*Grenze International Journal of Engineering and Technology, June Issue*

# Improving Data Integrity and Security in Cloud Environments: A Blockchain-based Method

Mahima John[1] and Dr. Abhijeet R Raipurkar[2]
[1]Student of Master of Technology, Department of Computer Science & Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur, India
Email: johnm@rknec.edu
[2]Assistant Professor, Department of Computer Science & Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur, India
Email: raipurkarar@rknec.edu

***Abstract*—Despite the years' rapid advancements, valid computing and data security continue to be ongoing challenges. dynamic cloud infrastructure, surpassing manual or reactive approaches. Unfortunately, many obstacles, including overly complex computations and issues with scalability, hinder the efficiency of many suggested fixes. This article has discussed the addition of block chain approaches as a possible solution to these problems. A promising way to improve cloud computing's processing speed and storage security is through the decentralized distributed computing structure known as block chain. The study presents the use of remote agent technology to deploy a distributed virtual machine (VMA) model over the cloud. The VM agent mechanism effectively handles tasks pertaining to monitoring, verification, and dependable data storage, providing a crucial basis for creating the reliability and safeguards for block chain technology. In order to reduce the computational burden and quantum computing during the user verification phase, this work presents the lattice signature algorithm, which contains a cuckoo filter. It is suggested that the decentralized ledger network take the role of conventional centralized audits in order to improve security and transparency.**

***Index Terms*— Block chain, Sampling, Data Integrity Verification, lattice signature, Cloud Storage.**

## I. INTRODUCTION

Block chain technology appears to be a key component of the answer to problems [1]. Block chain is presently leading high-value areas and offering a wide range of applications due to its unique technological advantages and creative value ideas [2, 3]. Block chain is seen by many experts as having the ability to spark the next big Internet technology revolution [4].

Block chain solves the authenticity problem by acting as a publicly available ledger that guarantees enduring, tamper-proof data with time stamps [5]. Block chain technology helps to partially address the trust issue plaguing the modern economy by decentralizing the conventional Internet.

Transfers carried out on the ledger remove the need for parties to build relationships of trust; instead, confidence in the distributed ledger itself is sufficient.

Additionally, block chain addresses issues with data validity by promoting effective data convergence, sharing,

*Grenze ID: 01.GIJET.10.2.658*

and circulation [6]. It unlocks the door to the "the worth Internet," where the efficiency of value transfer corresponds with the exponential growth in information sharing facilitated by the Internet. The importance of block chain in creating a trustworthy credit society is emphasized in the study.

Big data and block chain integration provide answers to current issues in data research and development [7-9]. Presently available centralized solutions entail dependable third-party entities; nonetheless, questions are raised about their reliability and vulnerability to malevolent actions. The study promotes the use of block chain technology in conjunction with big data to offer decentralized solutions that lessen the problems associated with centralization [10].

The article notes that Bit coin is credited with the development and broad adoption of block chain technology, which can be traced back to a Satoshi Takemoto's paper [11]. Block chain manages a distributed database that includes operations, barriers, smart contracts, and consensus methods. Secure, auditable records are guaranteed by its decentralized, clear, and tamper-proof nature, which is confirmed by all nodes that are involved.

According to the World Economic Forum, block chain technology is receiving international attention; by 2017, 80% of banks will be working on block chain projects, and more than 20 countries will have invested in this field of study [11]. Three phases are proposed in the paper for the actual implementation of the technology known as block chain [12]. Phase 1.0 concentrates on enabling digital currency; Phase 2.0 incorporates smart contract technology into financial procedures; and Phase 3.0 expands the use of block chain to traditional IT industries, diverse social domains, and the creation of potentially transformative life and social norms [13–20].

## II. LITERATURE STUDY

The ongoing development of the Internet has given rise to a number of innovative and rapidly developing fields, such as cloud computing, 4G/5G, IoT (Internet of Things), and software-defined networking, or SDN, [7]. Numerous linked sectors have initiated major research endeavours in these areas. Within the broad domains of cloud computing, scholars have explored subjects including data integrity validation in addition to data and picture retrieval studies [8–9]. This section aims to provide a thorough review of pertinent research in block chain, cloud data integrity verification, and lattice signatures—all crucial topics covered in this study. The wide range of research and developments in these fields make a substantial contribution to the field's overall technical innovation.

