International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 17 (2017) pp. 6735-6740

Multi-Level Privacy Preservation and Transmission of Medical Records Using Cryptographic Technique

Karmel A

Associate Professor,School of Computing Sciences & Engineering,
VIT University, Chennai Campus, Chennai, TamilNadu, India.
Orcid: 0000-0003-2706-2239

Saravanan

Student, School of Computing Sciences & Engineering,
VIT University, Chennai Campus, Chennai, TamilNadu, India.

Abstract

Clinical Medical Records (CMR) is a Collection of patient's details. The normally dispersed gathering of CMRs crosswise over different clinical focuses proposes the need to incorporate the records in a centralized information archive. This is important to investigate numerous information diagnostic calculations which are not bolstered on circulated database. De-identification is the process used to prevent a person's identity form being connected with information. Cryptographic technique to facilitate patient de-identification as the records encrypted and decrypted to records and Addressing the key management using hashing techniques to construct public and private keys. End to end secure data design for the gathering preparing and dissemination of scrambled restorative information.

Keywords:

INTRODUCTION

Health care includes a various arrangement of information gathering frameworks, for example, medical records, health reviews, managerial records, and even wearable gadgets that mirror a patient's present area and level of health. This information is gathered and utilized by different entities, for example, doctor's facilities, medicinal services focuses, doctors, and health designs and the information is regularly gathered in various areas. For gathered medicinal information to be powerful and enhance understanding treatment it must be transported from a gadget, aggregated, and investigated to create comes about.

Organizations contribute scrambled genomic succession records into an incorporated store, where the overseer can perform inquiries, for example, recurrence counts, without unscrambling the information and we can utilize Linkage disequilibrium (LD) system growing fantastic SNP marker maps. At the point when connected to disease–gene mapping, LD is assessed through affiliation examination that requires the correlation of haplotype frequencies between the influenced and the control people.

An information digging strategy for LD mapping called Haplotype Pattern mining (HPM). We assess the effectiveness and precision of HPM inside our system that can be imparted to mind suppliers and security concerns connected to restorative information make it trying to make minimal effort data models that permit secure information transmission. The protection of the comparing people and the general security of the framework and de-recognizable proof techniques those evacuate unequivocally, and conceivably, distinguishing highlights.

Therapeutic information might be dissected and utilized years after gathering at various areas since information sources and care suppliers frequently work on various time scales and are topographically appropriated. The requirement for circulated and long haul restorative information stockpiling along these lines requires a viable security model to assign information get to. Current information get to designation models don't give end-to-end insurance. Viable assignments demonstrate must keep information encoded constantly and maintain a strategic distance from the need to share unscrambling keys to dodge security vulnerabilities.

We exhibit a protected data engineering and model to execute such a model with end-to-end information encryption while limiting information access to assigned beneficiaries. Current security models ensure restorative information with encryption innovation. An extra issue with current security models is that information is just encoded when it is sent to a pre-endorsed beneficiary. This point-to-point transmission time encryption requires earlier coordination of unscrambling keys.

That coordination necessity makes security defenselessness information hoodlums can endeavor to take the keys. In any case, the sharing of restorative information requires either

decoding the information or the sender and beneficiary sharing unscrambling keys.

RELATED WORKS

Privacy, confidentiality, and security are often used interchangeably to refer to the protection of personal health information. According to the Committee on Maintaining Privacy and Security in Health Care Applications of the National Information Infrastructure (National Research Council, 1997) privacy refers to an individual's desire to limit the disclosure of personal information while confidentiality refers to a condition in which information is shared or released in a controlled manner. Physicians and nurses, for example, have professional codes of ethics which include confidentiality. Security consists of measures that organizations implement to protect information, and includes technology as well as policies and procedures for safeguarding patient data.

There have been previous efforts to design network security architecture for the health care system that reduce the risk of data leakage. We propose a framework and associate query illustrate how genomic data can be collect and queries in an encrypted manner and securely store, share and query clinical genomics data using secure cryptographic methods.

There are several limitations to the current implementation. What is it about genomic data that makes it "identifiable"? To date, various privacy protection strategies have been designed to remove identifying information prior to sharing genomic data.

For the most part, existing genomic data privacy techniques can be roughly grouped into two approaches with distinct benefits and drawbacks: 1) data de-identification 2) query execution and 3) proxy re-encryption.

Secure Information Architecture

Our design's model usage is a Java-based web arrangement permitting 1) a patient to distribute scrambled information to a PRE-empowered server and 2) a specialist to safely get to, utilize, and designate information through that server. We actualized the model's cross section based PRE cryptosystem. Figure 1 outlines these runtime work processes.

