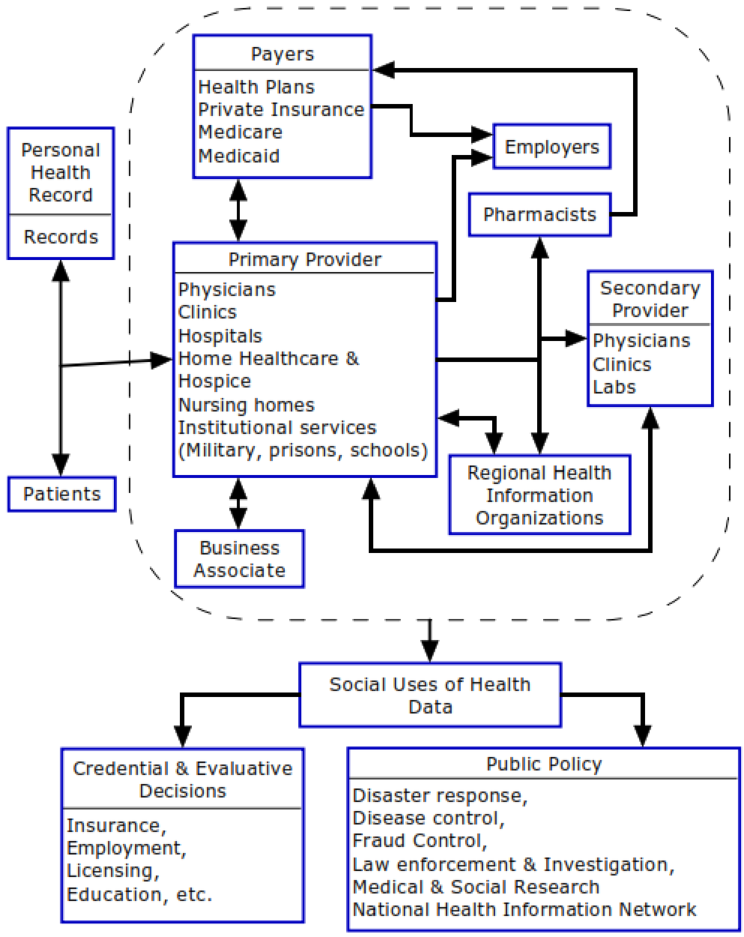

Fig. 1: Information flow in the healthcare system

Patient medical records could be used for various purposes, not only for the diagnosis and treatment facility e.g. those can be used for the efficiency and improvement within the healthcare system, for making the public policies, to conduct surveys and research for the advancement in medical science (Hodge, 2003). Patient records could also be shared with the payers, such as Insurance companies, Medicaid or Medicare for the justifications of payment related services which are done by the physicians. Healthcare providers may also use such information to manage their operations for improved service quality. Healthcare providers share information with the regional health information organization. Medical information is also used by the government for the public health management, medical research, hospital certification, public policy development and amendment and for the social and welfare system management. All these Actors in HIPAA are involved in the flow of such information.

1.1. Actors/roles in HIPAA

There are number of actors in HIPAA that participate in different kinds of activities, each actor has a distinct role and purpose in line with the regulations of HIPAA. These actors play specific roles in HIPAA as they are connected directly and/or indirectly. Around the globe there are different kinds of laws, why we need these laws? It can be answered by saying that all the participants of some event constrained by a specific law have to play their roles purposefully subject to some kind of condition(s). These participants might be the organizations or individuals. So if we see HIPAA rules closely we find that there are different kinds of actors involved, for whom the HIPAA rules are made e.g. Hospitals, Health Plan Providers, Insurance Companies, Patients, Doctors, Nurses, Law Enforcement Agencies and many more. All HIPAA rules are related to such actors ensuring execution of their work liable to some purpose(s) and/or condition(s), these actors communicate with each other in harmony with the set condition(s) for some specific purpose(s).

The definition of actors/stakeholders can be observed from HIPAA in 160.103, 164.500 and 164.501. The roles related to these Actors can be observed from HIPAA in 164.502 to 164.532. Detail and the maximum number of actors/stakeholders that are involved in HIPAA are explained in (Khan et al., 2012a). The rules related to these actors are present in different parts of sections of HIPAA and all the actors are connected with each other directly or indirectly.

The purpose for which data may be required for action is very important, because these purposes have even more issues which are very important for the information security (Ashley et al., 2002a; 2002b; Byun et al., 2005). Usually through the Role Based Access Control System (Sandhu et al., 1996), actors/stakeholders are allowed or denied to get the information which is related to their job function known as role. These roles are basically assigned to the Users/Actors where purposes are related to the data that is used for different kinds of transactions. For example if an Actor wants to access the data it must imply data purpose (reason for data usage), so purposes are the constraints with actors subject and the targeted objective. If one actor (requester) wants access to some kind of information from another actor (responder), the responder after checking the purpose and condition(s) will take an action accordingly and sends the response as output to the requester. The high level flow of information in accordance with the HIPAA rules for actors is shown in Fig. 2.

