DOI 10.4010/2016.716
ISSN 2321 3361 © 2016 IJESC

Research Article Volume 6 Issue No. 3

Scalability and Privacy Preservation for Health System Using Data Anonymization

Siddhesh Prabhu[1], Shreyas Kr Singh[2], Dewashish Mandal[3], Manish Gangawane[4], Manisha Wazalwar[5]
UG Scholar[1, 2, 3], Professor[4, 5]
PVPPCOE, Dept. of COMP, Mumbai, India.
Sid174455@gmail.com[1], Shreyassingh.98@gmail.com[2], Devmandal6@gmail.com[3]

Abstract:
Data anonymization focuses on privacy protection. In this paper, we are discussing a health system which will hide the identities of the patients from other users and also it will give prediction for best doctors in the nearby locality based on ratings provided by the user. First we will take patients information and store it in a database. Then we will use AES encryption algorithm for privacy protection. When a new patient registers in our system and search for a doctor, our system will recommend best doctors in the nearby locality for the patient based on previous users rating. For this recommendation purpose we are using Apriori algorithm**.**

Keywords: Data anonymization, Privacy Protection, Scalability, AES, Apriori, Recommendation.

I.INTRODUCTION

A personal health record, or PHR, is a health record where health data and information related to a patient is maintained by the user. The intention of a PHR is to provide a complete and accurate summary of an individual's medical history which is accessible online. The main aim of the proposed system is to provide secure patient-centric PHR access [1]. We are using AES encryption algorithm to encrypt each patient's PHR file. Using this system the patient gets best doctors, hospital related information and can save personal health record data on a server. We are using Apriori algorithm for recommending best doctors. The doctor will also be able to provide prescription to the patient.

II. EXISTING SYSTEM

The existing system requires high cost for building and maintaining the database. In existing system many PHR services are provided by third-party service providers. So the main problem in existing system is whether the patients could actually control the sharing of their personal health information, especially when they are stored on a third-party server which people may not fully trust. Also in existing system features like recommendation of best doctors is not available.

III. PROPOSED SYSTEM

The drawback of existing system is overcome by our proposed system. In our proposed system we will encrypt each patient's PHR file using AES encryption algorithm. So the unauthorized users cannot easily breach the information or alter any information. The patients can control sharing of their personal health information. Also the system will provide recommendation of best doctor using Apriori algorithm. Using this system the doctor can easily track patient's health record according to which the doctor can also provide prescription to the patient.

Data anonymization process proceeds in following 2 steps:

1. Encryption using AES algorithm
2. Recommendation of best doctor using Apriori algorithm.

Encryption using AES:

Advanced Encryption Standard (AES) is the first step of the process. AES algorithm is used to provide encryption for the login credentials and other sensitive information of the user [2].

The AES design consists of the secret key, plain text, cipher block and cipher text. A key length of 128 bit is applied on the plain text block and after the processing of cipher block, the resulting 128 bit cipher text is obtained.

AES consists of 10 rounds for 128 bit key. There are 9 regular rounds out of the 10 rounds while the 10^{th} round is different. The round of cipher block processing includes Shift Rows, Sub Bytes, Mix Columns and Add Round Key. The inverse cipher rounds include Inverse Shift Rows, Inverse Sub Bytes, Inverse Mix Columns and Add Round Key. Thus we obtain the cipher text which will be in encrypted form.

Recommendation using Apriori:

It is an algorithm for frequent item set mining and association rule learning over transactional databases. It was proposed by Agrawal and Srikant in19994 [3]. It uses a "bottom up" approach, where frequent subsets are extended one at a time (a step known as candidate generation), and groups of candidates are tested against the data. It uses breadth-first search and a Hash tree structure to count candidate item sets efficiently. The algorithm terminates when no further successful extensions are found.
It is a level wise algorithm which works in an iterative fashion to discover all frequent itemsets in a database. It uses prior

knowledge of frequent itemset properties [4]. Frequent itemsets are the sets of items that satisfy minimum support threshold. This algorithm takes only categorical input and associates attributes present in the dataset. There is a property associated with this algorithm called "Apriori Property" which states that any subset of frequent itemsets is also a frequent itemset. For example, if {x,y,z} is a frequent set then the sets { {x},{y},{z} }, { { x,y },{ x,z },{ y,z } } must also be frequent. The execution of this algorithm is done in two phases. In the first stage, the candidates are generated and in the next phase frequent itemsets are generated [5]. The generated large itemsets are used to produce association rules from database.

IV. SYSTEM DESCRIPTION

In this system first the patient needs to fill all the necessary personal information. This personal information includes all the necessary individual data regarding that patient like his name, address, mobile no, etc. This personal information will be stored in a database. To maintain database privacy and to provide security over the database Advanced Encryption Standard (AES) is used. It will encrypt each patient's PHR file. At the time of registration there are two different domains used for registration such as personal domain and public domain. Personal domain contains the patient's registration and public domain contains doctor's registration. It will provide scalable access over whole database [6].

After registration the patient will enter his health problems. Then based on previous users rating using Apriori algorithm our system will recommend best doctor for the patient. For apriori algorithm implementation first we will take previous users doctors rating. Then we will set minimal value of rating for each doctor. From this we will then find which doctor is mostly referred by the user. Then we will recommend that doctor to the patient. So using this system the patient can get best doctor in their locality.

After allocating doctor to the patient, the patient will send his medical history along with his current health problem to the doctor. Our system will make sure that only the patient and the allotted doctor can access patient's personal information. So the unauthorized user cannot breach the information from the database. This was possible in previous system due to third-party semi-trusted servers. The doctor after receiving patient's PHR file, he will provide some prescription to the patient. The patient in his user account can see this prescription and can be treated accordingly.

