IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 09, Issue 5 (May. 2019), ||S (XI) || PP 61-66

Trust-Level Security Based Authentication In Health Data Integration

Varsha Katiwal[1], Ms. Nisha Balani[2], Ms. Priyanka Dudhe[3]

M.Tech, Department of Computer Science & Engineering,
Jhulelal Institute of Technology, Nagpur, India
E-mail: varshakatiwal3@gmail.com [1] , n.balani@jit.org.in[2] , p.dudhe@jit.org.in [3]

Abstract : *— We are requires various health data domains the incorporation of healthcare data from diversified sources. Maintaining record linkage during the integration of medical data is an important research issue. we have given different solutions to this problem that are applicable for developed countries where electronic health record of patients are maintained with identifiers like social security number (SSN), universal patient identifier (UPI), health insurance number, etc. These process cannot be used correctly for record linkage of health data of developing countries because of missing data, ambiguity in patient identification, and high amount of noise in patient data information. This concept motivates us to create a trust level security authentication. It means, this healthcare database will be fully secured using cryptography algorithm of encryption and decryption using AES algorithm and authentication will be controlled on "Trust Level Security". We have proposed privacy preserved secured record linkage architecture that can support health data of developing countries such as Bangladesh. Our technique can unidentified identifiable private data of the patients while maintaining record linkage in integrated health repositories to facilitate knowledge discovery process.*

Keywords: *Data Security; Health Data Warehouse; Privacy Preserved Record Linkage; Data Mining*.

I. INTRODUCTION

Data required making proper medical decisions are trapped within fragmented and heterogeneous health systems that are not properly integrated. So, the integration of these health records into a single warehouse is necessary [1][2]. Healthcare data hubs are highly beneficial in many fields such as tracking health patterns, evidence-based medicine, personalized treatments, etc. Clinical diagnostic equipment creates a large amount of health records and related documents every day. These worthy healthcare data are reserved in different healthcare information systems such as Picture Archiving and Communications System, Hospital Information System, Radiology Information System, etc. in public hospitals, private clinics, and diagnostic centers. For maximum benefit from integrated health data repositories (IHDR), linkage of records is essential. Discovering effective knowledge (e.g., correlations among diseases) from medical dataset requires maintaining record linkage. Record linkage is the process of identifying record pairs from different information systems which belong to the same real-world entity. Given two repositories of records, the record-linkage process consists of determining all pairs that are similar to each other. The similarity between two records is defined based on domain-specific similarities over individual attributes constituting the record. [3][4][5]. Protecting the privacy of patients while maintaining effective record linkage, that is Privacy Preserved Record Linkage (PPRL), is currently an important focus of the researchers [3][6][7]. Health data containing protected health information (PHI) such as name, date of birth (DOB), and address can be made linkable easily with the help of PHI. But retaining PHI in healthcare data is very risky. These data are highly lucrative to hackers. Sell value of medical records containing PHI are 100 times more than credit card numbers and Social Security Numbers (SSN).

II. Related Work

In order to complete this research successfully, we have gone through following research papers to get ideas:

we studied that, Data warehousing methodologies share a common set of tasks, including business requirements analysis, data design, architectural design, implementation and deployment. Clinical data warehouses are complex and time consuming to review a series of patient records however it is one of the efficient data repository existing to deliver quality patient care. Data integration tasks of medical data store are challenging scenarios when designing clinical data warehouse architecture. The presented data warehouse architectures are practicable solutions to tackle data integration issues and could be adopted by small to large clinical data warehouse applications. [1]

we studied that, Clinical data warehouses offer tremendous benefits as a foundation for data mining. By serving as a source for comprehensive clinical and demographic information on large patient populations, they streamline knowledge discovery efforts by providing standard and efficient mechanisms to replace time-consuming and expensive original data collection, organization, and processing. Building effective data warehouses requires knowledge of and attention to key issues in database design, data acquisition and processing, and data access and security. In this article, the authors provide an operational and technical definition of data warehouses, present examples of data mining projects enabled by existing data warehouses, and describe key issues and challenges related to warehouse development and implementation. [9]

Record linkage to integrate uncoordinated databases is critical in biomedical research using *Big Data*. Balancing privacy protection against the need for high quality record linkage requires a human–machine hybrid system to safely manage uncertainty in the ever changing streams of chaotic *Big Data*. [3]

Healthcare organizations in Bangladesh own a large amount of data in diverse health information systems. Potential and useful hidden knowledge can be discovered if integration of this huge medical data is performed in national level. The integration process requires linkage of patients' records among different heterogeneous sources. To facilitate effective data mining, it is essential to preserve record linkage in health data warehouse by retaining identifiable attributes. On the other hand, identifiable health data have high risk to patient privacy and also increase the chance of attacks by cyber criminals. In this paper, we have provided a practical solution of privacy and security problems for developing national health data warehouse of Bangladesh. Our developed technique can anonymize identifiable private data of the patients while maintaining record linkage in national warehouse to facilitate knowledge discovery process. For this purpose, we have used encrypted mobile number, gender and name-value of patients to produce Patient Identification Key. Our system is being implemented to protect privacy of sensitive health data in health data warehouse [8].