### *A. Verification of Cloud Data Integrity*

Researchers are paying more attention to remote data integrity checks as cloud storage grows in popularity. Users must confirm the availability and integrity of data in the cloud during this procedure. Numerous solutions have been presented to mitigate the security issues associated with cloud storage. These solutions fall into two categories: Proofs of Retrievability (POR) and Provable Data Possession (PDP) schemes. PDP systems use a "challenge-response" process to confirm the accuracy of data storage, whereas POR schemes can retrieve lost or damaged data.

Ateniese et al. [3] presented the PDP system, which takes into account both trusted third-party audit and public audit utilizing RSA algorithms for data integrity verification. .. The e-CSP returned file blocks label data minimizes computational and transmission overhead in this method. It does not, however, permit updates to dynamic data. Ateniese et al. [11] have since presented an extensible PDP providing dynamic auditing, but with restrictions on the amount of verifications and data updates.

Erway et al. [12] presented a structural data on skip-lists for a PDP system based on hierarchical authentication in their study. While this technique allows for full data updates, data privacy is not adequately protected. In order to ensure user privacy, Wang et al. [13] optimized this approach using random masks; nevertheless, data block placement support was lost. Vulnerability problems with forgery attacks inside this technique were noted by Zhou et al. [14]. Zhu et al. [16] have developed a data structure based on the index hashing table (IHT) that reduces the overhead of storing verification information and facilitates operations Yang et al. [17] for the preservation of data privacy, whereas Ni et al. [18] revealed its weakness.

Xu et al. [19] improved the accuracy of data verification results by introducing a fraud-resistant algorithm utilizing dual evidence modes. Nonetheless, Shen et al. [20] proposed a method for ensuring data integrity while resolving compute and storage overhead concerns, permitting safe file sharing without jeopardizing privacy. Validation efficiency in a verifiable data integrity technique was enhanced by Li et al. [21].

Zhu et al.'s [22] integrity verification system for the Internet of Things, or IoT, proved the efficacy and security of the ZSS signature, a brief signature approach. Even if it's taken for granted that TPA believes in open audit verification techniques, concerns regarding TPA's reliability and internal operations arise. Methods have been proposed to counteract malicious TPAs.

Huang et al. [23] implemented a method that improved audit authorization by utilizing numerous Third-Party Auditors (TPAs) and adding receiving & time servers to improve audit timeliness and accuracy. Wu et al. [24] separated TPA verification into two categories: user-handled and TPA-managed, respectively, sophisticated computations and easy verification methods. Xiao et al. [25] have used trusted hardware in their work to secure servers that store cloud data for audit logs; nevertheless, more testing is required to confirm the security and reliability of this technique. These solutions take care of any problems with TPA fidelity while fulfilling the criteria of the cloud user audit.

### *B. Work Based on Lattice-Signatures.*

Significant progress in lattice cryptographic systems within the subject of cryptography theory has led to the appearance of notable research results in recent times. Researchers have been aiming to produce short signatures and reduce the amount of verification keys at the same time, addressing a major issue with lattice signature systems. (2010) has made a revolutionary contribution. Boyen presented the first scheme of short-signature inside the consensus paradigm, which evolved from lattice (provable) [26]. Peikert and Micciancio [27] subsequently improved Boyen's approach by strengthening the hash function, which increased overall safeguarding.

There have been improvements in validation signing and key size in later digital signature techniques with proved security. Yang and colleagues have developed a signature scheme that is derived from identity by using the distinct algebraic structure of the ideal lattice. The objective of this approach is to enhance productivity and optimize the signature procedure.

The anonymity for a single identity was not achieved by this approach, while unforgeability under fixed selecting messages and selected identity was satisfied. Later, Zhang and colleagues developed a lattice signature technique for verification keys protecting O (logn) matrices in the context of the standard model. Thus, the general security is strengthened.

M rental prices, makespan indicates total execution time, and security is quantified based on attack probability and damage probability of violating constraints.

Tang et al. [8]: Tang et al. propose a discrete virus optimization (DVO) algorithm for multi-objective workflow scheduling in the cloud optimizing cost, makespan, and availability. DVO mimics virus propagation and convergence toward optimum antibodies.