The encryption and unscrambling operations are performed via self-ruling cryptosystems sent in every customer's web arrangement and the re-encryption operation is performed by the cryptosystem conveyed in the PRE Server.

For patient and specialist utilize cases, a Java Servlet sent in their separate customers scrambles or unscrambles information by utilizing a middleware executing the Java Native Interface (JNI) structure to interface with cryptosystem APIs for encryption and decoding.

The PRE Server's re-encryption utilize case is activated by a demand from Doctor 1 to a Java Servlet on the PRE Server that uses the JNI middleware to interface with the cryptosystem's re-encryption API. The JNI middleware referenced is examined later in this Section.

The model's cryptosystem operations in are performed utilizing open/private key sets created by every customer's cryptosystem. Every customer stores its key match and transmits their open key to the Policy Server, This is on the grounds that lone encoded information is transmitted between all customer segments and the PRE Server.

The model's Confidentiality factor is upgraded by its usage of quantum registering safe grid encryption systems to figure the scrambled information transmitted between the PRE Server and its customers. This advantage is most certainly not accessible in on-advertise open key security models, for example, RSA which have been crushed by quantum figuring methods.

The model gives the structure to this component in light of the fact that each PRE Server customer creates an open/private key combine. Actualizing this element requires including cryptographic mark calculation and confirmation APIs to the cryptosystem. Cryptographic hashing or cryptographic checksum APIs could give the premise to the essential mark work APIs.

The executed cryptographic mark convention will permit PRE Server customers to register and send a mark component over information they transmit subsequently giving the information beneficiary a way to confirm information uprightness.

The CIA group of three's Availability necessity will be confirmed through future improvement and testing exercises including model segments. Adaptability testing will require recognizing potential bottleneck indicates due the convenience of different clients and simultaneous solicitations.

At least, this will include streamlining transmission capacity utilize and runtime preparing for (1) demands made by customer parts to play out the cryptographic operations (2) the key dispersion conspire.

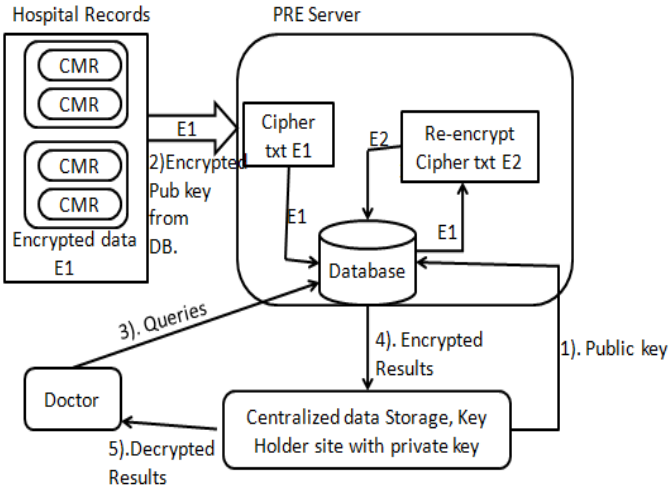


Figure 1: Prototype PRE server information Architecture

De-Identification

De-identification is the process used to prevent a person's identity from being connected with information. Security based on de-identification advocate the removal, or encryption of person specific identifiers, Such as name and social security number, initially associate with genomic records.

ARX anonymization tool is used for de-identification process to transforming structured, sensitive personal data using selected methods from the broad area of statistical disclosure control. Four perspectives of ARX anonymization tool

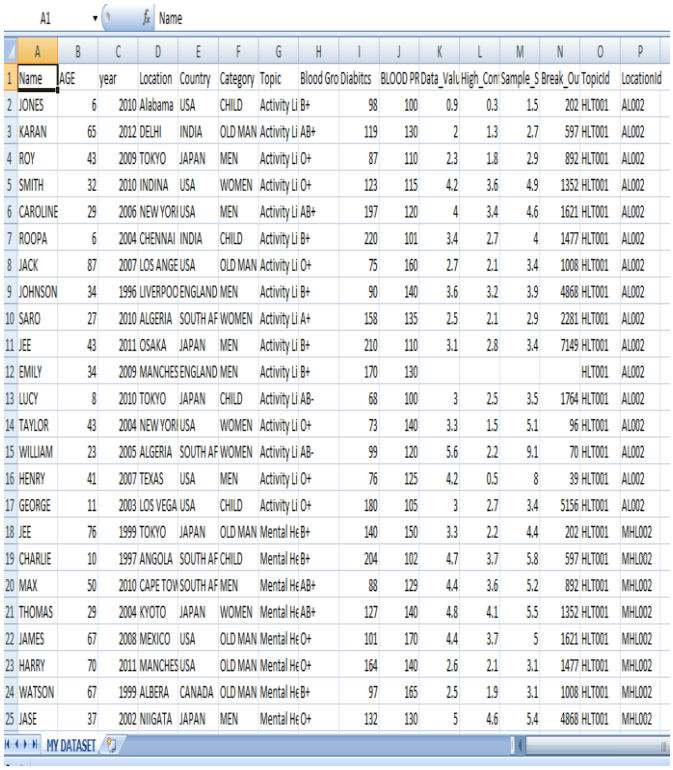
- Configuring privacy models, utility measures and transformation methods.
- Exploring the solution space.
- Analyzing data utility.
- Analyzing privacy risks.

