CryptoSecurity: Applying Homomorphic Security Schemes to Encrypted Data in Cloud Computing

Ms. Deepika Bhatia[1], Dr. Meenu Dave[2]
{ deepika.bhatia@vips.edu, meenu.s.dave@gmail.com }

[1,2] Department of Computer Science and Engineering, Jagannath Univeristy Jaipur , India

Abstract. Cloud computing is a platform which can be used for data storage and processing. Users can outsource their data to cloud environment where various operations can be performed on this data. As the data is private, different concerns are there for security and privacy of user's data over the distributed network. Nowadays, data security is a major concern in higher education. Various techniques are available for mobile devices and used in IOT. Major issue is of data communication arises while using such kind of devices. So the concern arises over the security of data during transfer using public environment. Size and power level available in these small handheld devices is also a major challenge. Mobile environment also creates a big concern over the security of user's data. The paper focuses on various challenges, issues and various technologies available to secure user's data. This paper presents a comparative analysis of various homomorphic encryption techniques. The paper presents the analysis of various homomorphic encryption algorithms such as RSA, ElGamal, Paillier, enhanced homomorphic encryption scheme, AHEE, BGV etc. Comparative study of these public key cryptographic techniques is presented in this paper and also it is shown that ECC method is extensively used nowadays by various Governments as it is more secured while exchanging the information in cloud environment. It can also be improved in future scenarios.

Keywords: Cloud, Homomorphic, Security, Encryption, Asymmetric, Elliptic Curve Cryptography etc.

ICIDSSD 2020, February 27-28, New Delhi, India

DOI 10.4108/eai.27-2-2020.2303142

1 Introduction

In current scenario, various electronic crimes are expanding day by day. Data kept in public domain is not secure. As we know that private domain can be very costly, so users find public domain more conducive. Security becomes the preliminary concern for such kind of data. Cloud computing is an interesting ware-house to store and administer the data. Different organisations store their data on the cloud. But the main concern is about the privacy of user's data. The data outsourced by various companies; is very sensitive and needs to be secured from various attackers. Various researchers worked in this direction and provided solutions related to security issues. For providing privacy to user's data on the cloud environment; different encryption techniques are available. These methods also support queries done on the encrypted data. It also helps to hide the relative knowledgeable information or data; from the cloud service providers. Various network based tools and applications are used to reacquire data from cloud. Security concerns [16] are also there in the network which is being used for data transfer by various users. So, unauthorized access is also a concern here. Data surveillance goals are divided into three main categories such as in-corruption which means integrity of data, confidentiality and availability.

1. Data integrity techniques - It shows the righteousness of data. Data should not be altered by illegitimate users. There are various techniques which offer data integrity. RAID method provides redundancy of data. Digital signature technique confirms integrity as if the hash value is changed by the illegal user, the output of verification technique will not be a proper match.

2. Confidentiality techniques - These methods deals with protecting user's data from unauthorized and illegal access. For example:-Homomorphic encryption, RSA, 3DES, Random number generator methods provide data security. Different authentication techniques include schemes such as user validation that can be done using passwords, and public key infrastructure. Other methods include Kerberos.

3. Data availability - It states that data should be accessible and available to users. Attacks such as denial of service create such problems.

In this paper, we mainly focused on Asymmetric public key encryption techniques related to data confidentiality. Homomorphic Encryption, Asymmetric Cryptographic Technique, is a technique in which computation is done over the cloud on already encrypted data and when decryption is done, the user will get the result which could have been actually obtained from performing the same operations on the original data. Figure 1 below shows various homomorphic encryption techniques:-

With the help of homomorphic encryption technique, the hacker will be unable to get relevant information from user's data. For example:- assume that user wants to send his data for storage to cloud provider. User encrypts his data and sends it. Now later on, the user may need any computation to be performed on that data. Cloud service provider performs those operations and returns the data to the user. Now finally user decrypts that result and gets the desired output which he might have been obtained from applying computations on original unencrypted data. Homomorphic encryption techniques can be classified as either semi-

homomorphic or fully homomorphic or somewhat homomorphic. Partial HE schemes uses either additive/multiplicative operation on encrypted data. Somewhat HE scheme uses limited number of these operations. And fully HE scheme uses arbitrary number of multiplicative/additive operations on user's data.