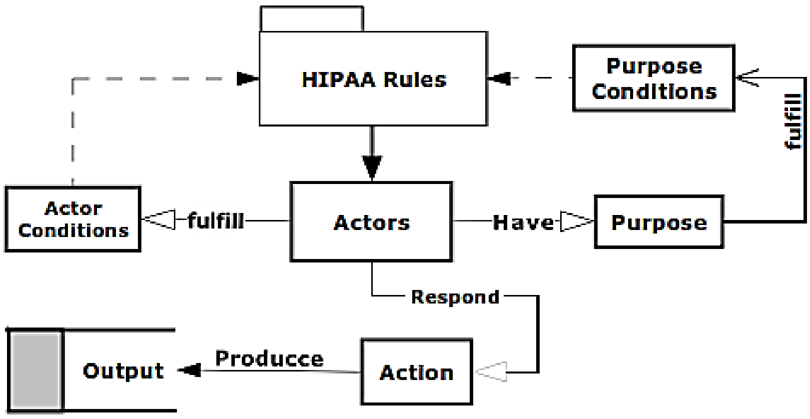

Fig. 2: Actor's rules information flow

In this paper we present how to reduce the complexity for the extraction of logical rules using WRM and RBDT. In the remainder of this paper section 2 briefly presents request flow procedure and steps to translate HIPAA rule set into a DDT (Khan et al., 2012b). Section 3 explains the complexity of HIPAA. Section 4 then explains the process of rules filtration and decision on medical data request processed through World Rules Model for HIPAA, also compares the complexity of algorithms and at the end of the same section an example for the researcher is explained.

2. Materials and methods

The rules that are defined in HIPAA are explained by the actors' responsibilities, purposes, conditions and actions (Wilson, 2006). All rules have been placed in different sections of HIPAA, if we wish to find the rules related to one actor, we have to check all the sections of HIPAA. For example there are many rules in one section of HIPAA that are referring to other sections of HIPAA for the same actor, hence making it easy to relate such rules collectively with a particular actor. On the other hand there are some rules those are not referred by other rules, but it does not lessen their importance even if they stay in a different section.

HIPAA complexity increases very much as there are many types of purposes for the data, and these purposes are constrained with the Actor's request. Each actor has to qualify that the specific actor is eligible for a given purpose or not. The examples of purposes in HIPAA include Treatment, Payment, Health care operations, Health care fraud, Abuse detection, Compliance, Billing, Access and Amendment etc. All these purposes are appertaining to different Actors. For normal request in HIPAA according to (Khan et al., 2013) one has to follow multiple steps for the response.

Step 1: Check from the list of Actor/Requester, who is requesting. i.e. **ReqT**

Step 2: Check what are the purpose(s) of the Actor from the list of Purposes. i.e. **PPT**

Step 3: Check for the requested Patient Record Item (PRI) from the list of PRI. i.e. **PRI**

After step 3 for further processing we need to process the request in accordance with the information obtained as a result to these three steps, then initially verify that the requester is eligible for the mentioned purpose for which data have been requested?, next for that purpose the requested PRI is needed or not?, then to verify whether PRI is accessible by the particular actor?, it is also verified that none of PRI conditions is contradictory to the actor's authority. For processing we need to multiply three values in order to get the total no. of comparisons required i.e. ReqT * PPT * PRI

Step 4: Check from the list of Conditions that are applied on the requesting actor i.e. CT hence the complexity further increases as:

ReqT * PPT * PRI * CT

Step 5: Check what possible action(s) may be taken in compliance to fee and time constraint for Purposes involved. i.e.

AT * TFT so **ReqT * PPT * PRI * CT + (AT * TFT)**

Step 6: Check for the Special Instruction how to Release the record from the list of Information Procedure. i.e. **IPT**

Step 7: Check for the Record Release what to Release from the list of Record Release Procedure. i.e. **RRT**

Finally the complexity of comparison for generating the response can be presented as shown in the following equation and flow chart in Fig. 3.