An informational collection can be transported in into the instrument and trait meta information can be determined, including information sorts and characteristic properties as far as protection dangers. Second, speculation orders for semi identifiers or touchy properties can be made semi-consequently with worked in wizards or imported into the device from CSV documents. Third, protection models, the strategy for measuring information utility and further parameters, which control the change procedure, can be determined [1][2].

Amid the de-recognizable proof process, ARX builds an inquiry space comprising of an arrangement of changes that can be connected to the datasets. This inquiry space is then portrayed in view of the given security models and utility measure. This point of view enables clients to peruse the arrangement competitors distinguished by ARX and to choose intriguing changes for facilitate examination.

To evaluate the reasonableness of a particular change for a given use situation, this viewpoint underpins looking at changes of the info informational collection to the first information. To this end, it joins different graphical portrayals of aftereffects of uni-variate and bi-variate insights and backings cell-by-cell correlations [3].

To assess the appropriateness of a yield informational collection for machine learning purposes, the point of view likewise permits to dissect the order precision that can be accomplished with a nonexclusive calculated relapse technique [4][5].



Name	AGE	year	Location	Country	Category	Topic	Blood Gro	Diabitcs	BLOOD PR	Data_Valu	High_Corr	Sample_S	Break_Ou	TopicId	LocationId
JONES	6	2010	Alabama	USA	CHILD	Activity Li	B+	98	100	0.9	0.3	1.5	202	HLT001	AL002
KARAN	65	2012	DELHI	INDIA	OLD MAN	Activity Li	AB+	119	130	2	1.3	2.7	597	HLT001	AL002
ROY	43	2009	TOKYO	JAPAN	MEN	Activity Li	O+	87	110	2.3	1.8	2.9	892	HLT001	AL002
SMITH	32	2010	INDINA	USA	WOMEN	Activity Li	O+	123	115	4.2	3.6	4.9	1352	HLT001	AL002
CAROLINE	29	2006	NEW YORI	USA	MEN	Activity Li	AB+	197	120	4	3.4	4.6	1621	HLT001	AL002
ROOPA	6	2004	CHENNAI	INDIA	CHILD	Activity Li	B+	220	101	3.4	2.7	4	1477	HLT001	AL002
JACK	87	2007	LOS ANGE	USA	OLD MAN	Activity Li	O+	75	160	2.7	2.1	3.4	1008	HLT001	AL002
JOHNSON	34	1996	LIVERPOO	ENGLAND	MEN	Activity Li	B+	90	140	3.6	3.2	3.9	4868	HLT001	AL002
SARO	27	2010	ALGERIA	SOUTH AF	WOMEN	Activity Li	A+	158	135	2.5	2.1	2.9	2281	HLT001	AL002
JEE	43	2011	OSAKA	JAPAN	MEN	Activity Li	B+	210	110	3.1	2.8	3.4	7149	HLT001	AL002
EMILY	34	2009	MANCHES	ENGLAND	MEN	Activity Li	B+	170	130					HLT001	AL002
LUCY	8	2010	TOKYO	JAPAN	CHILD	Activity Li	AB-	68	100	3	2.5	3.5	1764	HLT001	AL002
TAYLOR	43	2004	NEW YORI	USA	WOMEN	Activity Li	O+	73	140	3.3	1.5	5.1	96	HLT001	AL002
WILLIAM	23	2005	ALGERIA	SOUTH AF	WOMEN	Activity Li	AB-	99	120	5.6	2.2	9.1	70	HLT001	AL002
HENRY	41	2007	TEXAS	USA	MEN	Activity Li	O+	76	125	4.2	0.5	8	39	HLT001	AL002
GEORGE	11	2003	LOS VEGA	USA	CHILD	Activity Li	O+	180	105	3	2.7	3.4	5156	HLT001	AL002
JEE	76	1999	TOKYO	JAPAN	OLD MAN	Mental He	B+	140	150	3.3	2.2	4.4	202	HLT001	MHL002
CHARLIE	10	1997	ANGOLA	SOUTH AF	CHILD	Mental He	B+	204	102	4.7	3.7	5.8	597	HLT001	MHL002
MAX	50	2010	CAPE TOW	SOUTH AF	MEN	Mental He	AB+	88	129	4.4	3.6	5.2	892	HLT001	MHL002
THOMAS	29	2004	KYOTO	JAPAN	WOMEN	Mental He	AB+	127	140	4.8	4.1	5.5	1352	HLT001	MHL002
JAMES	67	2008	MEXICO	USA	OLD MAN	Mental He	O+	101	170	4.4	3.7	5	1621	HLT001	MHL002
HARRY	70	2011	MANCHES	USA	OLD MAN	Mental He	O+	164	140	2.6	2.1	3.1	1477	HLT001	MHL002
WATSON	67	1999	ALBERA	CANADA	OLD MAN	Mental He	B+	97	165	2.5	1.9	3.1	1008	HLT001	MHL002
JASE	37	2002	NIIGATA	JAPAN	MEN	Mental He	O+	132	130	5	4.6	5.4	4868	HLT001	MHL002

