Big Data Security Issues and Quantum Cryptography for Cloud Computing

Vijey Thayananthan
Computer Science Department,
Faculty of Computing and Information Technology,
King Abdul Aziz University,
Jeddah 21589, Saudi Arabia

Aiiad Albeshri
Computer Science Department,
Faculty of Computing and Information Technology,
King Abdul Aziz University,
Jeddah 21589, Saudi Arabia

ABSTRACT

Enhancement of security in cloud computing is challengeable when big data such as medical and government confidential information involves in a cloud environment. The managing key for individual user participated in the mobile cloud will be the problem when big data is handled without efficient security key management. Especially, mobile cloud nodes where users handle their data can be anywhere else in the cloud environment. So, managing keys for random users and finding correct locations of users' nodes and their identities are interesting areas for the research. To solve this problem related to key management, cloud services need efficient key management that employs quantum cryptography using Grover's algorithm. Moreover, in this research, we introduce effective authentication technique in order to improve the level of security (with minimum complexity) In this paper, theoretical model, Grover's algorithm and quantum cryptography with the PairHand protocol are considered as methodologies. As summarized results, security levels of users and services providers are considered to verify the services used in a cloud environment. In spite of same quantum properties characterize the light, data security issues are still not matured with the light technology. Hiding the big data in the light through this research methodology will be the conclusion of the future security.

General Terms

In this paper, we consider quantum cryptography (QC) as a general term. Throughout this research, software PairHand protocol is considered to improve the cloud computing.

Keywords

Big data security; PairHand protocol; Key Management; Quantum cryptography; Cloud services

1. INTRODUCTION

Big data security depends on many threats they are such as interception of data transmissions, data leakages during the uploading and downloading, etc. Credentials are vulnerable to many threats to big data security. Password-based authentication and storing cloud access credentials in random unsecured locations are some important examples.

There are many solutions for which security issues are being processed in cloud environments which include hybrid cloud computing [1] and cloud infrastructure management using adaptive resources [2]. Providing a high level of security for big data is a challenging development because owners' identity, the location of the port and server, and properties of big data should be addressed dynamically. Establishing security for an individual user is another challenge. To develop these challenges, key management (KM) in the cloud environment will be one of the choices. Efficient KM in random locations where big data being approached to the cloud storage is considered. Even strong security is established by the users, data storage and transmission in cloud environments cannot be controlled by the users without cloud service providers. This challenge is also one of the interesting topics. According to [3], size of the big data has been doubling every 2 years since 2011. Very soon (maybe by 2020), the size of big data will be 2.5 zettabytes (2.5 x 10^{21} bytes) of information could be encrypted using quantum cryptography (QC).

In this research, QC and KM with Grover's algorithm (GA) are considered in the designed theoretical model. Reading big data in transit and poor authentication used during the communication encryptions are also affecting the big data security in cloud environments. Regarding the big data security, poor KM provides vulnerabilities in cloud computing. Following points are used to improve the KMs. Firstly, hardware security modules can be located randomly in cloud environments. Secondly, KM interfaces are kept available and accessible.

The rest of the paper is organized as follows. Section 2 focus on literature review which explains the big data security based on QC, service provider's priority and PairHand technique. In section 3, we focus on the proposed theoretical model that uses the QC. Here, Grover's algorithms and PairHand protocol are also considered. Section 4 provides the necessary results dealing with security levels of service provider and users. Key lengths details of big data security are also given. Section 5 deals with security analysis of big data used in cloud computing. In section 6, overall conclusions are provided based on the results and the theoretical analysis.

2. LITERATURE REVIEW

According to [10, 21], applications of big data and revolutionary breakthroughs with current technology are considered. Despite many applications, big data security issues provide the challenging solutions to improve the cloud environments. Here, large individual organizations try to maximize their security when they employ the cloud computing. QC that provides the efficient security using quantum properties and principles has been developed for many latest applications. It provides not only the maximum security but also it improves the complex issues which depend on the number of steps. Despite many efficient search algorithms, GA used in this research reduces the key search operations. In this research, KM involved with GA controls the key exchange protocols employed between the cloud user and server [6-7, 11].

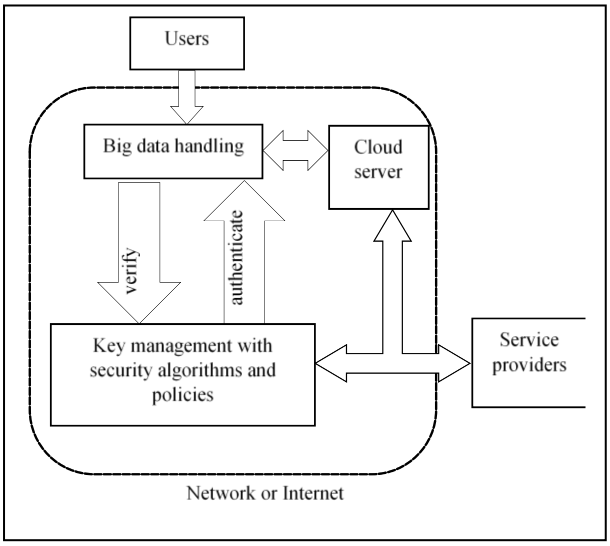

Figure 1: Basic cloud computing for big data security

As shown in figure 1, the user can send the big data through the cloud system that has basic procedures. Firstly, big data handling module distinguishes the type of the data which are size, sensitivity, etc. It can also send them to cloud server or KM module. Here, big data get the authentication dynamically and take the next procedure that may be either storage or up and downloading. Secondly, KM provides necessary security according to the service providers' instructions. Security during the transmission before big data reach the storage locations is not considered in this research.