(ReqT * PPT * PRI * CT + (AT * TFT)) * IPT * RRT

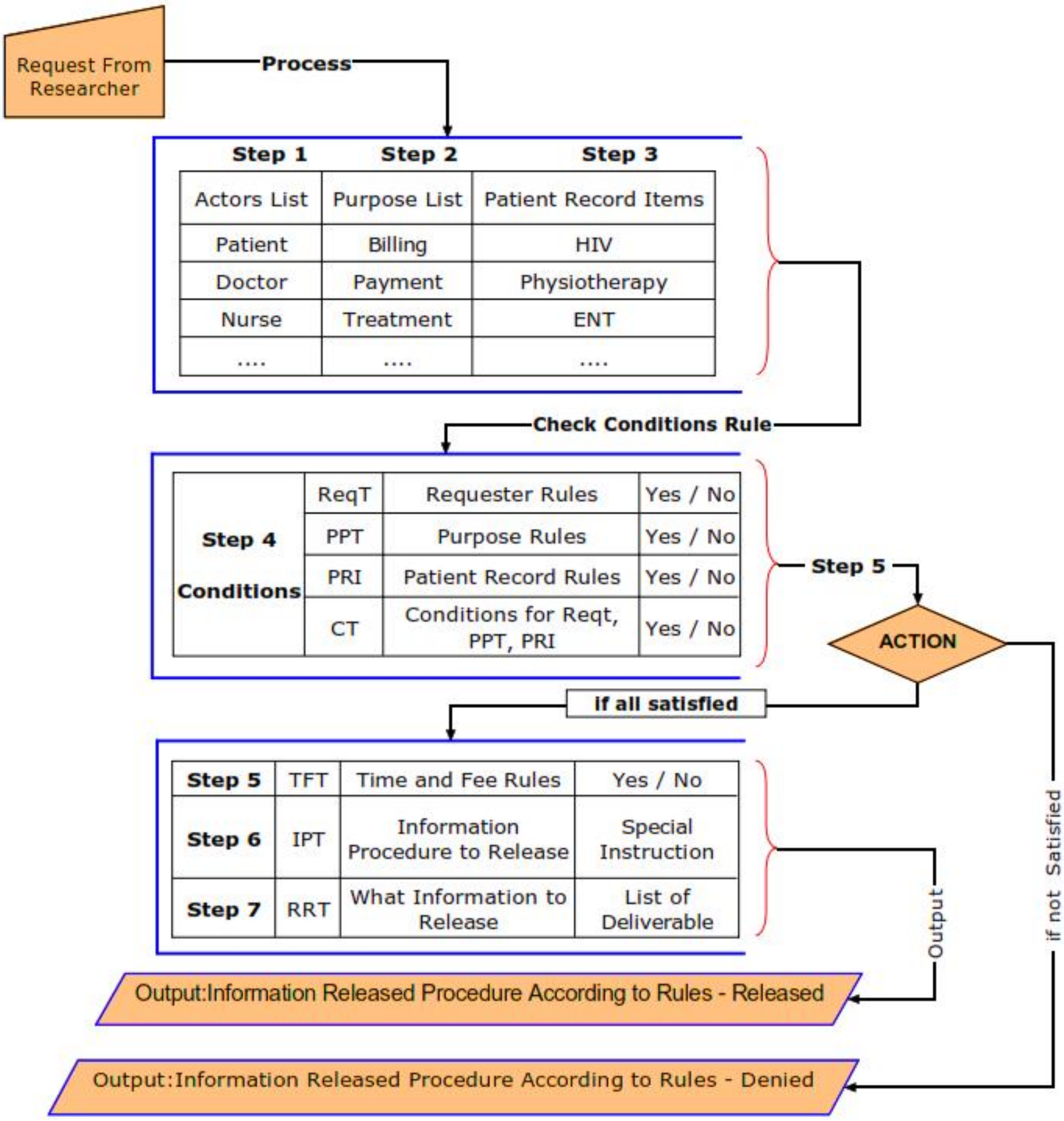

Fig. 3: Request flow for verification with HIPAA rules

2.1. Rules in HIPAA sections

In section 164 of HIPAA which covers most of the security and privacy related standards for exchanging PHI. It consists of 683 non repeated clauses starting from 164.502 up to 164.530, according to HIPAA Administrative 45 CFR. The total number of clauses and sub clauses that are used as a rule of each section are shown in Table 1. But in formalization process the rules will be increased dramatically, some time to explain the rule according to (Khan et al., 2013) modal; a single rule has to be broken into multiple rules for better understanding. For example in (Massey et al., 2010) the authors generate 263 rules from sections 164.520 to 164.526 (i.e. from Table 1: S. No. 8 to 11) according to their modal, while actually there are only 156 rules.

Similarly we find that there are only 683 rules that are from Section 164.502 to 164.534 if we go through clause by clause, but we have more than the total number of rules in our case. For example we have calculated the number of rules in 164.506 in Table 1: i.e.10 with clauses and sub clauses which are related to the treatment, payment and health care operation rules for disclosing the PHI, but the number of rules that can be generated from section 164.506 according to (Khan et al., 2013) is 16. In this section there is a referenced section i.e. 164.508, which is related to authorization.

2.2. Referenced and unreferenced rules for an actor

This is an important process of the approach which we adopt to link all rules which are related to actor that are referenced or unreferenced from the different sections of HIPAA. In this way it's easy to take the decision by the actor about the kind of information that can be released. For example an actor sends the request to another actor for information to release; this request will be checked in all the sections of HIPAA even if there exists no reference for that particular actor for this request. See Fig. 4.

Table 1: No. of Rules in HIPAA Privacy Section 164.502 to 164.534

S. No	HIPAA Section	No. of Rules
1.	164.502 General rules	-
2.	164.504 Organizational requirements	67
3.	164.506 To Carry out treatment, payment or health care operations	10
4.	164.508 For which an authorization is required	55
5.	164.510 Opportunity for the individual to agree or to object	26
6.	164.512 Authorization or opportunity to agree or object is not required	160
7.	164.514 Other requirements relating to uses and disclosures PHI	102
8.	164.520 Notice of privacy practices for PHI	67
9.	164.522 Rights to request privacy protection for PHI	21
10.	164.524 Access of individuals to PHI	37
11.	164.526 Amendment of PHI	31
12.	164.528 Accounting of disclosures of PHI	44
13.	164.530 Administrative requirements	45
14.	164.532 Transition provisions	13
15.	164.534 Compliance dates for initial implementation	5