Powell et al. [9]: Powell et al. proposed a two-step method separating the feasible configurations generation and optimal configurations selection processes. Feasible configurations satisfying hard and soft constraints are generated through an equivalent transformation-based constraint satisfaction approach. The paper claims this method can efficiently deal with complex constraints compared to conventional techniques. Pareto-optimal selections are then made from the feasible configurations using multiobjective evolutionary algorithms.

### *C. Blockchain Application*

With its open, transparent, and traceable characteristics, blockchain technology has become an indispensable instrument in a number of industries, including Internet of Things (IoT), healthcare, and finance. The implementation of cryptographic algorithms fosters confidence between users and the Cloud Service Providers, hence reducing the potential dangers associated with malicious third parties. In order to preserve confidence in cloud ecosystems, this guarantees the objectivity and integrity of data. Blockchain offers a strong basis for establishing trust through consensus procedures and encryption methods.

This study proposes a novel approach that uses blockchain to overcome the trust concerns related to TPAs: assurance and blockchain in cloud data integrity verification.

## III. RESEARCH METHOD

### *A . Blockchain*

Cryptographic algorithms are used to construct data blocks in blockchain, sometimes referred to as block chain, which is an open, decentralized database system that links data across time. With the use of decentralized design concepts, encryption algorithms, time stamping strategies, consensus processes, and incentive structures, it provides an entirely novel model in distributed computing that operates without requiring confidence in any one node.

The blockchain is made up of series of chronological blocks, each of which contains a record of all the legitimate transactions that have occurred inside the network. Usually, the block-header and block-body are located within each data block. A block-body includes the transaction counter and detailed transaction data.

The main characteristics of blockchain are: anonymity, which is attained by user addresses concealing true identity information; auditability, which is the ability to easily track and verify transactions using the UTO model in Bitcoin; decentralization, where individuals maintain consistent information without depending on a centralized mechanism; persistence, which makes transactions very difficult to modify, roll back, or delete once added to blockchain.

*B. Lattice*

Def. 1:

Let B= (b_1, b_m )⊂R^n signify a collection of linearly independent vectors. The lattice Λ created by B is Λ= {BC=∑_ (i=1) ^M ⟦b_i c_i |c_i∈Z, 1≤m≤n┤}. The basis length, referred to as ‖B‖, is then defined by the longest vector length inside the set of basis B, B is termed basic of a lattice Λ. The longest vectors length inside a set B is the basis length, represented by the symbol ‖B‖. The standard orthogonal basis is the set B= (b_1, b_n) that is produced by the conventional Schmidt orthogonalization procedure.

.Def. 2:

Since u∈Z_q^n, A∈Z_q^ (n×m), and q is a prime number, we may define the following: Λ_q^⊥ (A) {e∈Z^m s.t.Ae=0 (mod q)}, and Λ_q^u (A) {e∈Z^m s.t.Ae=u (mod q)}. Assume that m=5nlogq and that q ≥ 3. In that scenario, an algorithm with probability polynomial time exists. TrapGen (q, n) yields a resultant matrix T that is in T_q^ (m×m) and an A that is in Z_q^ (n×m). As a consequence, uniform A is Z_q^ (n×m), which is obtained from the lattice Λ_q^⊥ (A) be T and fulfills T̃≤O (√nlogq) and T≤O (n log) with some small probability.

*C. Technology Pertaining to Security in Cloud Data Storage*

In the realm of cloud computing, data storage is both a central component and a fundamental component. As such, when it comes to cloud computing security protection, data security comes first. Data information is an essential part of an organization's assets for both individual consumers and businesses. Cloud data storage security is a major problem. A typical security paradigm for storage services in the cloud is shown in Figure 5. User data security standards cover a range of hostile actions that might endanger cloud servers. These include anti-eavesdropping, anti-tamper-proofing, anti-discard, and anti-abuse procedures. Align with the CIA's core values of availability, integrity, and secrecy. Important technologies that might endanger cloud servers include control over access, data encryption, data recovery, and data integrity verification. These technologies are essential to guaranteeing the security of data in the cloud storing environment.

*Technique for verification correctness*

a1 Verification procedure in a user accuracy

The user participates in the key verification stage after obtaining credentials from a storage provider. In order to evaluate the formula for verification and establish the data integrity, the user uses their set of parameters and public keys in this step.

