2018 IEEE 25th International Conference on High Performance Computing Workshops (HiPCW)

Privacy Preserving in Blockchain based on Partial Homomorphic Encryption System for AI Applications

Sharath Yaji*, Kajal Bangera† and B. Neelima‡,
*† *Dept. of Computer Science and Engineering,*
‡ *Dept. of Information Science and Engineering*
NMAM Institute of Technology, Nitte, India
**sharathyaji@nitte.edu.in*
†bangerakajal12@nitte.edu.in
‡neelimareddy@nitte.edu.in

***Abstract*—The synergy between artificial intelligence and blockchain is increasing in the computing environment. To realize this blockchain technology making its way into applications such as healthcare, financial services, Internet of Things and much more., that use artificial intelligence making it more defendable to attacks. The current blockchain technology uses different encryption algorithms such as SHA256, MD5 etc. The blockchain attacks such as collision attack, primage attack and attacks on wallet motivated us to experiment on partial homomorphic encryption to enhance the strength of blockchain technology. This article considers i) Goldwasser-Micali and ii) Paillier encryption schemes for the comparative evaluation study with a focus on data privacy techniques. We believed and proved that the above two encryption schemes that were considered have less processing time and provide more strength to the possible attacks. While we present our preliminary results in this study, we discuss the pros and cons of the Goldwasser-Micali, Paillier and non-homomorphic encryption schemes that are expected to add value to blockchain technology to be used in Artificial Intelligence (AI) applications.**

***Keywords*-Blockchain, Partial Homomorphic Encryption, Goldwasser-Micali, Paillier Encryption System, Privacy Preserving.**

I. Introduction

A growing linked list of records linked cryptographically forms a blockchain. These digitalized public ledgers are decentralized and widely used by cryptocurrencies, used for both public and private business models. These blocks are set in a chronological order allowing its users to track the digital transactions. There is no central record keeping as the blockchains are decentralized and each of its users saves the copy of the block to itself. Each of these block or a ledger contains various parameters including data, hash, and hash of the previous block. This hash being unique, and changes in the blocks reflect the changes in the hash. It is also secured with a set of cryptographic keys, private and public key when together combined gives a digital signature. Figure 1 represents transactions that consist of the hash, public key, private key, and the digital signature. The time taken to mine a block is referred to as the block time. The expected block time is ideal whereas the average block time varies on the type, content, size, and number of the blocks. Recent studies shown in Proof of Work (PoW) is time-consuming to produce but verification is easy for data miners [2]. Homomorphic encryption is likely to play a vital role in optimizing the creation of hash and time consumption in blockchain technologies [3]. Recent studies show that blockchain wings spread over various industries like advertising, business development, healthcare, logistics etc. extending to all these innovations that are similar to AI and AI itself. High-level transactions such as financial transactions require similar security to preserve the privacy of the user. For many other industries where privacy of data is important, our research will come in handy. Furthermore, the recent progress made in machine learning makes AI a perfect ally for the blockchain to realize more secured and protected applications deployment. Motivated by the privacy and security issues related to the current blockchain, the following objectives are set in for the work in this paper:

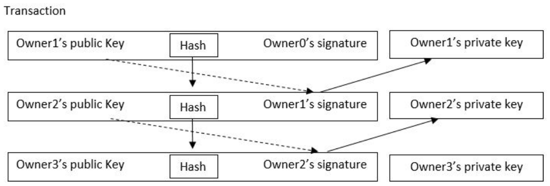

Figure 1. Existing blockchain approach

- To introduce the homomorphic and non-homomorphic encryption schemes in the era of privacy and security of data concerns.
- To analyze and show the importance of homomorphic encryption schemes that can enlarge the use of blockchain technology.
- To correlate the encryption schemes, with lesser processing time and strong, with the need of blockchain technology to provide security and privacy in the emerging applications.