Figure 1: Input Genomic Data

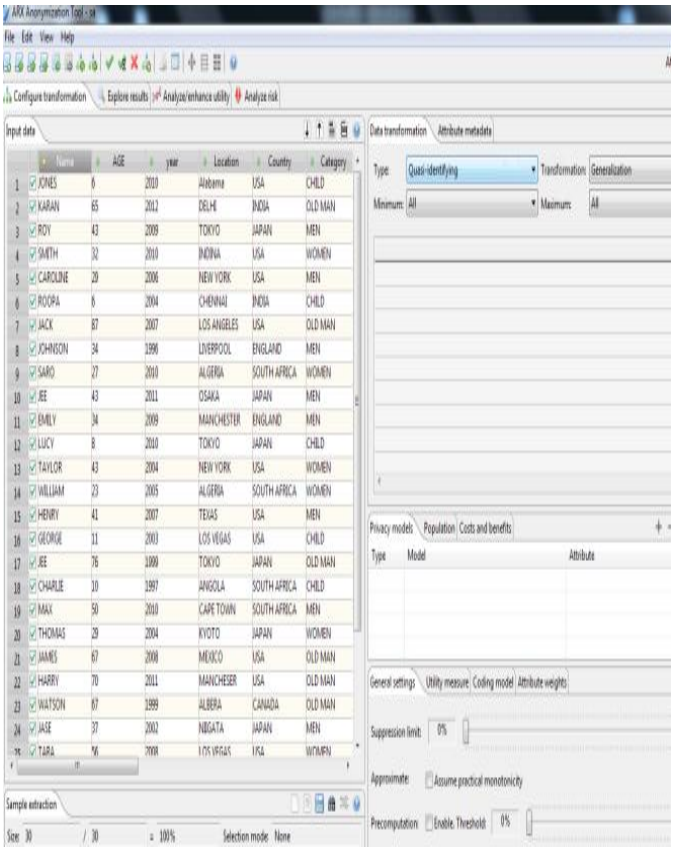


Figure 2: Data Transformation

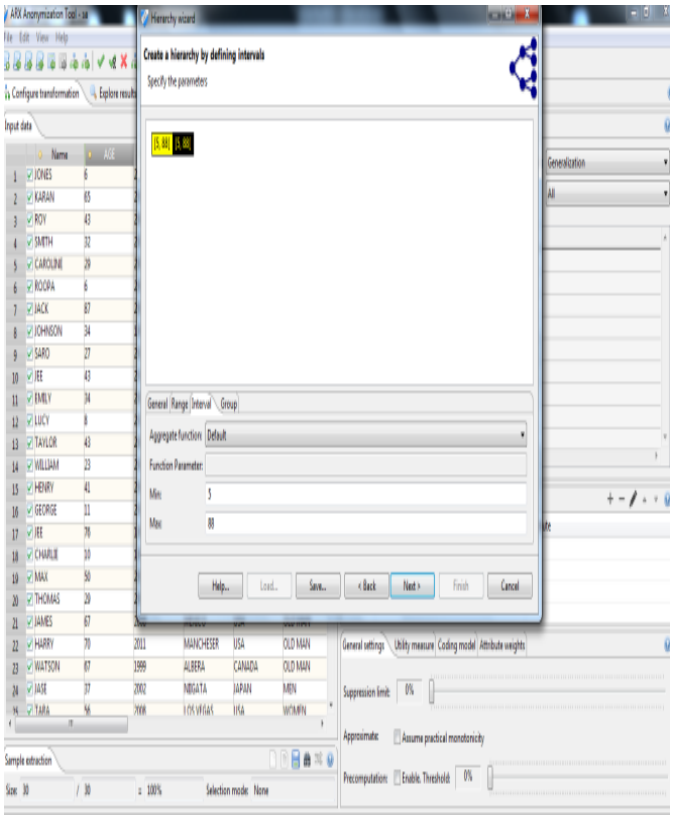


Figure 3: Specify the type of hierarchy

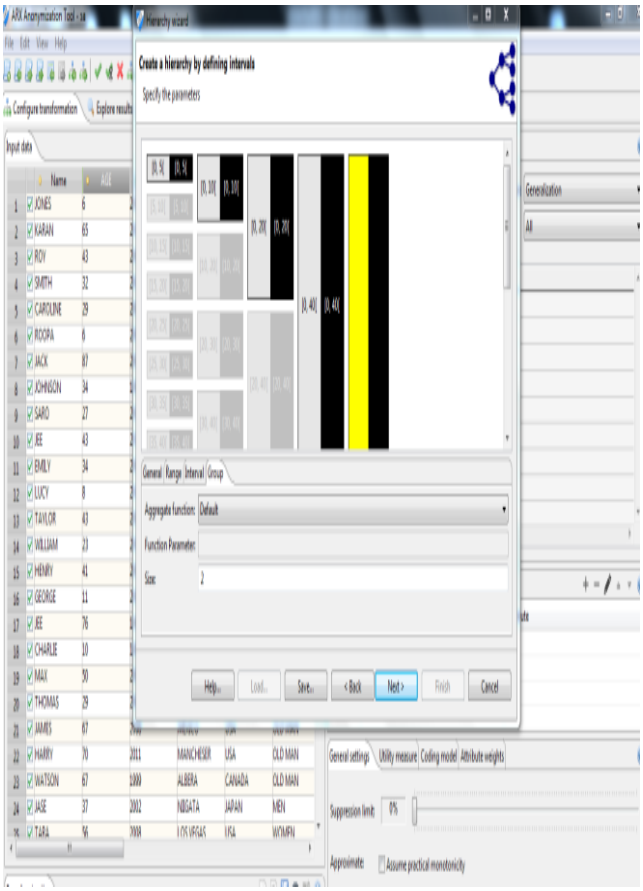


Figure 4: Specify the Parameters

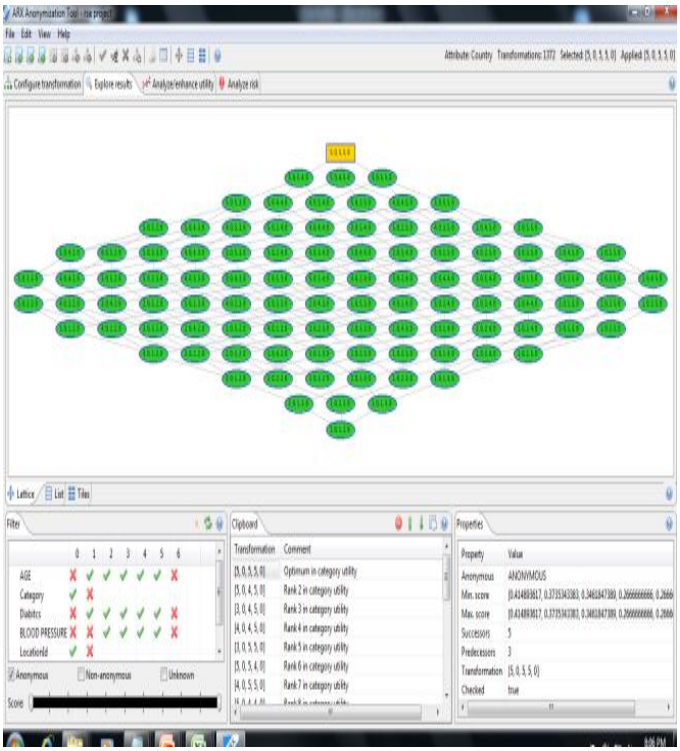


Figure 5: Explore Results

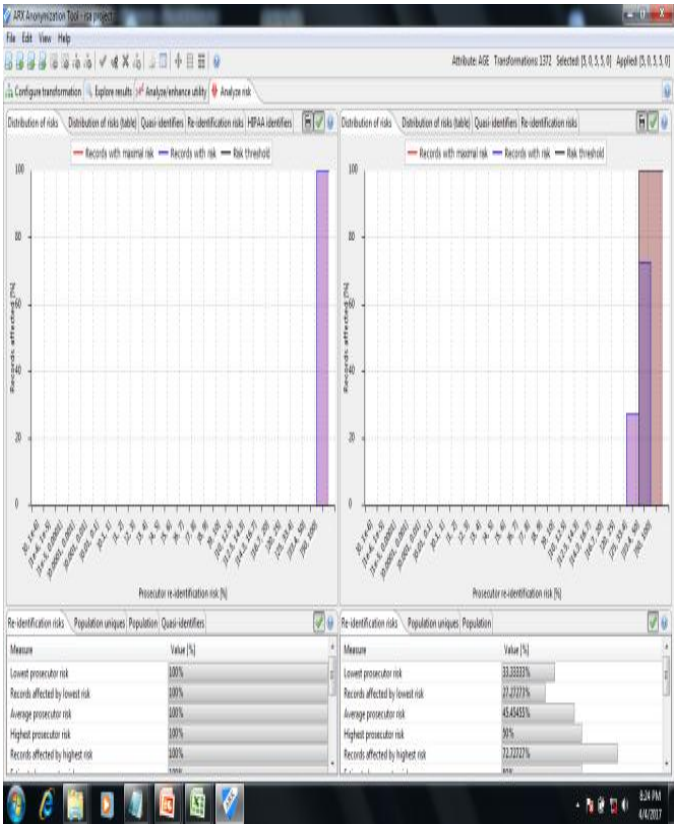


Figure 6: Analyzing risk

Different measurements reflecting security dangers are exhibited. Models actualized by ARX incorporate re-

recognizable proof dangers for the prosecutor, columnist and advertiser aggressor models and gauges of populace uniqueness, which can be ascertained utilizing diverse measurable models.

In addition, the viewpoint additionally gives access to a technique to recognizing traits which must be changed or adjusted when de-distinguishing information in consistence to the Safe Harbor strategy for the US Health Insurance Portability and Accountability Act (HIPAA identifiers) and a technique for discovering potential semi identifiers[6][7].

