JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2866 - 2871
https://publishoa.com
ISSN: 1309-3452

A Cryptography Based Secure Privacy Preservation of the Sensitive Data in Health Care Domain

S.Joseph Gabriel[1*], Dr.P.Sengottuvelan[2*]

[1*]Research Scholar, Department of Computer Science, Periyar University,Salem
talk2anette67@gmail.com
[2*]Associate Professor, Department of Computer Science, P.G ExtensionCentre,
Periyar University, Dharmapuri
sengottuvelan@gmail.com

(Corresponding Author: S.Joseph Gabriel)

ABSTRACT

With the advent of technical advancements in the field of digitalization there are a lot of data available in the web. But there are also many sensitive data which requires protection from adversaries. Privacy preservation is very crucial to preserve the sensitive data that is stored inthe cloud. Cryptographic techniques provide a good mechanism to protect the sensitive data.In this work we discuss about how we use the cryptographic techniques to preserve theprivacy of data and we develop a better framework for protecting the sensitive data. In thiswork we utilized a Diffie-Hellman key exchange algorithm for securely communicatingbetween the client and server environment and how this algorithm provides a secure measureis being discussed and it is compared with the RSA algorithm and also a frame work foreffective management of the details of the patients and providing proper support to thepatients and how there is a tremendous change brought by the ubiquitous computing and howit can be used to enhance the healthcareis also being discussed. There is a comparison of thecomputation time of the algorithms is being considered and which algorithm is optimal isbeingsuggested.

Keywords-Cryptography,Diffie-HellmanKey Exchange,RSA,Ubiquitous Computing

Introduction

Data mining can be viewed as a method for obtaining information from a huge volume ofdata. Information mining manages the sort of examples that can be mined. Based on the sortof information to be mined, there are two classifications of capacities engaged with DataMining namely Classification and Prediction, Distinct Function. Cryptography is a procedureof making sure about data and correspondences through utilization of codes with the goal thatlone those individuals for whom the data is processed.Accordingly forestalling unapprovedadmittance to data. The prefix "crypt" signifies "covered up" and postfix graph signifies"composing". In Cryptography the strategies which are used to shield data are gotten fromnumerical ideas and a bunch of rule-based counts known as calculations to change overmessages in manners that make it difficult to interpret it. These calculations are utilized forcryptographickeyproduction,advancedmarking,confirmationtoensureinformationsecurity, navigation on web and to secure secret exchanges, for example we can consider theonline transactions. This work discusses the framework of the datamining which is applied tothe healthcare domain and how the data is securely shared using the cryptographic algorithmsand its performance are being monitored. This research paper is organized as follows thesection 2 describes about the recent developments in the research, section 3 describes aboutthe methodology being developed and section 4 describes about the algorithms used andsection5 describesaboutthe performanceandsection 6 concludesthework.

Related Works

Privacy preserving data mining is gaining significant importance in recent period because oftheincreaseintheuseofdigitalmeans.Therewerevariousworkswhichusescryptographictechniques and some works which doesn't use these techniques and the direction of their searchis obtained out of this survey.

Consider a circumstance where various clinical foundations wish to lead a joint investigationforsomemutualpreferenceswithoutrevealinginconsequentialinformation.Inthiscircumstance, researchconcerning indications, finding and medication reliant on variouslimits is to be coordinated and all the while security of individuals is to be

guaranteed.Cryptographicproceduresareundeniablyinferredforsuchcircumstanceswherevarioussocial occasions collaborate to deal with results or offer non-sensitive mining results andhence avoiding revelation of fragile information.Cryptographic systems find its utility insuch circumstances considering two reasons: First, it offers an inside and out described modelfor security that consolidates procedures for exhibiting and estimating it. Second an immensegame plan of cryptographic counts and works to complete security protecting data miningfiguring are available in this space. The data may be passed on among different partnersvertically or equally. In vertically allocated data among different colleagues, the individualcomponents may have different credits of same course of action of records and in case of on alevel plane isolated data, solitary recordsare spread out over various components, all ofwhich has comparative course of action of attributes. By far most of the security sparing dispersed data mining counts reveals nothing other than the decisive result. Kantarcioglu and Clifton melded cryptographic methodologies to ensure security in association rule miningover equitably distributed data to restrict information sharedand at thesame time addingclose to no overheads tothe mining task. Lindell and Pinkas have analyzed how to deliverID3 decision trees on a level plane allocated. Yang et al. in their research work have analyzed a response for equally separated data where each customer has a private access just to theirown record. Vaidya and Clifton were the fundamental who thought how secure connectionruleburrowingshouldbefeasibleforverticallydistributed.DuandZhanpresentedaresponse for creating ID3 on vertically divided examining two events for mining. Vaidya andClifton developed a Naive Bayes classifier for sparing insurance on vertically allocated.Vaidya and Clifton in proposed a procedure for gathering over verticallyallocated data. Allof these procedures are almost established on an extraordinary encryption known as SecureMultipartyComputation(SMC) advancement.Cryptographictechniquesensurethatthechanged data is exact and secure anyway this strategy fails to pass on when more than two orthree social events are incorporated. Likewise, the data mining results may break the securityof individual records. There exist a nice number of courses of action in case of semi-genuinemodelsyetontheoffchancethatthereshouldbeaneventofmalignantmodelslessassessmentshavebeen made.

