HUGE RECORDS SOLITUDE BY ASPECTS OF BIG DATA IN BIOMEDICAL RESEARCH

P. Sarathkumar[1], J. Sreerambabu[2], N.Santhosh[3]

[1]PG Scholar, Master of Computer Applications Department, Thanthai Periyar Government. Institute of Technology, Vellore-2.
[2] Head of the Department, Master of Computer Applications Department, Thanthai Periyar Government. Institute of Technology, Vellore-2.
[3] Assistant Professor, Master of Computer Applications Department, Thanthai Periyar Government. Institute of Technology, Vellore-2.

Abstract

Biomedical research involves studying of patient records and records are used for research purpose for inappropriate use of patient records it may have chance to leakage of sensitive information's like SSN (Social Security Number) .It means patients information's are maintained in that SSN number only , we have to protect SSN number and maintain patient record, we want patient records for research purpose and also for research purpose it's very important so, without using SSN number we have to retrieve patient details in Electronic Data Record(EDR).

To maintain patient record in secure manner we introduce record linkage approach for research purpose, huge record means large number of data technically it refers to big -data(huge data). This process have to link that huge data to record linkage approach find possible- matches, non -possible match. If we require a patient detail for a research we require cancer patient information's We identify using name, age, address, DOB, blood pressure etc. if same required record matches is found forward that particular patient record to doctors for research, without using SSN we retrieve that record. Record linkage works by calculation comparing two datasets.

Keywords: *Data privacy, Record Linkage, SSN, EDR, Patient Records*

----------***----------

1. INTRODUCTION

In this Application we implement record linkage methods for maintain patient record in secure manner it focused on privacy protection techniques for electronic health record by using record linkage SSN number is maintained in this method And patient records information's are secured, and for research purpose easily we found some patient record information's without leakage of information's We discussed multiple aspects of big data privacy in the context of biomedical research. The "big" part of data privacy is because healthcare data often contain large scale clinical and genomic data, which are big in size and large in dimension. There are some unique challenges and off-the-shelf tools have difficulties in handling them By using record linkage it's possible and easy to maintain secure manner Record linkage works by calculation comparing two datasets. Finally using record linkage we maintain the patient record and we used to retrieve information's without using SSN number.

To maintain patient record in secure manner we introduce record linkage approach for research purpose, huge record means large number of data technically it refers to big -data(huge data). This process have to link that huge data to record linkage approach find possible- matches, non -possible match.

Purpose of the System

- Big data is large number of data difficult to handle them by using these we easily maintain that record and privacy is maintained in this model, privacy-preserving technologies for record linkage.
- In this yak out is used for securing some methods like privacy guarantee, scalability and linkage quality is maintained.
- The laws and government policies can be used to ensure obedience and provide guidelines in privacy protection. The Ethical, Legal, and Social Implications (ELSI) studies are essential for biomedical data privacy, which provides a new methodology to biomedical research by recognizing, investigating and tackling the ELSI of studies involving human subjects.

- All government laws are maintained and SSN number is not used to find patient details

2.SYSTEM ANALYSIS

Existing System

- In existing system all patient records and maintained in EHR but, there is no privacy security in patient records.
- They use the patient records for research purpose using SSN number and there may leakage of patient detail and there is no method for finding exact matching research matching patient matching requirements.
- Despite efforts in managing these information systems, the healthcare data often result to be fragmented, redundant, prone to errors, and heterogeneous, making the task of finding useful information among such data very challenging

Proposed System

- EHR linkage, the same patient records may change over time; therefore, the temporal information has to be taken into account for designing effective linkage solutions.
- In the database community recently few works have been proposed to linkage temporal records.
- While matching two datasets there is possible match occurs so that we easily used for research purpose for identifying new medicine without affecting patient privacy by using these method SSN model also secured and maintained Multiple aspects of big data privacy large scale clinical data.

3. DEVELOPMENT ENVIRONMENT

Hardware Requirement

RAM : 2GB

PROCESSOR : Intel(R)Pentium(R)2.10GHz

HARD DISK : 80GB Hard Disk space and above

MOUSE : PS/2 Mouse

KEYBOARD : PS/2 Keyboard

Software Requirement

PLATFORM : Visual Studio.NET 2015

FRONT END : VS.NET Framework2015.

BACK END : SQL Server 2014.

WEB BROWSERS : Firefox, Google Chrome.

DOCUMENT TOOL :MS-Office

4.MODULE DESCRIPTION

Admin Module

The Admin Module is a important module, In this module helps to enter the details about the patient records and its maintain and controlling of all the process. Then it is maintain the database.

Research Module

This Module is the research side wants to requesting the related patient record. And its enter the patient records, its sending to the admin then it is matching the database by using Record Linkage Approach

Enter Patient Records

In this Module helps to Admin and Research will enter the patient records depends up on related search done by the System .The Records are maintain in two Database A and B, All the patient records should be stored in secure manner.

Identify Matches

This module will compare both EDR patient record and probabilistic linkage approach to find possible, non-possible nearly possible matches would be identified. After identify matched nearly possible matches are manually identified by the Admin.

View Records

In this Module identify matches are displayed and possible matches, non possible matches, separated and matches are send to research purpose.