A1 Name

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Name	AGE	year	Location	Country	Category	Topic	Blood Gro	Diabitcs	BLOOD PR	Data_Valu	High_Con	Sample_S	Break_Ou	Topicid	Locationid
2	JONES	[5,20[	2010	Alabama	USA	CHILD	Activity Li	B+	98	100	0.9	0.3	1.5	202	HLT001	AL002
3	ROOPA	[5,20[	2004	CHENNAI	INDIA	CHILD	Activity Li	B+	220	101	3.4	2.7	4	1477	HLT001	AL002
4	LUCY	[5,20[	2010	TOKYO	JAPAN	CHILD	Activity Li	AB-	68	100	3	2.5	3.5	1764	HLT001	AL002
5	GEORGE	[5,20[	2003	LOS VEGA	USA	CHILD	Activity Li	O+	180	105	3	2.7	3.4	5156	HLT001	AL002
6	CHARLIE	[5,20[	1997	ANGOLA	SOUTH AF	CHILD	Mental He	B+	204	102	4.7	3.7	5.8	597	HLT001	MHL002
7	LENKA	[5,20[	2010	MUMBAI	INDIA	CHILD	Mental He	B+	109	115	4.1	3.7	4.4	7149	HLT001	MHL002
8	KELLY	[5,20[	1998	NEW YORI	USA	CHILD	Mental He	O+	189	135	4.3	3.7	5	1764	HLT001	MHL002
9	CAROLINE	[20,40[	2006	NEW YORI	USA	MEN	Activity Li	AB+	197	120	4	3.4	4.6	1621	HLT001	AL002
10	JOHNSON	[20,40[	1996	LIVERPOO	ENGLAND	MEN	Activity Li	B+	90	140	3.6	3.2	3.9	4868	HLT001	AL002
11	EMILY	[20,40[	2009	MANCHES	ENGLAND	MEN	Activity Li	B+	170	130					HLT001	AL002
12	JASE	[20,40[	2002	NIIGATA	JAPAN	MEN	Mental He	O+	132	130	5	4.6	5.4	4868	HLT001	MHL002
13	ROY	[40,60[	2009	TOKYO	JAPAN	MEN	Activity Li	O+	87	110	2.3	1.8	2.9	892	HLT001	AL002
14	JEE	[40,60[	2011	OSAKA	JAPAN	MEN	Activity Li	B+	210	110	3.1	2.8	3.4	7149	HLT001	AL002
15	HENRY	[40,60[	2007	TEXAS	USA	MEN	Activity Li	O+	76	125	4.2	0.5	8	39	HLT001	AL002
16	MAX	[40,60[	2010	CAPE TOM	SOUTH AF	MEN	Mental He	AB+	88	129	4.4	3.6	5.2	892	HLT001	MHL002
17	LARRY	[40,60[	2009	ALASKA	USA	MEN	Mental He	A+	201	120	5.4	2	8.8	70	HLT001	MHL002
18	KARAN	[60,80[	2012	DELHI	INDIA	OLD MAN	Activity Li	AB+	119	130	2	1.3	2.7	597	HLT001	AL002
19	JEE	[60,80[	1999	TOKYO	JAPAN	OLD MAN	Mental He	B+	140	150	3.3	2.2	4.4	202	HLT001	MHL002
20	JAMES	[60,80[	2008	MEXICO	USA	OLD MAN	Mental He	O+	101	170	4.4	3.7	5	1621	HLT001	MHL002
21	HARRY	[60,80[	2011	MANCHES	USA	OLD MAN	Mental He	O+	164	140	2.6	2.1	3.1	1477	HLT001	MHL002
22	WATSON	[60,80[	1999	ALBERA	CANADA	OLD MAN	Mental He	B+	97	165	2.5	1.9	3.1	1008	HLT001	MHL002
23	ROONEY	[60,80[	2012	LONDON	ENGLAND	OLD MAN	Mental He	B+	159	165					HLT001	MHL002
24	SMITH	[20,40[	2010	INDINA	USA	WOMEN	Activity Li	O+	123	115	4.2	3.6	4.9	1352	HLT001	AL002