Dwivedi, A. D., Srivastava, G., Dhar, S., and Singh, R. in their work proposed the utilizationof blockchain for safely overseeing the medical care information. In any case, the use ofblockchain is exorbitant and requires more data transmission and significant expense forcalculation

Erlingsson, Ú., Pihur, V., andKorolova, in their work built up a framework which hasRandomized aggregatable security protection is utilized for getting the insights of publiclysupportingfromcustomerprogramming.Thisphilosophypermitsalotofcustomerinformation to be concentrated yet without permitting to investigate the subtleties of theindividual trees. It gives high utility investigation of the information which is gathered by theclient. This strategy gives a decent measure of protection conservation and it is applied togenuinejust as engineeredinformation fortesting its validity.

Lin, J. L., and Liu, J. Y. C in their examination work comprehends about the issue of privacypreservingAssociationrulemining.Theydevelopedarandomizationstrategyfortheexchangeforensuringtheprotectionoftheinformation.Theyproposedacalculationforby

andbyacquiringtheitemsetwhichhappenseverynowandagainfrombothgenuineexchanges.Thiswork likewisegivessome improvement in themined outcomes.

Lohiya, S., and Ragha, L in their work randomized the first information and afterward thespeculation approach is applied over the information which is randomized. They found thattheirstrategyensurestheinformationandcombinestheprotectionwithadecentexactnessandthereis no much misfortune inthedataand makes the informationusable.

Sharma, S., Chen, K., and Sheth, A. in their work built up a customized medical services dataframework for illness observing and they examined their work and checked it how muchsecurityis obtained.

Zhang, C., Zhu, L., Xu, C., and Lu, R in their exploration work proposed anforecastframework where there is a protection of security where the clinical information will bescrambledandthe informationisput awayin the cloudworker anditisput awayforadditional handling like the preparation cycle by utilizing the single layer perceptron learningmodel.

Zhou,J.,Cao,Z.,Dong,X.,andLin,X.intheirworkproposedaprivacypreservinghomomorphic and they built up a protection saving capacity relationship coordinating fromthe clinical content mining and they found that their model accomplished high security andgreatexecution.

Zhu, Y., and Liu,L. (2004, August) intheir methodologybuiltup a plan for randomizationforprotectionsavingintheassessmentandtheyproposedasystemforrandomizationutilizing the blend models. The impact of randomization on information mining is estimatedby the measure of data which is lost and the corruption of the exhibition and they measure byreenactmentsand showed how theprotection isachieved.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2866 - 2871
https://publishoa.com
ISSN: 1309-3452

It empowers the retail areas to show clientreactionandencouragesthefinancialareatoforeseeclientbenefit.Itservesnumerouscomparable areas, for example, fabricating, telecom, medical care, car industry, training, andsome more. Information mining holds amazing potential for medical care benefits because ofthe remarkable development in the quantity of electronic wellbeing records. Already Doctorshold tolerant data in thepaperwhere the information was very hard tohold.

Digitalizationand development ofnewstrategiesless inhuman endeavors andmakeinformationeffectively assessable. For instance, the PC keeps a gigantic measure of patient informationwithprecision,anditimprovesthenatureoftheentireinformationtheexecutive'sframework. This is where information mining has demonstrated to be amazingly helpful.Researchersareusingvariousmethodologieslikegroups,characterization,choicetrees,neuralorganizations,andtimearrangementtodistributorsearch.Bethatasitmay,Healthcarehasreliablybeendelayedtojointhemostrecentexaminationintoregular practice. Information mining has been utilized seriously and generally by various businesses.Inmedicalcare,informationminingisturningouttobemoremainstreamthesedays.Information mining applications can inconceivably profit all gatherings who are associatedwith the medical care industry. For instance, information mining can help the medical careindustry in extortion recognition and misuse, client relationship the board, compelling patientconsideration, and best practices, reasonable medical services administrations. There is a lotofinformationcreatedbymedicalcareexchangesareexcessivelymindbogglingandenormousto be preparedanddissected by traditional strategies.