III. Current Implementation

After studying the literature we proposed a system which works on the following modules Figure 1: Architecture of project flow diagram

a. Designing

I. In this module we will design the GUI for insert/update/delete of patient record, so as to develop Health Information System. This will be Frontend Web User-Friendly GUI. We will develop individual GUI for Admin, Users, and Trust Level Authentication. figure no. 1 show the hospital registration figure no. 2 show the hospital main portal , figure no. 3 show patient profile .

Figure 1: Hospital Registration

Figure 2: Hospital Main Portal

Figure 3: Patient Profile

b. Health Information System

I. This system will contain data of patients belonging to different regions and religions of underdeveloped country like Bangladesh
II. Data is differentiated according to attributes like name, mobile no, age, sex
II. We will use MS SQL Server to maintain database of Health Information System. Figure no .4 show the SQl database

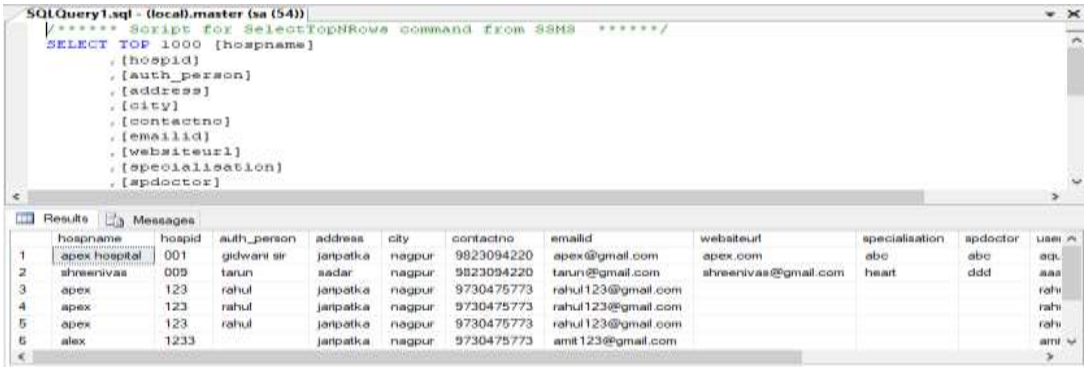
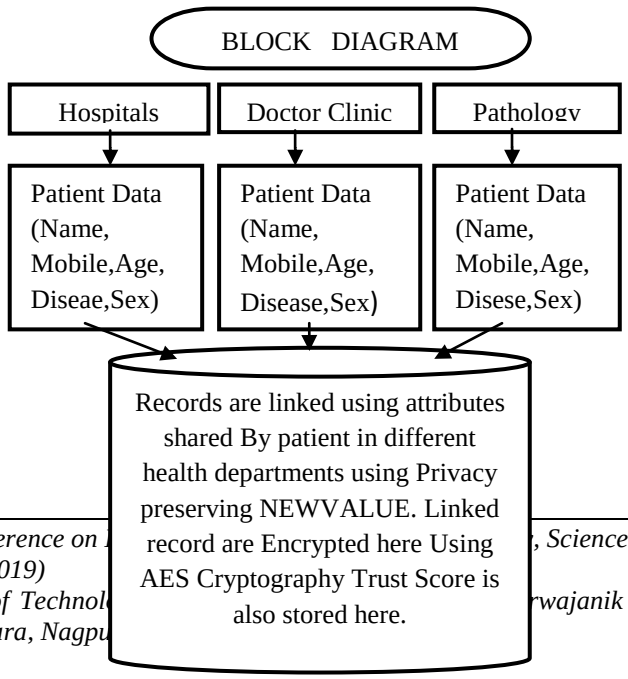


Figure 4: SQl database

c. Cryptography

I. Data will be stored in encrypted format to improve data security. We will use AES Algorithm for encryption and decryption
II. We will maintain key and share it according to trust level. Figure no .5 show the encrypted data.