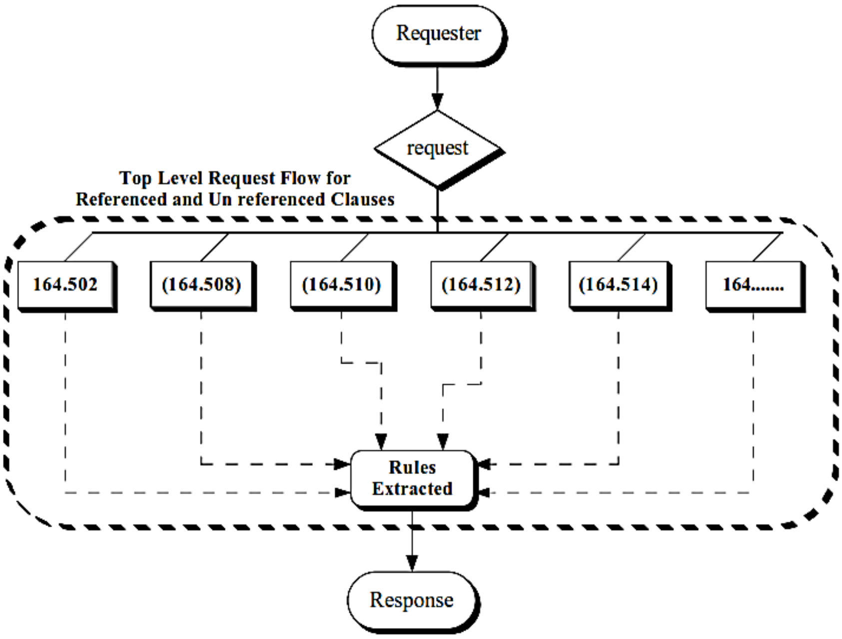

Fig. 4: Evaluation of requests between different sections

For example we have an actor Researcher who wants to get the PHI from another actor "Covered Entity", for that purpose all the rules related to Researcher will be checked by Covered Entity and there are some rules for the Researcher in section 164.508, 164.512, 164.514, that are referred but there are some rules in sections 164.532 those are not referred by any Researcher related rule. For that we have to collect all the rules related to researcher from all sections and then take the decision, so by this approach we have covered all the unreferenced rules for a particular actor.

3. Complexity of HIPAA

In (Khan et al., 2013) we have developed the World Rule Model, which has been used for data releasing decisions on the entire formalized legal text of HIPAA. Privacy rules are formalized from legal text into 8 different tags or classes. Then, the tags are analyzed, converted to "YES" or "NO" questions using the Rule Base Decision Tree (RBDT). These 8 classes are (ReqT, PPT, PRI, CT, AT, TFT, IPT and RRT). Fig. 3 shows the process of generating data for the construction of RBDT. Next, information retained in transponder will link requester with related tags logically based on requester. Whereas, transponder will work as a compliance checking agent against the requested information, and to make decision about how information is released by checking information procedures and then responds to the requester according to HIPAA privacy rules.

Formalizing requires tremendous effort, time and a lengthy process of data analysis. In addition, dedicated human power to implement this job is expensive and could cost a lot of money. Meanwhile, if the entire legal text is formalized, the problem of querying data produces undesirable false hits. In other words, using the approach proposed in (Khan et al., 2013) would require extensive comparisons with rules that might not produce a single successful hit and this would consume resources in an inefficient way. It happens for different reasons. First of all there are number of actors those are involved in HIPAA, secondly there are number of purposes to request the data in HIPAA by these actors from each other, similarly for the request there are number of conditions applied on each request according to the purpose of request for that actor. Thirdly there are number of patient related record items that are requested by these actors and some of these record items need to provide with some conditions, some of them are totally denied and some of them may be disclosed without any conditions. So these extra conditions have to be resolved before releasing the record internally from the records of related patients. So here we can imagine that after formalization of HIPAA we have to resolve the high level of comparison of rules in one request among Actors, Patient Record Items (PRI), Purposes, Conditions, Conditional record Items, Conditionals Purposes and etc.

Fig. 4 explains about the reference and cross reference issues in HIPAA, but if we analyze the diagram by the requester's point of view, the related rules in different sections of HIPAA are formalized according to WRM shown in Fig. 5.