Theframeworkisasbelow

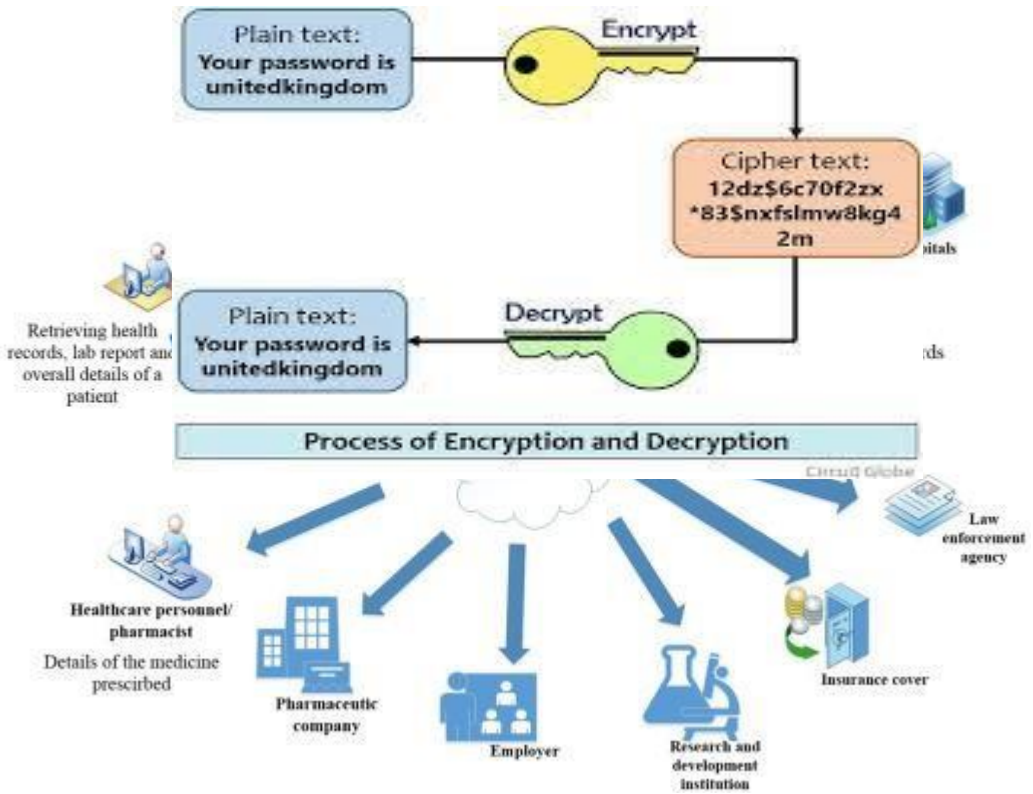

Figure-1.FrameworkforCryptographybasedprivacypreservationinhealthcaredomain

1. Cryptographicalgorithmsforsecurecommunication

Cryptography is concerned with process of conversion of a plain text into a text which isunintelligible to humans. It is a phenomenon of storing and transmission of data and only thepersons intended to obtainthe data alone can obtain the data and otherscannot access thedata.

Keybenefitsarrivedout ofCryptographyareasperthefollowing:

Confidentiality

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2866 - 2871
https://publishoa.com
ISSN: 1309-3452

Data must be obtained to by the individual for whom it is proposed and no other individualaside fromthem can get to it.

Integrity
Datacan'tbechangedawayorprogressamongsenderandexpectedcollectorwithnoexpansionto databeing distinguished.

Non-renouncement
Themaker/senderofdata can'tdenytheirexpectationtosenddataatlaterstage.

Authentication
Thecharactersofsender andcollector areaffirmed.Justasobjective/rootofdataisaffirmed.KindsofCryptography:

Therearemainlythreecryptographytypes:

1. SymmetricKey Cryptography:
It is an encryption framework where the sender and collector of message utilize a solitaryregular key to encode and decode messages. Symmetric Key Systems are quicker and morestraightforward however the issue is that sender and recipient need to by one way or anothertrade key in a protected way. The most well-known symmetric key cryptography frameworkisData Encryption System (DES).