Figure 5: Encrypted data



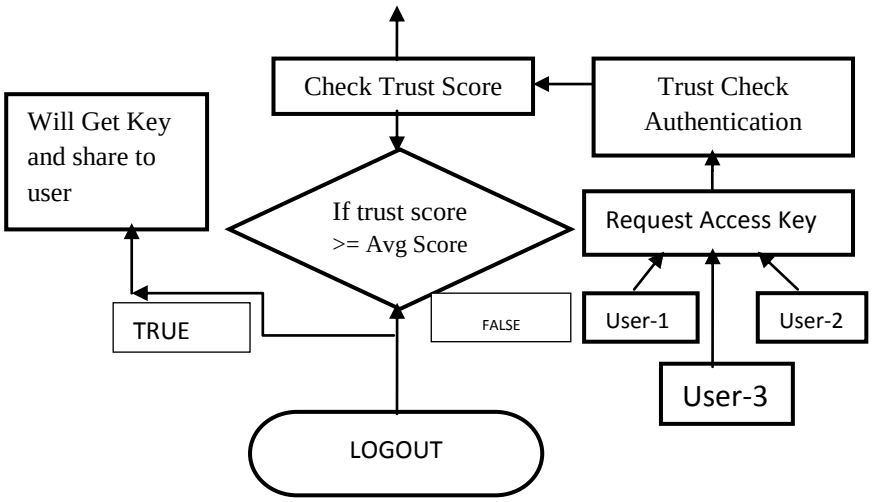


Figure 7: Architecture of project flow diagram

IV. CONCLUSION

Preserving record linkage by retaining identifiable attributes in national health data warehouse plays a vital role for effective data extraction. Once Health data warehouse is developed with record linkage, we provide Trust Level Authentication process to access patient records. It means that if any researcher or organization needs to access this data, then he/she must have at least above average trust level. Higher the trust level higher will be scope to access this data. Health data warehouse development is a complex and time-consuming process but is essential to deliver quality health services.

References

[1]T.R. Sahama, and P.R Croll, "A Data Warehouse Architecture for Clinical Data Warehousing," Australasian Workshop on Health Knowledge Management and Discovery, 2007.

[2] J. H. Weber-Jahnke & C. Obry "Protecting privacy during peer-to-peer exchange of medical documents", Inf Syst Front (2012), Springer, 14:87–104

[3] H.C. Kum, A. Krishnamurthy, A. Machanavajjhala et. at., "Privacy preserving interactive record linkage (PPIRL), " J Am Med Inform Assoc vol. 21, 2014, pp. 212–220.

[4] J. Liang, L. Chen, and S. Mehrotra, "Efficient Record Linkage in Large Data Sets," In Proc. of the Eighth International Conference on Database Systems for Advanced Applications, 2003.

[5] Your medical record is worth more to hackers than your credit card.

[6] P. Christen, "Automatic record linkage using seeded nearest neighbour and support vector machine classification," In Proc. of the 14th ACMSIGKDD International Conference on Knowledge Discovery and Data Mining, 2008.

[7] S. I. Khan and A.S.M.L. Hoque, "Privacy and security problems of national health data warehouse: a convenient solution for developing countries," In Proc. of the International Conference on Networking Systems and Security (NSysS). IEEE,2016.

[8]C. F. Andrea, C. Danielle, T.M. Matthew et. al., "Development and evaluation of a de-identification procedure for a case register sourced from mental health electronic records," BMC Medical Informatics and Decision Making, Vol. 13, 2013, pp.13:71.

[9] J.A. Lyman, K. Scully, and J.H. Harrison, "The development of health care data warehouses to support data mining," Clin Lab Med. 28,1 2008, pp. 55-71

[10] P. Christen, "Automatic record linkage using seeded nearest neighbor and support vector machine classification," In Proc. of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2008.

[11] E.A. Sauleau, J. Paumier, and A. Buemi, "Medical record linkage in health information systems by approximate string matching and clustering," BMC Med Inform Decision Making, Vol. 5, 2005, pp.32– 44.

[12] A.B. McCoy, A. Wright, Kahn M. et al., "Matching identifiers in electronic health records: implications for duplicate records and patient safety," BMJ Qual Saf Vol. 22, 2013, pp.219–24.

[13] C. F. Andrea, C. Danielle, T.M. Matthew et. al., "Development and evaluation of a de-identification procedure for a case register sourced from mental health electronic records," BMC Medical Informatics and Decision Making, Vol. 13, 2013, pp.13:71.

[14] N. K. Abel, P. C. John, L. J. Kathryn et al., "Design and implementation of a privacy preserving electronic health record linkage tool in Chicago," Journal of the American Medical Informatics Association, 2015, pp.1-9.