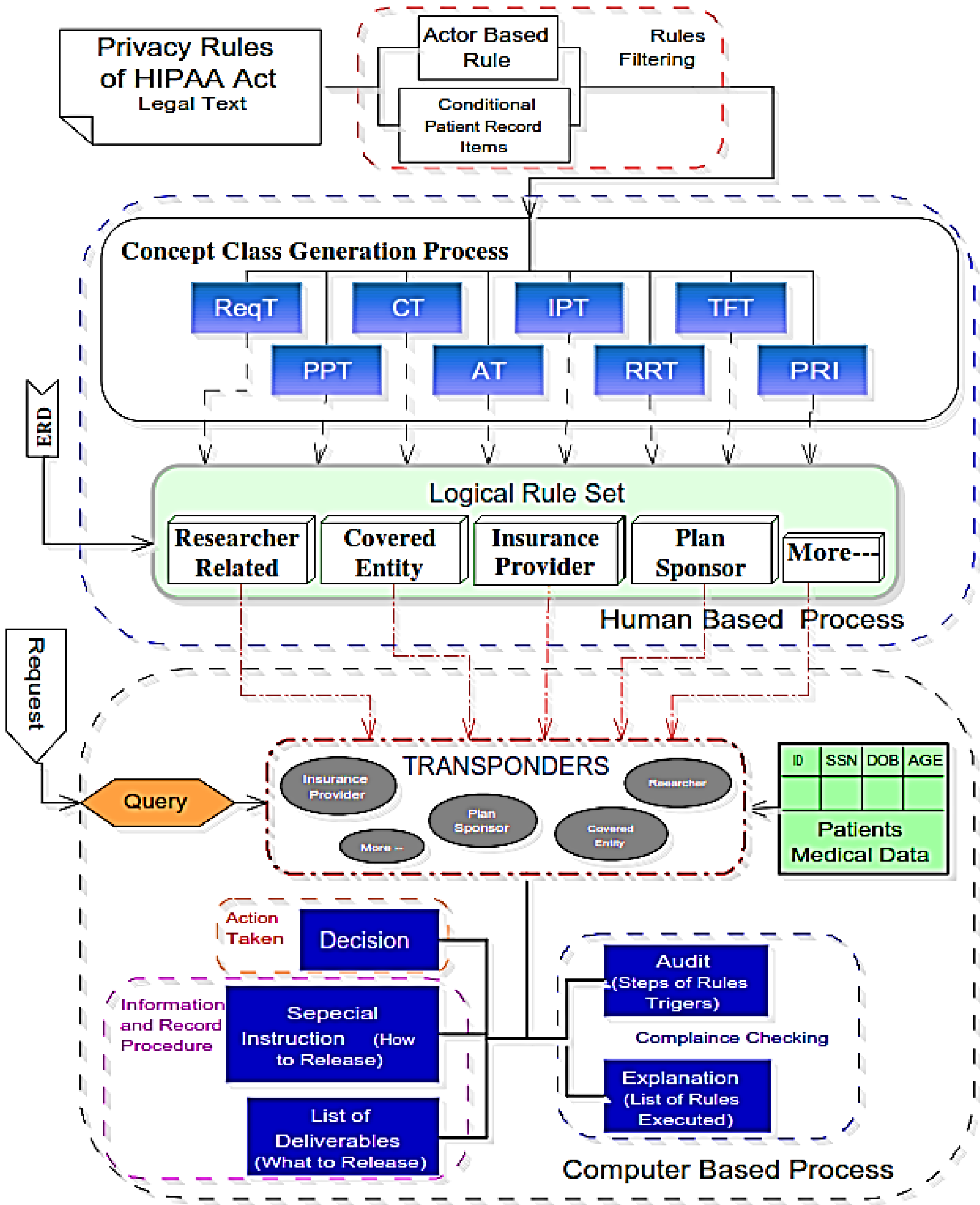

Fig. 5: Filtering rules based on requester through HIPAA world rule model

HIPAA has different number of Sections like (164.524), and each Section has the Clause (164.524.a), Super Sub Clause (164.524.a.1), Sub Clause (164.524.a.1.i) and sub of Sub Clause may exist (164.524.a.1.i.A). For request we have to search in depth up to sub of sub clause step by step from the top to bottom level of each section. If we extract rules from HIPAA as clause by clause from different sections, we have to use the Fig. 6. The complexity of HIPAA can be analyzed by this algorithm before formalization.

This algorithm is only for the searching of rule from HIPAA Sections, and after that each rule has to be compared with the purpose of request, conditions for that purpose, conditions for the requester and the patients' record items if that are required in request.

In (Khan et al., 2013) our first step to formalize the legal text and make the separate classes for these rules and all the rules are under any one of the class. So with this approach we reduce the searching of rule complexity from each section, as we have assigned our own tag value to each clause and under that sub clause on single level. So now for each request we have to search only from the eight classes for the request one by one. We have reduced the search spaces but complexity is still there for the comparison of related purpose, conditions and PRI. In section 4 we discussed how we reduce this complexity of comparison and get more reliable results (Fig. 6).

```
Data: This algorithm extract rules form different sections of HIPAA
Result: Generates Rules as clause by clause
Section S = 1 → N;
Super subClause SSubC = 1 → P;
Sub Clause SubC = 1 → Q;
Sub of SubClause SofSub = 1 → R;
while S ≤ N do
    i=1; j=1; k=1; l=1;;
    while i ≤ M do
        Search_Rule and save the Rule;
        if any C has SSubC then
            while j ≤ P do
                Search_Rule and save the Rule ;
                if any SSubC has SubC then
                    while k ≤ Q do
                        Search_Rule and save the Rule ;
                        if any SubC has SofSub then
                            while l ≤ R do
                                Search_Rule and save the Rule ;
                            end
                        end
                    end
                end
            end
        end
    end
end
```

Fig. 6: HIPAA rule extraction algorithm from Legal Text

3.1. Time complexity

In this subsection we present a brief analysis of HIPPA rule extraction algorithm. Let T(n) be the total time taken by the algorithm to extract a rule. We also assume each statement in the pseudo code takes a constant time to execute therefore single statements outside loop's bodies are not considered. Only the loops are taken into consideration for this analysis. So the total time can be expressed as in Eq. 1.

$$T(n) = \sum_{S=1}^{N} N_s\left(\sum_{i=1}^{M} M_i\left(\sum_{j=1}^{P} P_j\left(\sum_{k=1}^{Q} Q_k\left(\sum_{l=1}^{R} R_l\right)\right)\right)\right) \quad (1)$$

Worst Case: In worst case all if statements get executed i.e. rules are completely extracted in sub of sub Clause. Since each term takes constant time in each loop therefore (Eq. 2),

$$\sum_{S=1}^{N} N_s = n, \qquad \sum_{i=1}^{M} M_i = n, \sum_{j=1}^{P} P_j = n, \sum_{k=1}^{Q} Q_k = n, (\sum_{l=1}^{R} R_l = n \quad (2)$$

Then the total asymptotic time of this algorithm in worst case is $T(n) = O(n^5)$

Best Case: Best case occurs when the desired rule completes in super sub Clause i.e. the first if statement in second loop is false every time. In this case the time complexity $T(n) = O(n^2)$.