2 Literature Review

In 1978, Rivest, Shamir and Alderman [1] proposed RSA algorithm. The main objective of the technique was to derive semantic security. In this method, the message is padded with random bits before encryption. The main application of the RSA is in internet banking, credit card system and internet security. This method is based upon the multiplicative property of Homomorphic Encryption. There was a drawback of this algorithm. The padding with random bits led to loss in the multiplicative property exhibited by RSA encryption algorithm.

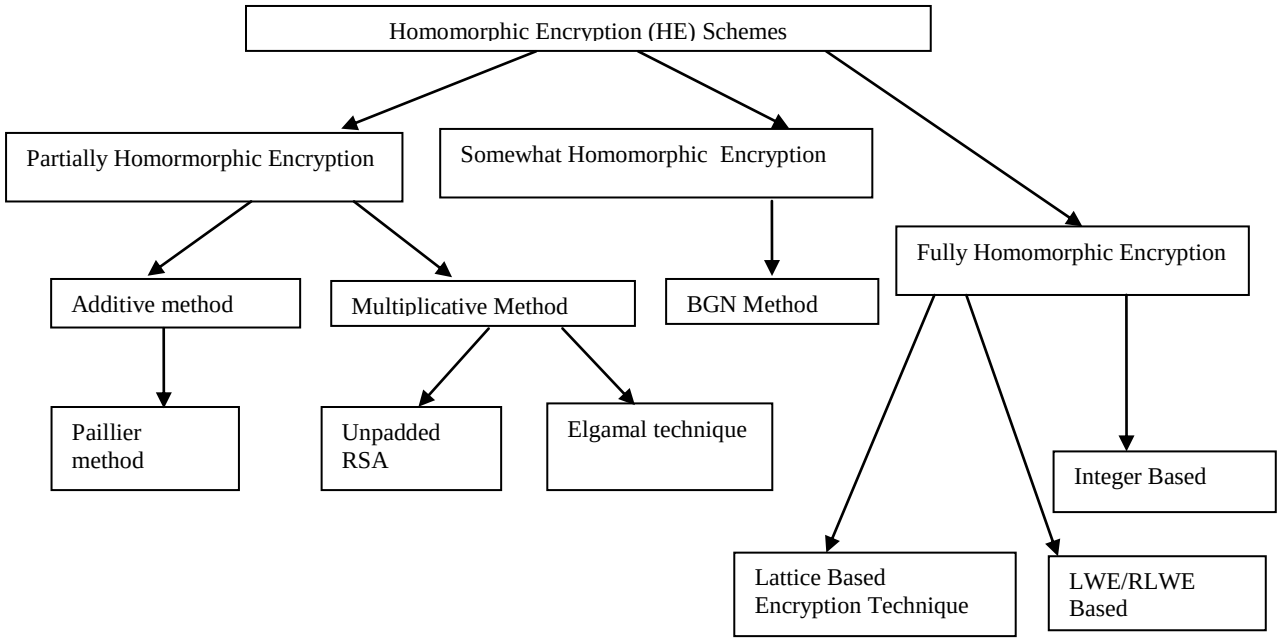

Fig. 1. Classification of various Homomorphic Schemes

Taher ElGamal [2], in 1985, advised a new scheme for data encryption. This asymmetric method is based upon Diffie-Hellman Key Encryption technique and uses public key cryptography. As shown in the table1, this algorithm is better than RSA as it is more secured, uses lesser key length and also avoids incidental attacks on user's private data.

In 1999, Pascal Paillier proposed asymmetric [29] method for data encryption. It generated so many cipher text from one plain text. So data security got improved. Mainly in e-voting system this technique was deployed.

In 2009, Craig Gentry [3] proposed a Fully Homomorphic encryption technique. As the message sent by user is first encrypted and then various operations are done on user's data. After each function, there is a certain noise that gets added to encrypted data. Now this data when decrypted will give unwanted results. So to overcome this problem, Gentry proposed a scheme called Bootstrapping that removes this noise. In this method, he decrypted the cipher-text partially and got the desired results. He succeeded in converting somewhat homomorphic scheme into fully homomorphic scheme as now many number of operations can be performed on the user's data. These operations are of multiplicative/additive types.