25	SARO	[20,40[	2010	ALGERIA	SOUTH AF	WOMEN	Activity Li	A+	158	135	2.5	2.1	2.9	2281	HLT001	AL002

ARX

Figure 7: De-identification Output

Query Execution

One of the more typical undertakings that genome based researchers need to perform is to decide what number of tests fulfill contain qualities. At the point when the extent of the database is relied upon to stay steady.

Be that as it may, as biomedical research examinations turn out to be more mind boggling and the amount of populace based information develops, it will be important to infer more effective system for secure inquiry execution.

Our outcomes with an example inspecting procedure demonstrate that running check questions more than 50000 arbitrarily picked SNP groupings might be adequate to correctly assess the first tally inquiry result.

Proxy Re-Encryption

A conclusion to end secure data engineering for the accumulation, preparing and conveyance of scrambled restorative information utilizing a grid based variation of proxy re-encryption (PRE).In this schemes are similar to traditional symmetric or asymmetric encryption schemes, with the addition of two functions: 1) Delegation 2) transitivity.

Delegation

Grid based Proxy Re-Encryption (PRE) gives a powerful way to deal with designate decoding capacity. There are existing PRE outlines in view of cross section encryption. That current cross section based PRE plot utilizes almost indistinguishable calculations for key era, encryption, and unscrambling as the Homomorphic Encryption (HE) conspire in this manner permitting utilization of comparative confirmations of security.

The essential usage distinction amongst HE and PRE is one of parameterization. Investigation from [9] shows that the PRE plot we present can be parameterized to utilize ring measurements of n = 1024 and figure content module of under 32 bits to encode 1024 plaintext bits.

Tentatively, these parameterizations result in 1) the disengaged re encryption handle running in less than 40 ms for every kb of information and 2) the encryption and unscrambling forms running in 6.1 and 7.9 ms for each kb of information.

These outcomes show the attainability of utilizing the PRE display for commonsense medicinal information applications and notwithstanding for time-basic restorative execution prerequisites.