4. Results and discussion

4.1. HIPAA rules filtration through world rules model

We basically present two problems. The first is related to the formalization process of the entire privacy rules of HIPAA which we have resolved in (Khan et al., 2013), where the second is the number of comparisons when querying data.

To solve the first problem, we propose partial formalization of legal text. Instead of formalizing the entire text of privacy rules, only rules that are related to certain discipline will be formalized (Fig. 7).

Rules from different sections of HIPAA are extracted using algorithm in Fig. 7, which is implemented in python. All the clauses in different sections of HIPAA are separated in XML "rule" tags. The information included in the "rule" tags about clauses for the decision is also added as XML tags. Each clause contains some keywords for logical rules. These keywords tags may about a requester, the purpose of information access, condition, action and cross referenced rules etc. These keyword tags may vary in the "rule" tag of each clause depending upon the information in the specific rule or clause. These logical rules are categorized into different medical entities for the fast retrieval of information. XML logical rules are separately stored into the transponder of WRM modal. A Sample of rule generated in XML from HIPAA is shown below.

```
<rules>
 <rule ruleid="164.502.1">
  <Request>access</Request>
  <Requester>patient</Requester>
  <Entity>hospital</Entity>
  <AccessLevel>permission</AccessLevel>
  <Condition>hospital_Law_2</Condition>

<CrossReference>164.534.a.1</CrossReference>

<CrossReference>164.524.a.2</CrossReference>
 </rule>
 </rules>
```

The conditions related to different actors are marked separately as references to check the "rule" for that condition. For example the conditions tag in the above "rule" tag is hospital_Law_2, which is referring to some other rule for condition test.