Van Dijk, M., Gentry, C., Halevi, S. and Vaikuntanathan [5] proposed fully homomorphic encryption over the integers. The author proposed DGHV method. But it failed to retrieve cipher text from plaintext. In the same year, 2010, Gentry [6] proposed FHE using unlimited additions and multiplications. It is based upon ideal lattices. The scheme was named as GEN10. DGHV method is less secure as it has to send the private key to the server.GEN10 was more secured than DGHV scheme as it doesn't takes the help of private key 'p' but takes a random number 'q' instead of that. This scheme also failed to retrieve cipher-text. So in 2012, J. Li et. al. [4] proposed SDC Scheme which is much more secured than DGHV and GEN10. The author proposed cipher-text retrieval algorithm. This method was able to read cipher-text from plain-text.

In 2012, Guangli [7] proposed Algebra Homomorphic Encryption scheme based on updated Elgamal (AHEE). This scheme provided better protection of user's data. The author focused on secured multiparty computation. AHEE was based upon mixed homomorphic encryption technique. Its real time application can be seen in electronic voting system. AHEE algorithm has a property that when two different encryptions are performed on the same plaintext, the results are different. So security is enhanced in this case. But this method was partially secured. And also doesn't prove to be efficient.

Brakerski et. al. [30] proposed BGV scheme which encrypted bits. This scheme basically deals with lattices. Integer vectors and integer polynomials are dealt with various versions of BGV scheme. This scheme is asymmetric in nature. This algorithm is very good as its various parameters are compared in the following table. It is a follows mixed homomorphic property. In this noise level is very much reduced by using the key and modulus switching methods. Table I below presents the comparative analysis of various Homomorphic Encryption Techniques:-

Table I. Comparative Analysis of Homomorphic Encryption Techniques

Scheme and year	**Security**	**Contribution/Gap**	**Performance**
RSA[1] (1978)	Semantic	Larger key length (1024-4096)	Speed is fast. But Data loss is there.
ElGamal[2] (1985)	Secured and based on DHKE	Smaller key length than RSA's.	Better than RSA as avoids trivial attack

Paillier[29] (1999)	A Semantically secured method unless it hold DCRA property	Security and confidentiality can be improved in future.	Only applies additive property of Homomorphic encryption. It can be improved.
Gentry[3] (2009)	More secured based on ideal lattices.	Public key size can be reduced in future. Implementation is complex and not realistic.	Improved as noise reduced
DGHV[5] (2010)	Secured	Public key size is smaller than Gentry but still key failed to retrieve cipher text from plaintext.	Simpler than Gentry 2009 scheme.
GEN10[6] (2010)	Secured	This scheme also failed to retrieve cipher text from plaintext.	Better than Gen09.
SDC[4] (2012)	More secured than DGHV and GEN10	Could retrieve cipher text from plaintext.	SDC is better in performance as compared to GEN10 and DGHV
Guangli method[7] (2012)	Secured than Elgamal	Partly secured and less effecient.	Partly secured method.
ECC-Elliptic Curve Cryptography[16] (2016)	Secured Technique.	This technique is difficult to implement. It leads to severe security issue if not implemented in a proper way.	Storage and bandwidth got saved. Also Data security issues can be resolved by reviewing the security code, static code analysis and penetration testing.

In 2013, Graepel et. al. [8] , used training algorithm for machine learning which can be presented as low degree polynomials. The predicted results will be sent to the user and user can now decrypt it and read it. This led to the infusion of machine learning in homomorphic encryption technique. So security got improved. But drawback of this system was accuracy. The result predictions were error prone. It showed bad results in application areas like speech or image recognition. Table II below shows the properties and applications of various Homomorphic schemes (HE) :-

Table II. Properties of various Homomorphic Cryptographic Schemes

Serial no.	**Algorithm/ Scheme**	**Additive HE**	**Multiplicative HE**	**Mixed-HE**
1	Paillier	Yes	No	No
2	RSA	No	Yes	No
3	Elgamal	No	Yes	No
4	BGV	No	No	Yes
5	EHC	No	No	Yes

6	AHEE	No	No	Yes
7	ECC	No	No	Yes

In 2014, Yan Zhang et el. [9] proposed an image regeneration method which is very secured. This method used Paillier scheme and thus its properties to implement the secured homomorphic encryption. The technique was based on the factor that to enhance image security, its attributes and content are both encrypted. Care should be taken while decryption so that features and content should remain as it is. In this paper, three important features such as colour, texture and shape are extracted from the image for matching process.

N. Emmadi [10] in 2015, studied the complexities of various algorithms like Bubble sort, Insertion, Bitonic, and odd-even merge sort. The author observed that when data is encrypted by fully homomorphic encryption algorithm, the sorting shows worst case complexity depending upon the order of sorting. He also contributed the theory that odd-even Merge sort after some optimization outperforms the other three algorithms. The running time of this algorithm is less as compared to other sorting methods. Table III below shows various complexities of Sorting algorithms [10] and their complexities in plain and encrypted domains.