The user does independent computations for both sides of the verification equation when any data remains intact, indicating unaltered parameters. To finish the validation procedure, the equation on the left is assessed. $e(\sigma, g) = e\left(\prod_{i=1}^{r_c} \sigma_i^{\omega_i}, g\right) = e\left(\prod_{i=1}^{r_c} (h(f_i) \cdot u^{f_i})^{k \cdot \omega_i}, g\right)$

$$= e\left(\prod_{i=1}^{r_c} (h \cdot (f_i) \cdot u^{f_i})^{\omega_i}, g\right)$$

$$= e\left(\prod_{i=1}^{r_c} h \cdot (f_i)^{\omega_i} \cdot \prod_{i=1}^{r_c} u^{f_i \omega_i}, g\right)$$

$$= e\left(\prod_{i=1}^{r_c} h \cdot (f_i)^{\omega_i} \cdot u^{\sum_{i=1}^{r_c} f_i \omega_i}, v\right)$$

It becomes clear that the formulas on both sides are equivalent. Interestingly, since the constants u and v are available to the general public, differences between the two solutions to equation (3) only occur when these quantities change.

$$proof = \left\{\mu, \sigma, \left(h(f_i)\right)_{1 \le i \le r_c}\right\}$$

The following are the specific actions that have been delineated for this enhancement:
*Initialization:*
During the first stage, a user creates random public-private key pairings. For each user, n∈ (1, N) indicates. In the CA system, submit a certificate request to get a set of verified private and public keys. Later, a random number (ssk, spk)
selection of select$X_n \leftarrow Z_p$, & evaluate$v_n = g^{x_n}$, $n$ user will be using $V_n$ as a public key, at time$V_n \in Z_p$.

*Labelling of Data:*
Each user formulates equation (4) by dividing the transmission data into k pieces. A user computing signatures in the next formula (5) chooses u_n↞G.

$$F_n = f_{n,1} \| f_{n,2} \| \dots \| f_{n,k} \quad \dots . (4)$$

$$\sigma_{n,j} \leftarrow \left[ h(f_{n,j}) \cdot u_n^{f_{n,j}} \right]^{x_n} \quad \dots (5)$$

*Challenges:*
From the pairings of {1, 2, k}, the user selected subgroups in random R= {i_1, i_2, i_c}, and, supposing that components in R have been organized in ascending order, a subset of c components. The server query is started by the verifier, as shown in equation (6).

$$chal = \{(i, V_i)\}_{i_1 \le i \le i_c} \quad \dots (6)$$

*Key generation:*

The input received by server, as illustrated in Equation 6, and processes it for every user.$n \in \{1, \dots, N\}$, Calculating based on Equation (7) and Equation (8), the server processes the data, and in response, provides feedback to the verifier according to Equation (9).

$$\tau = \sum_{\{i, v_i\} i_1 \le i \le i_c} v_i \cdot f_{n,j} \in Z_p \quad \dots (7)$$

$$\sigma = \prod_{n-1}^{N} \left( \prod_{\{(i,v_i)\}_{i_j \le i \le i_c}} \sigma_{n,j}^{v_i} \right) \quad \dots (8)$$

$$= \prod_{n-1}^{N} \left( \prod_{\{(i,v_i)\}_{i_1 \le i \le i_c}} \left[ h(f_{n,j}) \cdot u_k^{f_{n \cdot j}} \right]^{x_n v_i} \right) \quad \dots (9)$$

$$\left\{ \sigma, \{\tau\} 1 \le n \le N, \left\{ h(f_{n,j}) \right\} \right\} \quad \dots (10)$$

*i. Key validation:*

After obtaining the keys, each user uses their public key to confirm that Equation (10) has been established. If the equation is true, it means that the data's integrity has not been compromised.

$$e(\sigma, g) = \prod_{n=1}^{N} e\left( \prod_{\{(i,v_i)\}_{r_1 \le i \le r_c}} h(m_{n,i})^{v_i} \cdot u_n^{\tau_n}, v_n \right) \quad \dots (10)$$

The precision of data integrity verification is equivalent to every individual user.
The following will happen if a tenant supplies k query attributes:

$$C_s^{m-MHT} =$$

After the selection process, the flow reaches another decision point to ascertain if a node has been chosen. If no node is selected, the search concludes. However, if a node is indeed chosen, the process continues the search from the selected node, and the cycle repeats, starting again from the decision on the proportion of nodes to develop. This iterative process allows for a systematic exploration and refinement of nodes within the search algorithm.