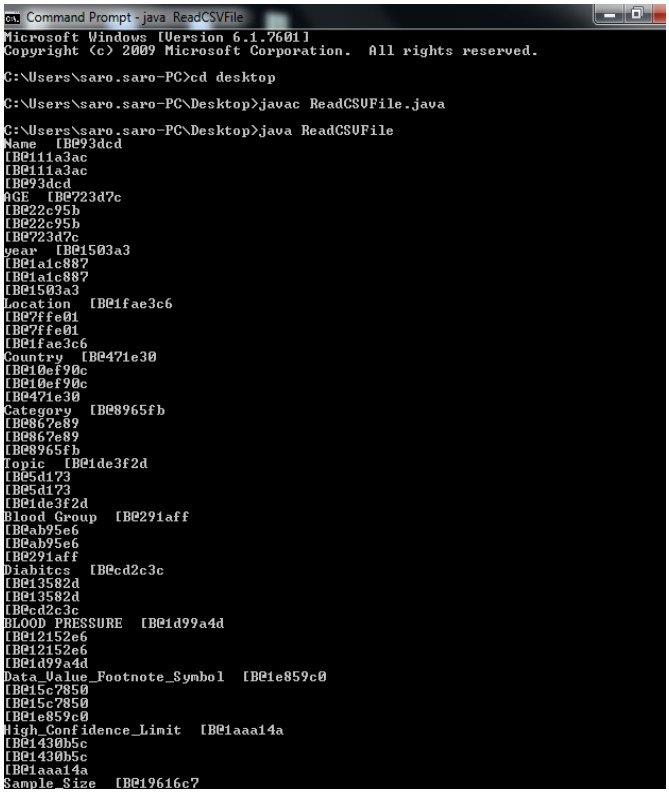


Figure 8: Data re-encryption results

CONCLUSION

We display secure data engineering for designating medicinal information get to that ensures information consistently by applying a grid based variation of Proxy Re-Encryption (PRE) to

1) give end-to-end encryption and

2) keep the need to share decoding keys. Exploratory outcomes show that our engineering tends to the lacks in current security models noted in Section I and the assessment measures noted in Section II.

While we concentrate on security and secrecy concerns, our design can be expanded to give Integrity and Availability assurances in this way fulfilling the CIA set of three. For instance, cryptographic marking strategies could fulfill Integrity and administration replication could fulfill Availability against equipment issues and refusal of-benefit assaults.

An essential advantage of our design's end-to-end cross section based encryption conspire is that it empowers medicinal services substances to safely utilize ease distributed computing situations to share information while additionally altogether decreasing defenselessness to insider assaults.

For instance, our engineering limits information get to just to the framework managers who have unscrambling keys notwithstanding when scrambled registering is facilitated on exclusive servers. Furthermore, our design averts access to unscrambled information until the point when it achieves its planned beneficiary.

The advantages of our design will along these lines diminish the operational expenses of exceedingly directed businesses, for example, human services, where administrative consistence confines the capacity to outsource information security calculations.

REFERENCES

[1] L. Burnett, K. Barlow-Stewart, A. Proos, and H. Aizenberg, "The 'Gene Trustee': A universal identification system that ensures privacy and confidentiality for human genetic databases," *J. Law Med.*, vol. 10, no. 4, pp. 506–513, May 2003.

[2] G. de Moor, B. Claerhout, and F. de Meyer, "Privacy enhancing techniques—The key to secure communication and management of clinical and genomic data," *Methods Inf. Med.*, vol. 42, no. 2, pp. 148–153, 2003.

[3] M. Canim, M. Kantarcioglu, and A. Inan, "Query optimization in encrypted relational databases by vertical schema partitioning," in *Proc. Secure Data Manage.*, 2009, pp. 1–16

[4] Y. Polyakov, K. Rohloff, G. Sahu, and V. Vaikuntanathan, "Proxy re-encryption, lattice encryption, software engineering, delegating access control," Submitted.

[5] W. H. Maisel and T. Kohno, "Improving the security and privacy of implantable medical devices," *New England journal of medicine*, vol. 362, no. 13, p. 1164, 2010.

[6] Arnab Deb Gupta, Yuriy Polyakov, and Kurt Rohloff, "Secure Access Delegation of Encrypted Medical Information" 2013.

[7] Mustafa Canim, Murat Kantarcioglu and Bradley Malin, *ieee transactions on information technology in biomedicine, vol. 16, no*. 1, 2012.

[8] Barbara Carminati, Elena Ferrari, Michele Guglielmi, " Secure information sharing on support of emergency management," IEEE International Conference on Privacy, Security, Risk, and Trust, and ieee international conference and computing, 2011.

[9] F. Wozak, T. Schabetsberger, and E. Ammenwerth, "End-to-end security in telemedical networks—A practical guideline," *Int. J. Med. Inf.*, vol. 76, pp. 484–490, 2007.

[10] K. Benitez and B. Malin, "Evaluating re-identification risks with respect to the HIPAA Privacy Rule," J. Amer. Med. Informat. Assoc., vol. 17, no. 2, pp. 169–177, 2010.