Table III: Complexities of various Sorting algorithms

Sr. No.	**Method/technique**	**Plain Text (best case)**	**Encrypted Text (any case)**
1	Bubble sorting	O(n)	O(n2)
2	Insertion sort	O(n)	O(n2)
3	Quick sort	O(n logn)	O(n2)
4	Merge sort	O(n logn)	O(n2)
5	Bitonic sort	O(nlog2n)	O(nlog2n)
6	Odd-even merge sort	O(nlog2n)	O(nlog2n)

In 2015 Zhihua Xia et al. [11] suggested an image retrieval method. The user extracts the local features from various images and outsources the image repository to the cloud. The cloud based image retrieval system, of the cloud tries to find similar images from the database without revealing or having any information about user's data. So the data is secured. The main disadvantage of this scheme was to extract proper features from image database and to send them to cloud. If this service is also performed by CSP, the burden will be relieved from the data owner and user in future.

In 2016, Ayantika Chatterjee et. al. [12], proposed that various data structure operations such as sorting, can be done on encrypted data. The author suggested Lazy Sort algorithm with lesser number of re-encrypting operations. These re-encrypting operations are very costly. So performance got improved relative to normal comparison sorting methods. In future, cost of

re-encrypting operations can be further reduced. Also hardware and scheduling costs can be taken care of.

During 2016 Nathan Dowlin et. al., suggested a new approach where data can be encrypted and sent to the CSP for homomorphic encryption. Some predictions, if required, can be done on data using neural network [13]. Thus user information is secured and information and results are not shared in original form. The user has to decrypt the result for further analysis. The problem with this approach was that Homomorphic Encryption made the process much slower. Also the parameters chosen were larger in number that made the process slower. These problems can be improved in future work.

In 2017, Keke Gai1et. al. [14], suggested a Fully Homomorphic encryption techniques base on real numbers. This was further advancement over the previous techniques which utilized integers for computations. It was named as FHE with RN. The method used KP that is Kronecker Products encryption scheme which used both additive and multiplicative properties of Homomorphic encryption techniques.

In 2018, Debasis Das [15] , used padding techniques into a hybrid encryption algorithm based on RSA. The author combined two schemes, Optimal Asymmetric Encryption Padding (OTAEP) together with Hybrid Encryption algorithm that is based on RSA Smaller and Efficient RSA (HE-RSA). This method helped to preserve the confidentiality of user's data and also focused on multi-party computations over this encrypted data.

Ming-quan et. al. [17], in 2016, proposed Elliptic Curve Cryptography (ECC) algorithm. ECC is a public key cryptography approach and is based upon the algebraic structure of elliptic curves over finite fields. Its main base was homomorphic encryption. By this operations and communication cost related to secure multiparty computation gets very much reduced. Also, it is very secured, uses less energy and communication cost is low. The author compared various schemes such as RSA, Paillier and ECC. The performance is better in ECC as compared to RSA and Paillier.

Anggriane et. al. [18], in 2016, implemented improved version of e-voting system based upon Paillier homomorphic encryption algorithm. This is an efficient method for calculating voting given by general public. Any information from a particular vote can't be extracted as data is in encrypted form. The calculations are done on this encrypted version of user's vote and result is sent for decryption by cloud service provider.

Ximeng Liu et. al. [19] in 2015, presented a new approach based upon Paillier homomorphic encryption technique clubbed with a threshold decryption scheme. This scheme was named as POCR, "privacy preserving outsourced calculation for real numbers". Various operations such as division, sorting, multiplication, GCD and equivalence testing were performed on integer data. After that, the scheme applied same functions on rational numbers.

In practical real world scenario, it can be seen that Fully homomorphic encryption techniques are very complex to implement [21] and also data security still remains the biggest concern [22]. So partially HE schemes are considered as good approaches to implement privacy preserving concept. Various authors proposed protocols such as secure comparison protocols [23], secure scalar product protocols [24], secure set intersection protocols [25],

secure vector comparison protocol [26], and secure TOP-K protocols [20][27]. Different applications were proposed by different authors. Li et al. [28] proposed a protocol that is helpful to match a profile amongst a group of people based upon some attributes. The protocol used was named as set intersection protocol. Also Liu et al. [27] created a TOP-K method that can identify user's top-k disease based upon user's symptoms. All is done in a secure privacy preserving manner.