## IV. Results and Discussion

The setup for these experiments involves a virtual environment for cloud storage services, which consists of five storage node servers and one cloud storage management server. Within this configuration, one of the nodes acts as the data integrity maintenance, while the remaining nodes acts as active nodes for cloud storage services. These storage nodes are responsible for storing data and facilitating data transfer during the process of data

integrity protection.Utilizing the Ethereum Wallet, the Meta Mask extension, and the Interplanetary File System (IPFS), we have created a distributed block-chain-based cloud platform.

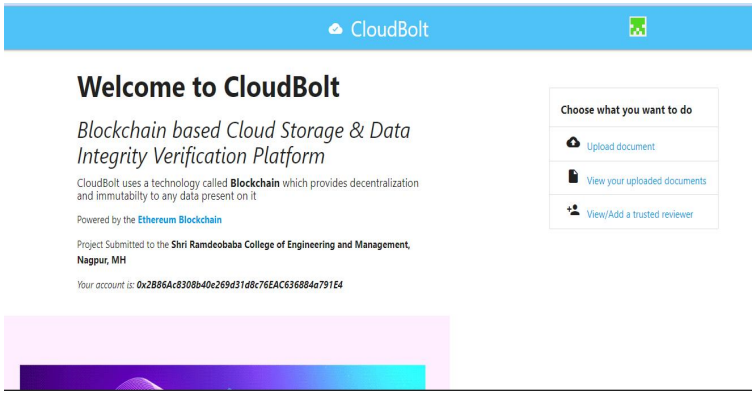


Figure 1. View of the Cloud Bolt Cloud and Data Integrity Platform

A little every piece of hash codes linked to a secure account, IPFS retains the data in units called Peers, guaranteeing data integrity via the use of a signature scheme

The structure of the cloud platform intended for data integrity testing is shown in Figure 1. The platform, called Cloud Bolt, is made using Python utilizing Flask, Web3, Ethereum Wallet, and ReactJS and NodeJS servers. The IPFS server may be accessed more easily with this configuration.

As seen in Figure 2, comparisons were made across the three distinct aspects of our test: the cost of creating signatures, proofs, and verifications. The results show that our method generates signatures with less computational cost than Zhu et al.'s ZSS signature and Wang et al.'s BLS signature.

By using bilinear mapping, both the BLS and ZSS signatures include a large number of exponential operation that incur high computing costs. Comparatively, our method's primary computational overhead is primarily attributable to RST (the rejection sampling Even with ZSS signature improvements, the greater efficiency of our system is constantly maintained.

We have developed the differential analysis to evaluate the efficiency of the signature algorithm, comparing our approach to the ZSS by Zhu et al. [22] and the BLS by Wang etal. [13]. During the verification phase, we assess and replicate the effectiveness of the three provided strategy.

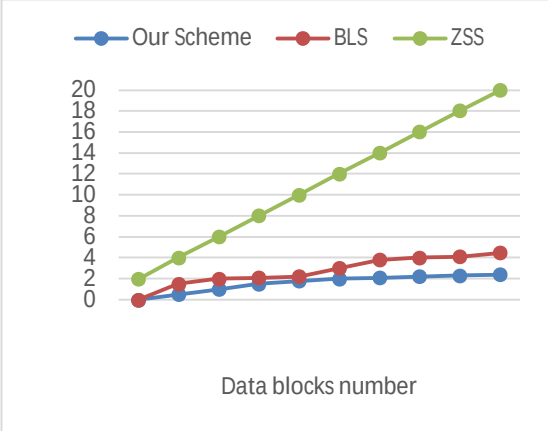


Figure 2. Signature creation difference

The comparison's proof generation time findings are shown in Figure 3, which displays a linear connection with many difficult blocks. In this phase, our technique is unique in that it achieves better performance with a limited data volume by needing just multiplication, as well hash, and modulus operations. Similar to the signature creation procedure, the BLS signature involves bilinear, hash, and exponential mapping operations. In contrast, the ZSS signature shows more computing overhead during the proof generation step because of a greater quantity of exponential computations, even if it avoids them.