978-1-7281-0114-9/18/$31.00 ©2018 IEEE
DOI 10.1109/HiPCW.2018.00021

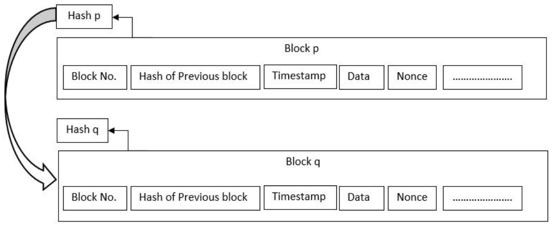

Figure 2. Parameters of a blockchain

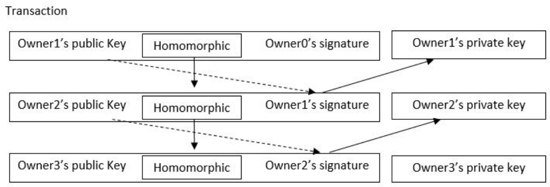

Figure 3. Proposed blockchain approach

Figure 2 shows the existing model of blockchain [4]. This uses hash encryption for hiding privacy of block owners or traders. It is proved that partial homomorphic encryption models are better than fully homomorphic [17]. As shown in figure 3 this paper considers partially homomorphic encryption (additive). Such as i) Goldwasser-Micali Encryption scheme and ii) Paillier Encryption Scheme for the evaluation. Choosing a better encryption scheme to be used in blockchain technology helps provide protection against attacker in getting the identity of the user and the message, and the transaction is done by sender and receiver, provides protection against attacker to obstruct and updated transferred messages or transactions, disables the generation of duplicate public-private key pair by the attacker, prevents the hijacking of transactions and prevents violation of confidentiality of trading parties. The paper summarizes the use of partial homomorphic encryption schemes to preserve privacy in blockchain resulting in better performance. As blockchains are used in artificial intelligence any advances in blockchains reflect on the entire system. The paper is organized as follows. Section II gives the background of blockchain technology and homomorphic encryption system while covering the related work. Section III lists the possible attacks to realize the challenges in visualizing the breakthrough of blockchain technology into artificial intelligence. Section IV gives the mathematical background to support the usage of homomorphic encryption schemes. The experimental set-up and the discussion of preliminary results are carried out in Section V and we conclude in Section VI.

II. Background

In the year 2008, Satoshi Nakamoto introduced digital currency (bitcoin) [4], also known as the cryptocurrency or electronic cash is decentralized, its transactions are recorded in a blockchain which is a public distributed ledger. This technology is valid to any digital asset transaction occurred online and is basically a dispersed database of records that are equal or been executed and shared among the different parties. The entered data can never be removed. The blockchain contains a certain irrefutable record of each and every exchange at any given point. In the advanced online world blockchain sets up an arrangement of making a circulated agreement which is a primary speculation of blockchains [5].

Blockchains provides an excellent backbone for AI algorithms. It reduces the cost of AI by providing security for data input making observation possbible at every step (learning and decision making process) of an AI application. Emerging blockchain reimagines the internet services and access the data of any desired system with a decentralized twist. Obtaining large datasets from and to applications is a challenging task and privacy is the first concern. Using blockchains to retain privacy makes the application simpler and efficient. This paper discusses the enhancement that results in privacy-preserving when homomorphic encryption scheme is implemented in the blockchain, that is in turn used in AI applications.

The fully homomorphic encryption scheme, allows computing arbitrary functions over encrypted data without the decryption key was introduced. Having many applications such as, allowing search on encrypted data and update files only when it is decrypted [3]. The partially homomorphic encryption scheme specifies a type of operation that can be applied to a ciphertext and in turn generates an encrypted result that is later matched with the result of the operation performed on the plaintext. Further, this encryption is subdivided into additive and multiplicative schemes. Further dividing additive scheme into three namely, Goldwasser-Micali Cryptosystem, Paillier cryptosystem, and Okamoto-Uchiyama Cryptosystem [6], [7], [8]. This paper discusses and deals with Goldwasser-Micali Cryptosystem and Paillier cryptosystem.

Fully Homomorphic :. In 1978, Ronald Rivest et al. suggested a proposal on the concept of homomorphic encryption [15]. Later, Craig Gentry proposed fully homomorphic encryption that computes an arbitrary number of additions and multiplications on encrypted data [16]. This scheme enables the programs for any desirable functionality, such as homomorphic property to run on encrypted data and produce an encryption of the result.

Partial Homomorphic : A partially homomorphic exhibits either additive or multiplicative homomorphism, but not both. In addition, the efficiency of some partial homomorphic encryption schemes is high enough for practical applications [17].

Goldwasser-Micali Scheme : Shafi Goldwasser and Silvio Micali mathematically proposed the concept of probabilistic cryptography [10]. The model gave the ability to encrypt the same text in many different ways without changing the modulus, followed by does bit level encryption.

Paillier Scheme : Paillier Scheme is an asymmetric algorithm for public key cryptography, Paillier crypto scheme was proposed by Pascal Paillier in 1999. This probabilistic scheme centered around the property it possesses and allows simple addition computation on several encrypted values and produces the encrypted sum. The encrypted summation can be decrypted without knowing the values that made up the summation.