3 Conclusion

The study of various homomorphic encryption schemes like Algebra homomorphic encryption scheme based on ElgGamal i.e. updated ElGamal, called as AHEE, DGHV and GEN10, RSA, Paillier, FHE based upon Neural Network, MPC with Homomorphic Encryption, HE- with real numbers(HE-RN), machine learning with HE, Optimal Asymmetric Encryption Padding (OTAEP), is presented in this paper. This paper also focused on some major issues related FHE such as intense amount of computational power is required. It is observed that BGV and AHEE schemes are better than other techniques in case of privacy preserving. FHE is costly and also takes intense processing time and implementation complexity. In future, we can propose to implement FHE in real life applications with better performance such as reduced processing time, better security, arbitrary numbers of additive and subtractive operations on data and also improved throughput.

4 Future Work

In future, fully homomorphic encryption techniques with arbitrary number of operations can be implemented based upon various performance parameters such as cost, time, key size, throughput and data security etc. It is observed that elliptic curve cryptography (ECC) technique is popular nowadays. The survey of various schemes presented above shows that ECC is asymmetric fully homomorphic cryptographic technique which is more secured, authentic, uses short key size, and also takes less computing overhead. We can implement ECC scheme in real world applications as it is more secured and efficient amongst various cryptographic schemes.

Acknowledgement

I would like to thank Dr. M. Dave, who guided me to conduct this comparative analysis. This work was supported by my institution, Vivekananda Institute of Professional Studies, New Delhi. I am thankful to the organization for providing me the laboratory set up for performing my experimentation for successful completion of my work. Also I would like to thank the anonymous reviewers for their valuable and helpful comments.

References

[1] Somov, R.L. Rivest, A. Shamir, and L.M. Adleman,"A Method for obtaining Digital Signature and Public- Key Cryptosystems," Comm. Assoc. Computing Machinery, vol. 21, no. 2, 1978, pp. 120–126.

[2] T. ElGamal, "A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms," IEEE Trans. Information Theory, vol. 31, no. 4, 1985, pp. 469–472

[3] C. Gentry, "Fully Homomorphic Encryption Using Ideal Lattices," Proc. 41st Ann. ACM Symp. Theory of Computing (STOC 09), 2009, pp. 169–178.

[4] J. Li, D. Song, S. Chen, X. Lu, "A Simple Fully Homomorphic Encryption Scheme Available in Cloud Computing", In Proceeding of IEEE, 2012, pp. 214-217.

[5] V. Dijk, M, Gentry, C. Halevi, S. Vaikuntanathan, "Fully homomorphic encryption over the integers."LNCS, vol. 6110,2010, pp. 24-43.

[6] C. Gentry," Computing arbitrary functions of encrypted data." Communications of The ACM, 53(3), 2010, pp. 97-105.

[7] X. Guangli, Y. Benzhi ,"An Algorithm of Fully Homomorphic Encryption", 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), IEEE, 2012, 978-1-4673-0024.

[8] Graepel, Thore, Lauter, Kristin, and Naehrig, Michael., "Ml confidential: Machine learning on encrypted data". Springer, In Information Security and Cryptology–ICISC 2012, 2013, pp. 1–21.

[9] Y. Zhang, Li Zhou, Y. Peng, J. Zhang, "A secure Image Retrieval Method Based on Homomorphic Encryption for Cloud Computing", in proceedings of the 19th International Conference on Digital Signal Processing, IEEE, 2014, pp. 1-11.

[10] N. Emmadi, P. Gauravaram, H. Narumanchi, and H. Syed, "Updates on sorting of fully homomorphic encrypted data," IACR Cryptology, IEEE , 2015, pp. 19-24.

[11] Z. Xia, Y. Zhu, X. Sun, Z. Qin, and K. Ren,"Towards Privacy-preserving Content-based Image Retrieval in Cloud Computing", IEEE transactions on computer computing, September 2015 , pp. 1-11

[12] A. Chatterjee and l. Sengupta , "Sorting of Fully Homomorphic Encrypted Cloud Data: Can Partitioning be effective? ", IEEE 2016, 1939-1374, pp. 1-14.