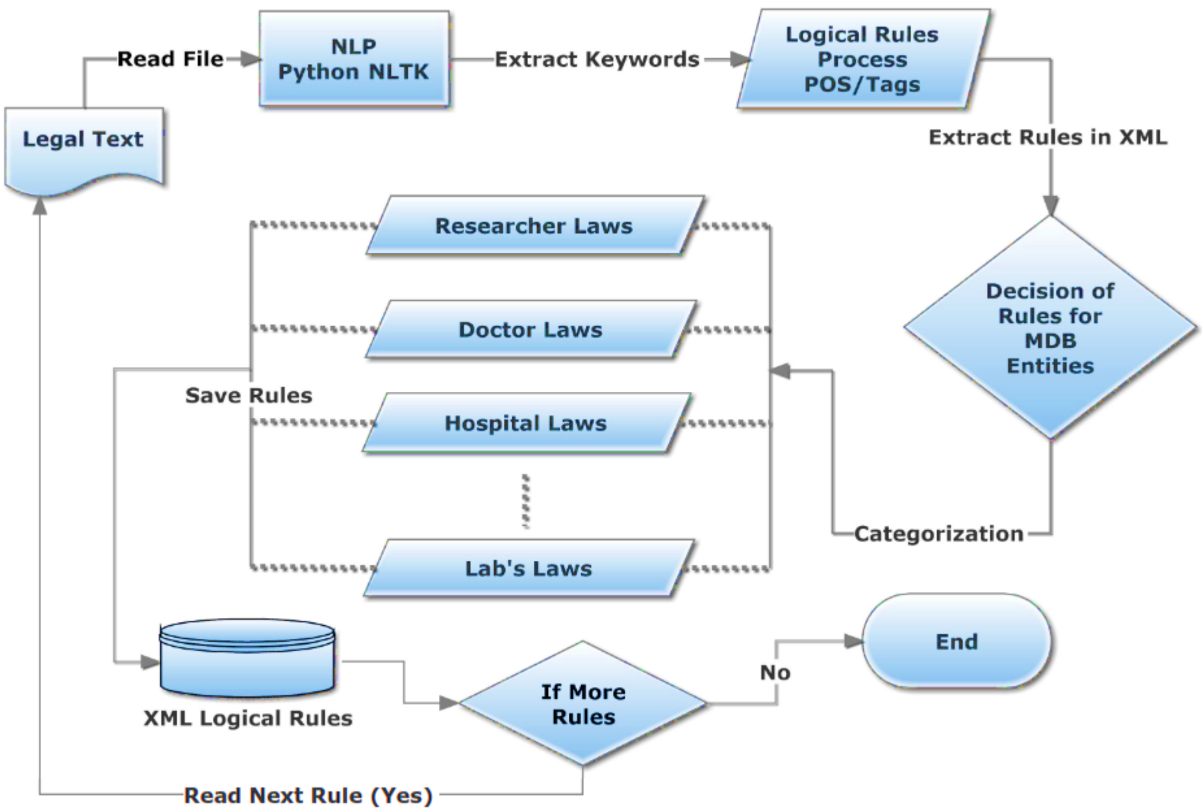

Fig. 7: XML rules filtering algorithm model

For example if the users want to access PHI as a researcher, then clauses related to researcher will be extracted from HIPAA. Nevertheless, the researcher as a requester of data has the right to access any patient record items based on the authority and the purpose of this access. It means all clauses that contain patient record items need to be formalized in this process. Having partial formalization would save time and money because we formalized portion of the entire legal text to build one block. Once more resources afforded, another block can be built, but for now only clauses related to a certain requester will be formalized and no more formalization to the patient record items is done, see rules filtering in Fig. 5. Remaining steps of formalization will be similar to the method discussed in (Khan et al., 2013). Example of researcher rule is shown in Fig. 8.

```
<rule ruleid="164.508.a.4">
  <Request>access</Request>
<Requester>researcher</Requester>
  <Entity>lab</Entity>
  <AccessLevel>permission</AccessLevel>
  <Condition>lab_11</Condition>

<CrossReference>164.512.b.3</CrossReference>
 </rule>
```

As for as the second problem is concerned, we have developed a new methodology to search rules set based on requester instead of making comparisons with all rules' sets. Availability of more requesters will not change the fact that only rules related to requester will be examined.

Data: Algorithm checks rules when a requester query for data through HIPAA World Rule Model
Result: Generate output rules according to Actors request
$arrule = 1 \rightarrow N$;
while $arrule \leq 1toN$ **do**
  if *by_rule = true* **then**
    if *cond_rule OR Ref_rule exist* **then**
      if *cond_rule OR Ref_rule = true* **then**
        Save the Rule goto step-2 ;
      else
        Exit;
      end
    end
  else
    Exit;
  end
  return result;
end

Fig. 8: Actor base rule filtering algorithm from Legal Text through HIPAA world rule model

4.2. Time complexity

Time complexity of the proposed Fig. 8 for extracting rules is O(n), since there is only one loop running n times. Again here we do not consider any statements outside loop's body since their time is constant. Also note that the worst and best case time complexity is same for this algorithm. Therefore we can express the asymptotic running time of the algorithm as T(n) = O(n)

4.3. Complexity comparison of actor base approach

Due to the complexity of legal text of HIPAA, the proper formalization of HIPAA has not been done but several interesting attempts have been made to convert general legal text into logic rules (Maxwell and Antón, 2009; Lam et al., 2009). As the complexity of HIPAA rules extraction clause by clause is very high i.e $T(n) = O(n^5)$ and that will grow exponentially if more sub of Sub Clause are in HIPAA. But in Actor based approach running time of the algorithm is T(n) = O(n), which is in a linear time to evaluate the request. Using this approach we will be able to get results more efficiently and accurately.

4.4. Actor based queries example

Query: "A researcher wants to access/retrieve disclose PHI regarding HIV patients who were admitted during in current year to improve healthcare. He/she does not have authority to access PHI but has been granted a waiver from the medical provider to fulfill this job". Flow of rules generated according to HIPAA for researcher is shown in Table 2 and the result will be generated according to these rules.

Symbols used in Fig. 7 are explained as follow; '∧' means "and within a rule", 'and' means "and between rules",'||' is used as "or", '!' for "not", '~' means "may". The resolution process reads rules as, If (Requester = "Researcher" and Purpose="Purpose Tags", and Items = "Patient Record Items"; check Conditions (Condition = "Conditions Tags" and ReqT Condition = "Requester Tags" then (Action= "Action Tags", and Record Release= "Record Release Tags" and Information Procedure= "Information Procedure Tags", and Time and Fee = "Time and Fee Tags")));

Table 2 shows that how logical rules are extracted and represent constraint with instructions that are expressed in HIPAA. For example, the first logical rule in Table 1 means, that if a requester is of type researcher (ReqT1) with purpose (PPT1), that is for research and is requesting for the patient record item (PRI). If he/she satisfies all the conditions " CT 1 ∧ CT 2 ∧ CT 3 ∧ CT 4, that might satisfy CT5 (CT 6 ∧ CT 7), might satisfy CT8 (CT9 ∧ CT10 ∧ CT11) and must have to satisfy CT12 (CT 13 ∧ CT 14 ∧ CT 15)" then an Action (AT4) for disclosing the information is issued as a result after satisfying all conditions. This decision will be related to deliverable that indicates that how the information will be released "~ IP T 1 ∧ IP T 2 ∧ IP T 3" and filters will be applied on released information "RRT1 RRT2". Tags details are explained in (Khan et al., 2013).