2. HashFunctions:
Thereis no utilization of anykey inthiscalculation.

3. AsymmetricKeyCryptography:
Under this framework a couple of keys is utilized to scramble and decode data. A public keyisutilized forencryptionand aprivatekey is utilized fordecoding.

Diffie-HellmanAlgorithm
Diffie Hellman calculation is basically a convention that is utilized for Exchange of keys.Utilizing this intuitive convention two gatherings will determine a typicalsecret key byimparting one another. The security of Diffie-Hellman calculation is predominantly foundedonthe trouble ofregistering the discrete logarithms.

Algorithm:
ThealgorithmofDiffie-Hellmankeyexchangeisasfollowsanditisexplainedwithgeneralstepsand exampleas follows

Begin
BobandAlicegetpublicnumbers P=23, G=9
Bob chose a private key a = 4 and Alice chosen a private key b = 3
Aliceand Bob process publickeys
Bob:x =(9^4 mod 23)= (6561 mod 23) =6
Alice: y = (9^3 mod 23) = (729 mod 23) = 16
Aliceand Bob exchangetheirpublickeys
Bob gets public key y =16 and Alice gets public key x = 6
Aliceand Bob process symmetrickeys
Bob: ka =y^amod p = 65536 mod 23 = 9
Alice:kb = x^b mod p=216 mod 23 =9
9 is the secret key which is shared between them.
Stop

ThisalgorithmiscomparedwiththeRSAalgorithmandthelevelofsecurityandthetimetakento access therecords by thepatientand themedical teams wereanalyzed.

Resultsanddiscussion
Theentireworkwasimplementedinjavaandthedatasetsfortheresearchworkwasobtainedfrom MIMICandthe time takentoobtainthe electronic healthrecord bythepatients and the medical team was calculated and the accuracy of the classification algorithmsusedwerecalculated andthe values arepresented in a Table-1

Table-1
PerformanceoftheCryptographicalgorithms

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 2866 - 2871
https://publishoa.com
ISSN: 1309-3452

S.No	Datasetused	Algorithmused	Keysize(in Bits)	Computational Time(m)
1	MIMIC	DiffieHellmen	768	4.6225
2	MIMIC	DiffieHellmen	1024	6.3250
3	MIMIC	DiffieHellmen	2048	10.3876
4	MIMIC	RSA	1024	8.3456
5	MIMIC	RSA	2048	14.5276
6	MIMIC	RSA	4096	18.8253

From the above table we can observe that the Diffie Hellman key exchange algorithm canperformwellwhenitiscomparedwiththeRSAalgorithmanditisinferred from theexperimentation study that increase in time is proportional to the key size increase and theDiffie Hellman key exchange algorithm proves effective for protecting the privacy of the datawhenit is exchanged.

5. Conclusion

Thus a privacy preserving framework is being developed which considers the usage of theDiffie Hellman key exchange algorithm for securely transmitting the data between the clientand the server and there are various variations being performed in the size of the keys whichare used for the encryptionand the same processisapplied for the RSA algorithmand itfoundfromtheexperimentalanalysisthattheDiffieHellmankeyexchangealgorithmperformsbetterintermsoflesscomputation altimeanditprovestobeaneffectivecryptographic algorithm. The usage of cryptographic algorithm is very much essential forensuring a safe communication between the client and server. Information mining can beconsideredasacalculationwhichdevoursinformationastheinformationandproducesdesigns like item sets, rules for grouping, rules for Association. Be that as it may, due todatamining there is a danger to the protection of the information in light of the assortment ofinformationandhowsafelytheinformationisputawayandrecoveredwithoutlosingsecurity.Sincealargeportionoftheassociati onsarerelieduponthecloud-basedadministrations for the capacity of information the undertaking of safeguarding the protectionis exceptionally pivotal. Hence this work develops a framework for securing exchanging theelectronic health record of patients in a secured fashion and also the computational time isalsovery less.

Conflicts of Interest

The authors declare they have no conflicts of interest.

References

1. Agrawal, R., & Srikant, R. (2000, May). Privacy-preserving data mining. In *Proceedings ofthe2000ACMSIGMOD Internationalconferenceon Managementof data*(pp.439-450).