[13] N. Dowlin, R. Gilad, B. Kristin, L. Michael, N. J. Wernsing, "CryptoNets: Applying Neural Networks to Encrypted Data with High Throughput and Accuracy Proceedings of the 33 rd International Conference on Machine Learning, New York, NY, USA,JMLR: W&CP volume 48, 2016, pp. 1-12.

[14] K. Gai1, M. Qiu, Y. Li, X. Y. Liu ,"Advanced Fully Homomorphic Encryption Scheme Over Real Numbers" , IEEE, 4th International Conference on Cyber Security and Cloud Computing, 2017, pp. 64-69.

[15] D. Das,"Secure Cloud Computing Algorithm Using Homomorphic Encryption And Multi-Party Computation ",IEEE, 978-1-5386-2290-2/18, 2018, pp. 391-396.

[16] D. Bhatia, S. Wankhede,"A Study Of Security Issues In Cloud Computing Architecture ",International Journal Of Advanced Research In Datamining And Cloud Computing Issn 2321-8754 ,Volume 3, Issue 5, May 2015, pp. 13-17.

[17] M. Q. Hong,W. B. Zhao, P. Y. Wang,"Homomorphic Encryption Scheme Based on Elliptic Curve Cryptography for Privacy Protection of Cloud Computing", 2016 IEEE 2nd International Conference on Big Data Security on Cloud, IEEE International Conference on High Performance and Smart Computing, IEEE International Conference on Intelligent Data and Security, 2016, pp. 152-157

[18] S. M. Anggriane, S. M. Nasution and F. Azmi,"Advanced e-voting system using Paillier homomorphic encryption algorithm", IEEE, International Conference on Informatics and Computing (ICIC). 2016, 978-1-5090-1648-8/16.

[19] X. Liu,K. K. R. Choo, R. H. Deng, R. Lu and J. Weng, "Efficient and Privacy-Preserving Outsourced Calculation of Rational Numbers" , IEEE, 1545-5971, 2015, pp. 1-14.

[20] B. K. Samanthula, Y. Elmehdwi, and W. Jiang, “k-nearest neighbour classification over semantically secure encrypted relational data,” IEEE Trans. Knowl. Data Eng., vol. 27, no. 5, 2015, pp. 1261–1273.
[21] N. P. Smart and F. Vercauteren, “Fully homomorphic encryption with relatively small key and ciphertext sizes,” in Public Key Cryptography–PKC 2010. Springer, 2010, pp. 420–443.
[22] L. Morris, “Analysis of partially and fully homomorphic encryption”, published in Semantic Scholars, 2013. pp. 1-4.
[23] H. Lin and W. Tzeng, “An efficient solution to the millionaires’ problem based on homomorphic encryption,” in Applied Cryptography and Network Security, Third International Conference, ACNS 2005, New York, NY, USA, June 7-10, 2005, Proceedings, 2005, pp. 456–466.
[24] R. Lu, H. Zhu, X. Liu, J. K. Liu, and J. Shao, “Toward efficient and privacy-preserving computing in big data era,” IEEE Network, vol. 28, no. 4, 2014, pp. 46–50.
[25] F. Kerschbaum, “Outsourced private set intersection using homomorphic encryption,” in 7th ACM Symposium on Information, Computer and Communications Security, ASIACCS ’12, Seoul, Korea, May 2-4, 2012, 2012, pp. 85–86.
[26] X. Liu, R. Lu, J. Ma, L. Chen, and H. Bao, “Efficient and privacy preserving skyline computation framework across domains,” in Future Generation Computer Systems, 2015.
[27] X. Liu, R. Lu, J. Ma, L. Chen, and B. Qin, “Privacy-preserving patient-centric clinical decision support system on naive Bayesian classification,” in IEEE Journal of Biomedical and Health Informatics, 2015.
[28] M. Li, N. Cao, S. Yu, and W. Lou, “Findu: Privacy-preserving personal profile matching in mobile social networks,” in INFOCOM 2011. 30th IEEE International Conference on Computer Communications, 10-15 April 2011, Shanghai, China, 2011, pp. 2435–2443.
[29] Paillier, Pascal. "Public-key cryptosystems based on composite degree residuosity classes." In Advances in cryptology—EUROCRYPT‘99, Springer Berlin Heidelberg, 1999, pp. 223-238.
[30] Z. Brakerski, C. Gentry, V. Vaikuntanathan,” (Leveled) fully homomorphic encryption without bootstrapping.” In: Goldwasser, S. (ed.) ITCS, ACM 2012, pp. 309-325.