Advantages:

1. Low cost
2. Data confidentiality
3. Scalability and high data utility
4. Recommendation system
5. Easy to use

The working of the system is explained below:

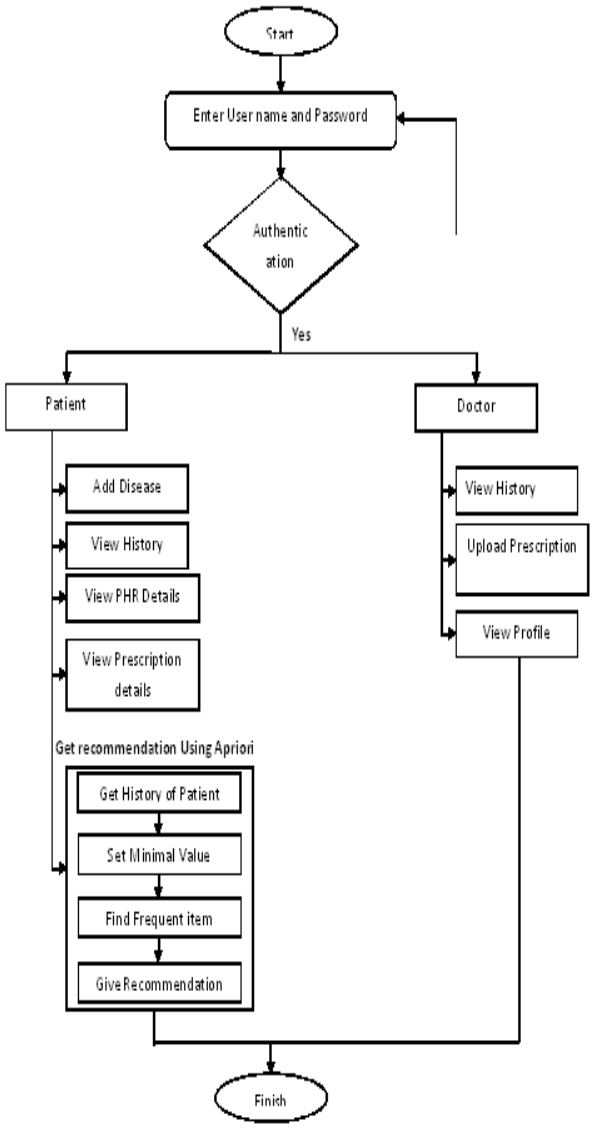
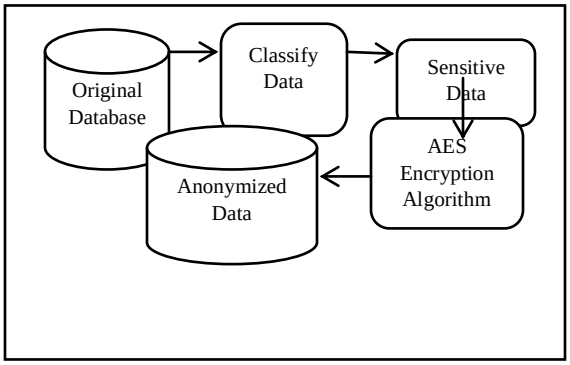


Fig.1.Working of Health System

System Architecture for privacy preserving:



Future modifications which can be done our system are:

1. Integrating mobile commerce. So that the users can use this system from handheld devices such as mobile. We can develop an app for our system.

2. In future, to provide high security and privacy for Personal Health Record (PHR), the existing Multi authority attribute based encryption could be further enhanced to proactive Multi authority attribute based encryption.

VI. CONCLUSION

In this paper, we have proposed secure sharing of personal health records between patient and doctor. Patients will have complete control over sharing of their personal health information. It greatly reduces the complexity of key management while enhancing the privacy compared with previous work. Our proposed system provides the advantage of data confidentiality, integrity and data utility over existing system. Our system also provides recommendation system which was not present in previous system.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude for the assistance and support of Mr. Manish Gangawane (Professor), Mrs. Pallavi Wazalwar (Assistant Professor) who made this project initiative a success. We would like to acknowledge throughout support and guidance of them in every step from conceptualization to implementation of system adding in successful implementation of project.

REFERENCES

[1] M. Li, S. Yu, K. Ren, AND W. Lou, "Securing personal health records in cloud computing: Patient-centric and fine-grained data access control in multi-owner settings," in SecureComm'10, Sept. 2010,pp. 89-106.

[2] J. Benaloh, M. Chase, E. Horvitz, and K. Lauter, "Patient controlled encryption: ensuring privacy of electronic medical records," in CCSW '09, 2009, pp. 103-114.

[3] Smitha T. and V. Sundaram, Association Models for Prediction with Apriori Concept, International Journal of Advances in Engineering & Technology, Nov. 2012, Vol. 5, Issue 1, pp. 354-360.

[4] P. Kasemthaweesaband, W. Kurutach, "Association Analysis with Complication States Based on Association Rules", 7th IEEE Conference on Industrial Electronics and Applications (ICIEA) 2012.

[5] M. Ilayaraja and T. Meyyappan, "Mining Medical Data to Identify Frequent Diseases using Apriori Algorithm", In Proceedings of the 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME), 21-22 February.

[6] S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving secure, scalable, and fine-grained data access control in cloud computing," in IEEE INFOCOM'10, 2010.