III. Blockchain Attacks

Modern cryptocurrency infrastructures are centralized and make use of third-party organizations that handles accounts, processes payments and provides security. But these system suffers from scalability and security breach, for eg.

i) In a centralized network, the authetication and payment activities are fully disturbed by data breaching of a centralized middleman.

ii) Centralized middleman may attack daily transactions and also predict the daily activities. The drawbacks of the centralized infrastructure are the motivation behind addressing some of the attacks related to the blockchain. Table I, discuss blockchain attacks list the existing and possible attacks, essentially contributing towards motivation for the work in this paper. These attacks are focused on both messaging and the occurrence of transactions during communications. SHA 256, SHA512 and so on are used to performe the hashing along with RSA to encrypt the block messages.

Table I
Blockchain encryption attacks

SN	Attacks	Description
1	Random number generator	Attackers can exploit vulnerabilities of OpenSSL through the poor random number. This can be utilized to generate public or private key-pair identical to victims to messaging keys, hijack transaction and signatures.
2	Collision attack	If an attacker would be able to generate two different public keys with the same address. Then hash value of plain text can be breached.[18]
3	Preimage attack	If only SHA256(m) is given, where m is the victims seed, the attacker uses a collision search on the elliptic curve secp256k1 points to find a seed message $m^{'}$ that hashes to the same private key SHA256(m) = SHA256($m^{'}$)
4	Sybil attack	Private subnetworks can be built by injecting fake puppet nodes, these nodes detaches victim nodes from the network.
5	Clock drifting attack	The attacker does this attack through drifting clock forward and backward.
6	Stopping particular transactions	The attacker occupies outbound connection slots of a certain user and controls transactions of the victim [19]
7	Race attack	This attack is performed through sending two transactions that conflict, sequentially to the network.
8	51 percent attack	Attacker controls computational power to modify and reverse past transaction. Later, perform double spending attack by generating blocks with malicious transactions.
9	Double-spending attack	This attack is performed by utilizing a single digital token or selling ownership over the same amount of introduced energy twice.
10	Attacks on wallet	Private keys of victim are used to get his money from wallet.

IV. Mathematical Background

This section studies and discusses the mathematical background related to the Goldwasser-Micali Scheme and Paillier Scheme.

A. Goldwasser-Micali Scheme

This encryption is performed through probabilistic algorithm. The cipher text is different after different encryption. Therefore reducing the dictionary text attacks [11].

This encryption method depends on determining whether a given value v is square mod X, given factorization of (p,q) of X, and this can be done as follows:

1) Compute $v_p = vmodp, v_q = vmodq$
2) If $v_p^{\frac{p-1}{2}} \equiv 1modp$ and $v_q^{\frac{q-1}{2}} \equiv 1modq$, then v is a quadratic residue of mod X.

Key Generation: Let p and q be two large prime numbers, these are unique and chosen at random

1) Compute $X = pq$
2) Find nonresidue v such as $\frac{v}{p} = \frac{v}{q} = -1$ and $\frac{v}{X}$ is +1.

The public key consists of (v, X) and the secret key is the factorization of (p,q).

Encryption : Let plaintext p be encrypted as cipher text c.

1) Encode p as a string of bits $(p_1, p_2, ..., p_n)$
2) For every bit p_i, generate random value y_i form the group of units modulo X (i.e $gcd(y_i, X) = 1$). Then cipher text $c_i = y_i^2 v^{p_i} modX$. The cipher text c_i are used as $(c_1, c_2, ..., c_n)$. Where n is size of cipher text.

In this method, only one bit can be encrypted at a time resulting the expansion to be huge. This type of scheme could be usable for bit encryption, however it takes too long for normal size encryption.

Decryption : Decryption process the plaintext p can be recovered as follows:

1) For each i, do factorization (p,q), check if c_i quadratic residue then p_i is 0, otherwise p_i is 1.
2) The decrypted message is the string $p = p_1, p_2, ...p_n$.

To check if c_i was encrypted as quadratic residue modulo X or if it was encrypted as a pseudosquare. Check, if c_i is encrypted as $v^2 mod X$ then it is a quadratic residue mod X, since that is the definition of a quadratic residue. If c_i = yv^2 then we must have a pseudosquare.

Security : This scheme is a probabilistic encryption based, encryptions on the same plaintext, several times will yield different cipher texts. The semantic property of this algorithm, secures under plaintext attack [14]. Goldwasser and Micali, can encrypt more than 1 bit at a time. For any given key k and security parameter N, this method allows the encryption of k bits of data into N + k bit ciphertext. This model gives the protection against considered attacks. The major disadvantage of this algorithm is, for a given security parameter N, the probabilistic encryption of each bit is N bits long, requires N random bits. In this model plaintext will be encrypted as large cipher text, this may reduce the performance to some extent.