2. Aggarwal,C.C.,&Philip,S.Y.(2004,March).Acondensationapproachtoprivacypreserving data mining. In International Conference on Extending Database Technology (pp.183-199).Springer, Berlin, Heidelberg.

3. Dwivedi, A. D., Srivastava, G., Dhar, S., &Singh, R. (2019). A decentralized privacy-preservinghealthcareblockchain forIoT.*Sensors*,*19*(2), 326.

4. Erlingsson,Ú.,Pihur,V.,&Korolova,A.(2014,November).Rappor:Randomizedaggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSACconferenceon computerand communicationssecurity*(pp.

1054-1067).

5. Evfimievski, A., & Grandison, T. (2009). Privacy-preserving data mining. In *Handbook ofResearch on Innovations in Database Technologies and Applications: Current and FutureTrends*(pp. 527-536).IGIGlobal.

6. Kantarcioglu, M., & Clifton, C. (2004). Privacy-preserving distributed mining of associationrulesonhorizontallypartitioneddata. *IEEEtransactionsonknowledgeanddataengineering*,*16*(9), 1026-1037.

7. Lin, J. L., & Liu, J. Y. C. (2007, March). Privacy preserving itemset mining through faketransactions. In *Proceedings of the 2007 ACM symposium on Applied computing* (pp. 375-379).

8. Liu, K., Giannella, C., & Kargupta, H. (2006, September). An attacker's view of distancepreserving maps for privacy preserving data mining. In European Conference on Principles ofDataMining andKnowledgeDiscovery(pp. 297-308). Springer,Berlin, Heidelberg.

9. Lohiya, S., & Ragha, L. (2012, November). Privacy preserving in data mining using hybridapproach.In *2012FourthInternationalConferenceonComputationalIntelligenceandCommunicationNetworks*(pp. 743-746).IEEE.

10. Pika,A.,Wynn,M.T.,Budiono,S.,terHofstede,A.H.,vanderAalst,W.M.,&Reijers,H.

11. A.(2019, September). Towards privacy-preserving process mining inhealthcare.

12. In *InternationalConferenceonBusinessProcessManagement* (pp.483-495).Springer,Cham.

13. Pika, A.,Wynn,M.T.,Budiono,S.,TerHofstede,A. H.,vanderAalst,W.M.,&Reijers,H.
14. A.(2020).Privacy-preservingprocessmininginhealthcare.*Internationaljournalofenvironmental researchand publichealth*, *17*(5),1612.

15. Qi, X., Mei, G., Cuomo, S., & Xiao, L. (2020). A network-based method with privacy-preserving for identifying influential providers in large healthcare service systems. *FutureGenerationComputer Systems*.

16. Sharma, S., Chen, K., & Sheth, A. (2018). Toward practical privacy-preserving analytics forIoT and cloud-basedhealthcaresystems.*IEEEInternet Computing*,*22*(2),42-51.

17. Vaidya,J.,&Clifton,C.(2003,August).Privacy-preservingk-meansclusteringoververtICallypartitioneddata.In *ProceedingsoftheninthACMSIGKDDinternationalconferenceon Knowledgediscoveryand data mining*(pp. 206-215).

18. Vaidya, J., & Clifton, C. (2004, April). Privacy preserving naive bays classifier for verticallypartitioneddata.In *Proceedingsofthe2004SIAMinternationalconferenceondatamining*(pp. 522-526).SocietyforIndustrialandApplied Mathematics.

19. Verykios, V. S., Bertino, E., Fovino, I. N., Provenza, L. P., Saygin, Y., & Theodoridis, Y.(2004). State-of-the-art in privacy preserving data mining. ACM Sigmod Record, 33(1), 50-57.

20. Yu, F., & Ji, Z. (2014). Scalable privacy-preserving data sharing methodology for genome-wideassociationstudies:anapplicationtoiDASHhealthcareprivacyprotectionchallenge.*BMC medicalinformatics anddecision making*, *14*(1), 1-8.

21. Zhang, C., Zhu, L., Xu, C., & Lu, R. (2018). PPDP: An efficient and privacy-preservingdisease prediction scheme in cloud-based e-Healthcare system. *Future Generation ComputerSystems*,*79*, 16-25.

22. Zhou, J., Cao, Z., Dong, X., & Lin, X. (2015). PPDM: A privacy-preserving protocol forcloud-assistede-healthcaresystems. *IEEEJournalofSelectedTopicsinSignalProcessing*,*9*(7), 1332-1344.