Table 2: Rules generated for researcher in RBDT 5

<table>
<tr><th>Requestor</th><th>Purpose</th><th>PRI</th><th>Conditions</th><th>Action</th><th>Information Release Procedure</th></tr>
<tr><td>Researcher</td><td>PPT1</td><td>Any PRI</td><td>" CT1 ^ CT2 ^ CT3 ^ CT4 " & "~CT5(CT6 ^ CT7) ^ ~CT8(CT9 ^ CT10 ^ CT11) ^ CT12</td><td>AT4</td><td>" RRT1 || RRT2 " & "~IPT1 ^ IPT2 ^ IPT3</td></tr>
<tr><td>Researcher</td><td>PPT1</td><td>Any PRI</td><td>" !(CT1 ^ CT2 ^ CT3 ^ CT4) " & " CT5(CT6 ^ CT7) ^~ CT8(CT9 ^ CT10 ^ CT11) ^ CT12</td><td>AT4</td><td>" RRT1 || RRT2 " & "~IPT1 ^ IPT2 ^ IPT3</td></tr>
<tr><td>Researcher</td><td>PPT1</td><td>Any PRI</td><td>"! CT1 ^ CT2 ^ CT3 ^ CT4" & " !CT5(CT6 ^ CT7) ^~ CT8(CT9 ^ CT10 ^ CT11) ^ CT12</td><td>AT1</td><td></td></tr>
<tr><td>Researcher</td><td>PPT1</td><td>Any PRI</td><td>"!CT1 ^ CT2 ^ CT3 ^ CT4 " & " CT5(!CT6 ^ CT7) ^~ CT8(CT9 ^ CT10 ^ CT11) ^ CT12</td><td>AT1</td><td></td></tr>
<tr><td>Researcher</td><td>PPT1</td><td>Any PRI</td><td>"! CT1 ^ CT2 ^ CT3 ^ CT4" & " CT5(CT6 ^ CT7) ^~ CT8(CT9 ^ CT10 ^ CT11) ^ !CT12</td><td>AT1</td><td></td></tr>
<tr><td>Researcher</td><td>PPT1</td><td>Any PRI</td><td></td><td>AT1</td><td></td></tr>
<tr><td>Researcher</td><td>!PPT1</td><td>Any PRI</td><td></td><td>AT1</td><td></td></tr>
<tr><td>Researcher</td><td>!PPT1</td><td>Any PRI</td><td></td><td>AT1</td><td></td></tr>
</table>

5. Conclusion

Formalizing legal text into logical rules is a complex and challenging task which consumes time and effort. This paper presents an improvement on our previously proposed WRM modal for decision, based on the extracted rules from legal text by reducing its complexity of comparisons. The rules related to specific actors are extracted separately using the Rule Base Decision Tree (RBDT) and Rule Extraction Algorithm from legal text. Using Actors based approach, reduces the time complexity of the decision process in rule generation process (algorithm) from *O* (n^5) to *O(n).* The proposed methodology can be used to formalize other legal texts and potentially open up a new area of computer assisted jurisprudence.

References

Ashley P, Hada S, Karjoth G, and Schunter M (2002a). E-P3P privacy policies and privacy authorization. In Proceedings of the 2002 ACM Workshop on Privacy in the Electronic Society. ACM: 103-109, Washington, USA.

Ashley P, Powers C, and Schunter M (2002b). From privacy promises to privacy management: A new approach for enforcing privacy throughout an enterprise. In Proceedings of the 2002 Workshop on New Security Paradigms. ACM: 43-50, Virginia, USA.

Breaux T and Antón A (2008). Analyzing regulatory rules for privacy and security requirements. IEEE Transactions on Software Engineering, 34(1): 5-20.

Breaux TD, Vail MW, and Anton AI (2006). Towards regulatory compliance: Extracting rights and obligations to align requirements with regulations. In 14th IEEE International Requirements Engineering Conference (RE'06): 49-58, Minneapolis, USA.

Byun JW, Bertino E, and Li N (2005). Purpose based access control of complex data for privacy protection. In Proceedings of the Tenth ACM Symposium on Access Control Models and Technologies. ACM: 102–110, Stockholm, Sweden.

Hodge JG (2003). Health information privacy and public health. The Journal of Law, Medicine and Ethics, 31(4): 663-671.

Khan I, Alwarsh M, and Khan JI (2012a). Quantitative charting of HIPAA section 164's legal universe privacy rules of HIPAA (Technical Report). Networking and Media Communications Lab, Kent State University. Available online at: http://www.medianet.kent.edu/techreports/TR2012-06-01-HIPAA-164-legaluniverse.pdf

Khan I, Alwarsh M, and Khan JI (2012b). Tr2012-16-02 a decision tree model of HIPAA section 164 of HIPAA legal requirements, Tech. rep., Technical Report, Networking and Media Communications Lab, Kent State University. Available online at: http://www.medianet.kent.edu/techreports/TR2012-06-02-HIPAA-164-decisiontree.pdf

Khan I, Alwarsh M, and Khan JI (2013). A comprehension approach for formalizing privacy rules of HIPAA for decision support. In Machine Learning and Applications (ICMLA), 12th International Conference, 2: 390-395, Miami, USA.

Lam PE, Mitchell JC, and Sundaram S (2009). A formalization of HIPAA for a medical messaging system. In International Conference on Trust, Privacy and Security in Digital Business. Springer Berlin Heidelberg: 73-85.

Massey AK, Otto PN, Hayward LJ, and Antón AI (2010). Evaluating existing security and privacy requirements for legal compliance. Requirements Engineering, 15(1): 119-137.

Maxwell JC and Antón AI (2009). Developing production rule models to aid in acquiring requirements from legal texts. In 7th IEEE International Requirements Engineering Conference: 101–110, Atlanta, USA.

Maxwell JC and Antón AI (2010). The production rule framework: developing a canonical set of software requirements for compliance with law. In proceedings of the 1st ACM International Health Informatics Symposium. ACM: 629-636.

Sandhu RS, Coynek EJ, Feinsteink HL, and Youmank CE (1996). Role-based access control models yz. IEEE Computer, 29(2): 38-47.

Wilson JF (2006). Health insurance portability and accountability act privacy rule causes ongoing concerns among clinicians and researchers. Annals of Internal Medicine, 145(4): 313-316.