B. Paillier scheme

The fundamentals of the Paillier scheme is composite residuosity and the scheme is additive homomorphic cryptosystem. Let $X = pq$ is a composite since X is a composite of pq.

Key Generation :

1) p and q are randomly chosen two large prime numbers, such that $gcd(pq, (p-1)(q-1)) = 1$. This is possible when prime numbers p and q are of equal length (i.e $p, q \in 1 \parallel 0, 1^{s-1}$).
2) Compute the security parameter $X = pq$ and $\lambda = lcm(p-1)(q-1)$
3) Choose integer i randomly, where $i \in \mathbb{Z}^*{}_{n^2}$
4) Ensure n divides the order of i by using $\mu = (L(i^{\lambda} mod n^2))^{-1}\ mod n$, Here L is $L(u) = \frac{u-1}{n}$
5) Public, Private keys are (n,i) and (λ, μ) respectively.

Encryption :

1) Let P be plaintext to be encrypted, here $P \in \mathbb{Z}_n$.
2) Choose random number r such that $r \in \mathbb{Z}^*{}_n$.
3) Compute Cipher text C as $C = i^P . r^n mod n^2$

Decryption :

1) Cipher text $C \in \mathbb{Z}^*{}_{n^2}$
2) Compute plaintext $P = L(C^{\lambda} mod n^2) . \mu mod n$

Both schemes have homomorphic properties. For example, Pillier supports the addition of multiple ciphertexts C, the addition of a plaintext P constant to a ciphertext C [13].

Security : The semantic security of this encryption plot was demonstrated under the decisional composite residuosity assumption and the DCR problem is obstinate. However, X Yi et al. proved that this scheme does not protect against adaptive chosen-ciphertext attacks. Later, Paillier cryptography was updated such that encryption incorporates the combined hashing of the message with a random number. This gave protection against adaptively chosen ciphertext attacks [14].

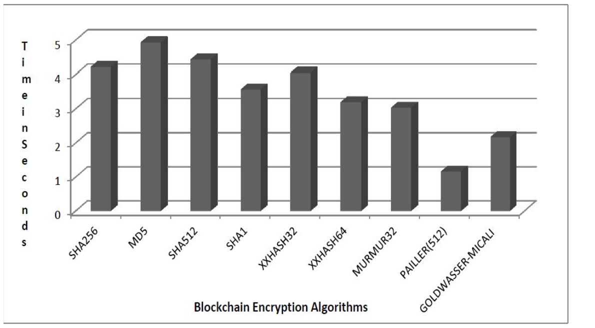

Figure 4. Execution time comparison of different encryption algorithms

Since both the encryption models does not create any hash value, *collision attack* and *preimage attacks* are not possible. In GM, private keys are generated through factorization so its hard to get private keys. This avoids *attacks on wallet*. In the year 2015 Jost et al. proposed improved version of Paillier scheme. The new methods reduce the bottleneck of noise calculation through precomputed noise [21].

V. RESULTS AND OBSERVATIONS

This section highlights the experimental set-up and discusses the observations.

A. Experimental Setup

The experimental set-up includes the hardware and software environment. The implementation used gcc compiler (5.4.0-6). The experiments were conducted in Intel Core i7-2600 CPU @ 3.40 GHz with 64-bit Ubuntu 16.04. It has NVIDIA Titan X GPU card with 4 GB RAM. One block encryption for each algorithm is considered towards the preliminary evaluation.

B. Discussion

We experimented creation of blocks with SHA256, MD5, SHA512, SHA1, XXhash32, XXhash64, murmur32 encryption algorithms, partial homomorphic encryption models, Goldwasser-Micali (GM) Cryptosystem and Paillier cryptosystem [7], [8]. The Figure 4 shows the elapsed time of the experiments. As discussed in the previous sections, the GM and Paillier encryptions schemes are stronger to provide data privacy and these algorithms perform better than others. This gives a chance of considering these algorithms to be considered in blockchain technology especially to be used in AI applications.

The requirements on privacy and security are satisfied as follows by the GM and Paillier encryption scheme as follows:

1) Blockchains encrypted through these (Paillier and GM) algorithms is made using the public key of the receiver and is then sent to everyone. The confidentiality of the receiver is preserved because no one knows to whom the message was sent.