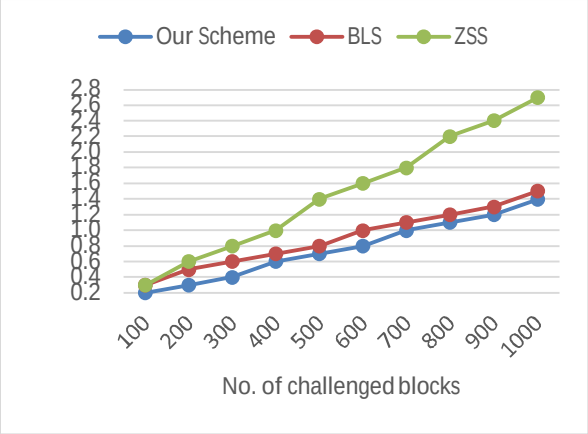


Figure 3. Comparative Analysis to verify generation time

## V. Conclusions

Ensuring the integrity of data stored in cloud settings has become crucial due to the increasing use of cloud storage and the ongoing growth of data inside these environments. This study uses block chain technology to present a reliable and effective method for ensuring the integrity of cloud data. Our methodology aims to overcome the constraints seen in conventional centralized audits by utilizing the block chain network's capabilities, thereby improving efficiency and security. Furthermore, our verification system is developed to withstand potential hazards from quantum cloud computing, based on the SIS issue assumption. By solving issues with user processing capacity, the integration of cuckoo filter and lattice signature technologies greatly streamlines the user's verification procedure. The outcomes of the evaluation highlight the verifiable effectiveness of our suggested plan. The goal of future study should be to achieve a more smooth integration of systems for integrity verification.

## Acknowledgment

We would like to express our sincere gratitude to the distinguished professors in the Department of Computer Science and Engineering at Ramdeobaba College of Engineering and Management, which is associated with Rashtrasant Tukdoji Maharaj Nagpur University (RTMNU). Their unwavering support was instrumental in making the research project a success. Throughout the course of my study on Improving Data Security and Integrity in Cloud Environments: A Blockchain based Method, I am grateful for their constant support and encouragement.

## References

[1] C. Peikert, “Trapdoors for lattices: simpler, tighter, faster, smaller,” Advances in Cryptology-EUROCRYPT 2012, vol. 7237, pp. 700–718, 2012.
[2] H. Tian, F. Zhang, and B. Wei, “A lattice-based partially blind signature,” Security and Communication Networks, vol. 9, no. 12, pp. 1820–1828, 2016.
[3] Y.-L. Gao, X.-B. Chen, Y.-L. Chen, Y. Sun, X.-X. Niu, and Y.-X. Yang, “A secure crypto currency scheme based on post quantum block chain,” IEEE Access, vol. 6, pp. 27205–27213,
[4] Y. Wu, X. Lin, X. Lu, J. Su., and P. Chen, “A secure light-weight public auditing scheme in cloud computing with potentially malicious third-party auditor,” IEICE Transactions on Information and Systems, vol. E99.D, no. 10, pp. 2638–2642, 2016.
[5] D. Xiao, L. Yang, B. Sun, and S. Zheng, “Provable data possession system for realistic cloud storage environments,” Journal of Software, vol. 27, no. 9, pp. 2400–2413, 2016.
[6] X. Boyen, “Lattice mixing and vanishing trapdoors: a framework for fully secure short signatures and more,” Public Key Cryptography-PKC 2010, vol. 6056, pp. 499–517, 2010.
[7] D. Micciancio and C. Peikert, “Trapdoors for lattices: simpler, tighter, faster, smaller,” Advances in Cryptology-EUROCRYPT 2012, vol. 7237, pp. 700–718, 20
[8] S. Nepal, S. Chen, J. Yao, and Dilakanathan, “Dias: Data Integrity as a Service in the Cloud,” in Proceedings of the 2011 IEEE 4th International Conference on Cloud Computing, pp. 308–315, IEEE, Washington, DC, USA, 2011.

[9] T. Ristenpart, E. Tromer, H. Shacham, and S. Savage, “Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds,” in Proceedings of the 16th ACM Conference on Computer and Communications Security, pp. 199–212, ACM, New York, NY, USA, 2009.
[10] Yan S and Zheng,” A user centric trust and reputation method for service selection” in proceedings of 2010 international symposium on intelligence information processing and trusted computing, IEEE 2010, 101-105.