5. SYSTEM ARCHITECTURE

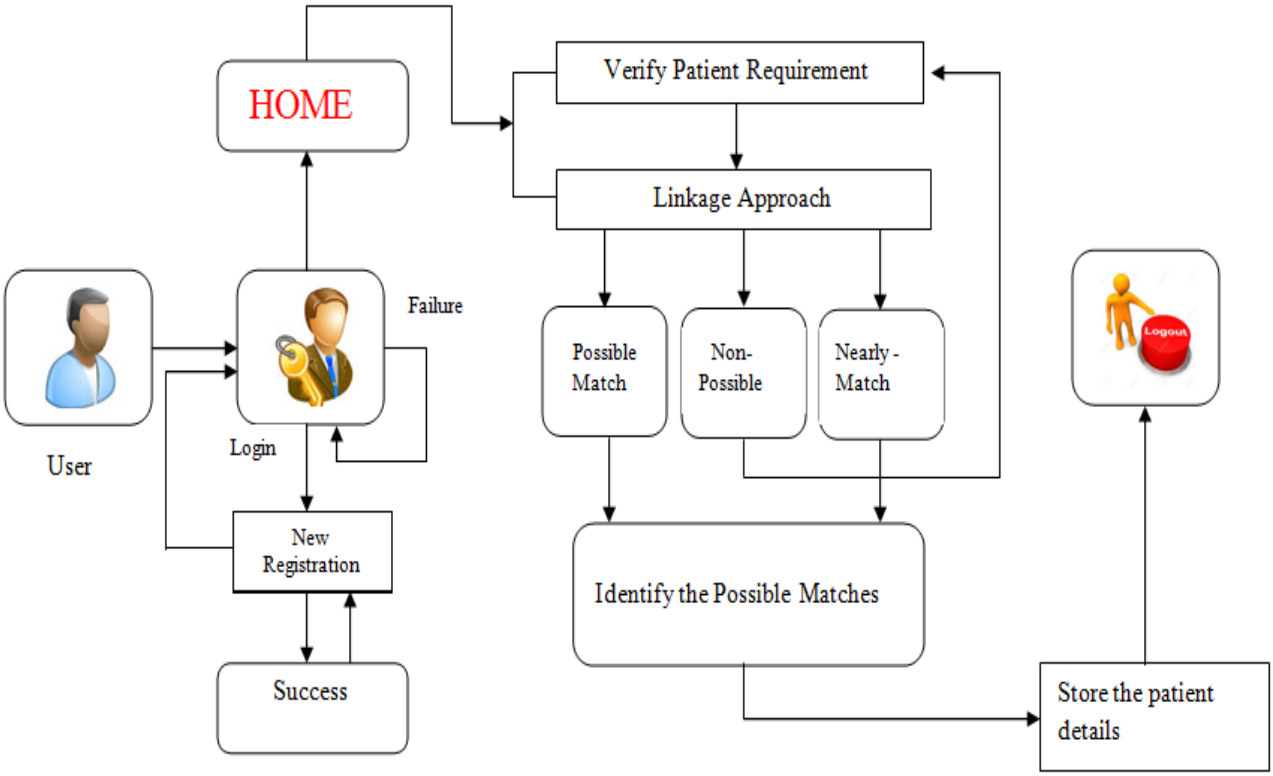

Fig 1: System Architecture

6. ADVANTAGE

- Big data is large number of data difficult to handle them by using these we easily maintain that record and privacy is maintained in this model, privacy-preserving technologies for record linkage.
- In this yak out is used for securing some methods like privacy guarantee, scalability and linkage quality is maintained.
- The laws and government policies can be used to ensure obedience and provide guidelines in privacy protection.
- The Ethical, Legal, and Social Implications (ELSI) studies are essential for biomedical data privacy, which provides a new methodology to biomedical research by recognizing, investigating and tackling the ELSI of studies involving human subjects.
- All government laws are maintained and ssn number is not used to find patient details.

7. CONCLUSION

- We discussed multiple aspects of big data privacy in the context of biomedical research.
- The "big" part of data privacy is because healthcare data often contain large scale clinical and genomic data, which are big in size and large in dimension.
- There are some unique challenges and off-the-shelf tools have difficulties in handling them.
- For example, the scalability concerns about fully homomorphism encryption and secure multiparty computing algorithms to deal with whole genome sequencing (WGS) data.

FUTURE ENHANCEMENT

- There are also challenges in safeguarding the outcomes of computation on high dimensional genomic data, which can easily exhaust the budget if not allocated carefully.
- We reviewed state-of-the-art privacy-preserving technologies for record linkage, synthetic data generation, and genomic data analysis.
- Despite of exciting progresses, there are many problems and emerging challenges need to be addressed and we believe good solutions to mitigate privacy risks in biomedical research require a joint effort from different communities (e.g, computer security, ELSI, biomedicine, genomics, etc.)..

ACKNOWLEDGEMENT

We express our thanks to all faculty and staff members of the Department of MCA, TPGIT. Also thanks to Surya informatics Pvt ltd, Chennai for their constant support and encouragement.

REFERENCES

[1] FOR .NET INSTALLATION
www.support.mircosoft.com
[2] FOR DEPLOYMENT AND PACKING ON SERVER
www.developer.com
www.15seconds.com
[3] FOR SQL
www.msdn.microsoft.com
[4] FOR ASP.NET
Asp.Net 3.5 Unleashed
[5] www.msdn.microsoft.com/net/quickstart/aspplus/default.com , www.asp.net