2) Since in blockchain messages and transactions are signed with a private key, here only a receiver would be able to access the actual message. Interruption and alteration would break signatures and integrity of the message.
3) Algorithms such as SHA256 and SHA512 depends on the hash partial preimage resistance property, the concept is similar to broken SHA1. However, both proposed partial homomorphic equations do not use hash functions, this increases resistance against collision and preimage attacks.
4) Paillier and GM are based on asymmetric cryptography concepts. This prevents transaction hijacking by signing transactions.
5) Both can be used for pseudonyms transactions. This will avoid violation of confidentiality of trading agents.

VI. Conclusions

The privacy of a trader in blockchain can be breached through different attacks. For attacks such as collision, preimage and attack on the wallet can be avoided through encrypting block using proposed Goldwasser-Micali and Paillier encryption schemes. With the evolution of the need for security and privacy in applications using artificial intelligence, blockchain technology should evolve by using stronger and cost-effective encryption schemes. The preliminary results from performance evaluation of Goldwasser-Micali, Paillier and non-homomorphic encryption schemes are considered in this paper. We observed that proposed schemes consume less time compared to non-homomorphic encryption schemes, that shows that stronger and cost-effective encryption schemes are possible to be added to blockchain technology to make it more suitable to security and privacy based applications, specifically using artificial intelligence. We further plan to extend this work by considering specific artificial intelligence applications and exprimenting possible attacks on it. These can be further protected by using a modified blockchain technology with better encryption schemes.

References

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.

[2] Gervais, Arthur, et al. *On the security and performance of proof of work blockchains*, Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. ACM, 2016.

[3] Gentry, Craig, and Dan Boneh. *A fully homomorphic encryption scheme*. Vol. 20. No. 09. Stanford: Stanford University, 2009.

[4] Nakamoto, Satoshi. *"Bitcoin: A peer-to-peer electronic cash system"*, 2008.

[5] Crosby, Michael, *Blockchain technology: Beyond bitcoin* Applied Innovation 2, 2016.

[6] Okamoto, Tatsuaki, and Shigenori Uchiyama. *A new public-key cryptosystem as secure as factoring*, International conference on the theory and applications of cryptographic techniques. Springer, Berlin, Heidelberg, 1998.

[7] Paillier, Pascal. *Public-key cryptosystems based on composite degree residuosity classes*, International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Berlin, Heidelberg, 1999.

[8] Shafi, Goldwasser, and Silvio Micali. *Probabilistic encryption*, Journal of computer and system sciences, 1984.

[9] Team, R. Core. *R: A language and environment for statistical computing*, 2013.

[10] Shafi, Goldwasser, and Silvio Micali. *Probabilistic encryption*, Journal of computer and system sciences, 1984.

[11] Bellovin, Steven M., and Michael Merritt. *Encrypted key exchange: Password-based protocols secure against dictionary attacks*, Research in Security and Privacy, Proceedings., 1992 IEEE Computer Society Symposium on. IEEE, 1992.

[12] Paillier, Pascal. *Public-key cryptosystems based on composite degree residuosity classes.* International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Berlin, Heidelberg, 1999.

[13] Richardson, Kert. *Progress on probabilistic encryption schemes* ,2006.

[14] Yi, Xun, Russell Paulet, and Elisa Bertino. *Homomorphic encryption and applications* Heidelberg: Springer, 2014.

[15] Rivest, Ronald L., Adi Shamir, and Leonard Adleman. *A method for obtaining digital signatures and public key cryptosystems*, Communications of the ACM, 1978.

[16] Gentry, Craig, and Dan Boneh. *A fully homomorphic encryption scheme*. Stanford: Stanford University, 2009.

[17] Bao, Fen. *Cryptanalysis of a provable secure additive and multiplicative privacy homomorphism*, proceedings of the international workshop on coding and cryptography. 2003.

[18] Sasaki, Yu, Lei Wang, and Kazumaro Aoki. *Preimage Attacks on 41-Step SHA-256 and 46-Step SHA-512*, IACR Cryptology ePrint Archive, 2009.

[19] Piasecki, P. *Design and security analysis of Bitcoin infrastructure using application deployed on Google Apps Engine* Wydzia Fizyki Technicznej, Informatyki i Matematyki Stosowanej. University of Warsaw, 2012.

[20] Shruthi, R., P. Sumana, and A. K. Koundinya. *Performance Analysis of Goldwasser-Micali Cryptosystem* International Journal of Advanced Research in Computer and Communication Engineering, 2013.

[21] Jost, Christine, et al. *Encryption Performance Improvements of the Paillier Cryptosystem*, IACR Cryptology ePrint Archive, 2015.