2.1 Big data security based on QC

Big data security based on QC increases the protection for improving cloud computing. Although QC has many benefits and advantages such as less complexity, big data needs more attention to improve the security levels through the security algorithms. According to [16], a symmetric key with the block cipher provides the better security levels of the cloud computing. Further, efficient design of the block cipher allows us to control the big data security. Despite the efficient design, GA provides the better solutions to minimize the complexity depended on the larger size of block ciphers and the steps for key searching [16-17]. Through the GA, big data security issues based on QC can be improved with maximum protections and other benefits such as storage capacity.

2.2 Service providers' priorities in a cloud environment

According to [20], following priorities are unavoidable when big data security are considered in a cloud environment. Firstly, the service provider should ensure that confidentiality, integrity, and availability are manageable with or without any conditions in anywhere around the cloud environments. Secondly, according to the service level agreements of resources used in the cloud computing, the service provider should check the features and functionalities of the cloud resources. Service providers offer all latest versions of the tools and resources to protect their services. Capabilities of self-service and achievable scaling are also provided through the automation and simplifications. Authentication is also one of the priorities.

2.3 PairHand authentication

According to [4-5], authentication procedures take many (more than 2) handshakes which increase not only the time but also complexity during the overall computations. Despite many authentication procedures, handover authentication procedure helps us to analyze the big data security in mobile and cloud computing environment. According to recent articles [15], PairHand protocol involves four phases. Initially, the mobile user $\mathrm{U_i}$ generates the signature S_i as in (1). In which, H_1 and H_2 are both hash functions, and k is used to create the private key.

$$S_i = H_2(U_i)kH_1(uid_i) \tag{1}$$

Where $U_i = uid_i \| ID_{APy} \| ts$, uid_i is authentication server chose pseudo-ID (AS) located in the cloud environment, ID_{APy} is the identification of the *APy* and the timestamp is *ts*. Then the generated signature is used for the access request message in which the $\mathrm{U_i}$ create a one-to-one connection to *APy*. Then, generate the shared symmetric key as in (2).

$$K_{i-y} = \hat{e}(kH_1(uid_i), H_1(ID_{APy})) \tag{2}$$

In the second phase, *APy* validates both signature $\mathrm{S_i}$ and timestamp *ts*. If succeed, *APy* generates value as in (3), arranges the code of authentication then forward all parameters (uid_i , ID_{APy} , Auth) to the i-th user $\mathrm{U_i}$.

$$K_{y-i} = \hat{e}(H_1(uid_i), kH_1(ID_{APy})) \tag{3}$$

Thirdly, $\mathrm{U_i}$ guarantee the identity of the connection by creating the verification code (as in (4)) which then compared with the authorization code sent by *APy*.

$$Ver = H_2(K_{i-y} \| uid_i \| ID_{APy}) \tag{4}$$

The user $\mathrm{U_i}$ compares ($Ver \overset{?}{=} Aut$), if they are equal, $\mathrm{U_i}$ generates a secure channel with authentication between the authentication server and each node of the cloud.

Finally, access point APy creates the connection that allows the users to establish the secure transmission. Although many users send the message with their signatures, the individual user will be able to transfer the message to AS in the cloud. Latency increases when re-authentication takes place during the handover procedure. Although proposed algorithm reduces the latency, most of the security solutions which help us to analyze the big data security in mobile and cloud computing environment.

3. PROPOSED THEORETICAL MODEL

As shown in Figure 2, we have developed the proposed model theoretically. All procedures in each layer and authentication key used in this model help us to establish the secure link between the two nodes of the cloud environment. Each node in cloud computing can be either a user or receiver, but when this node acts a sender and receiver, mobile cloud activities will have to face the handover problems. To improve these situations, mobile user and AS, in theory, can be developed with appropriate KM that provides the necessary security to big data. In this theoretical model, an authentication protocol is

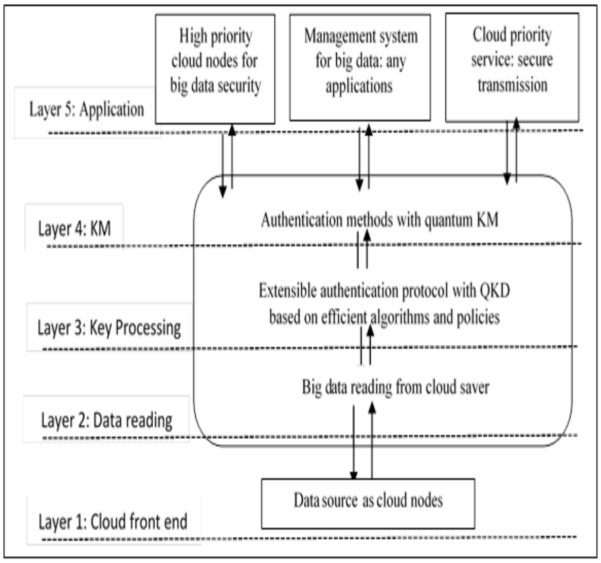
High priority cloud nodes for big data security
Management system for big data: any applications
Cloud priority service: secure transmission
Layer 5: Application
Layer 4: KM
Authentication methods with quantum KM
Extensible authentication protocol with QKD based on efficient algorithms and policies
Layer 3: Key Processing
Big data reading from cloud saver
Layer 2: Data reading
Data source as cloud nodes
Layer 1: Cloud front end

Figure 2: Proposed theoretical model for cloud computing

As shown in Figure 2, we have explained the layers used in this theoretical model.

Layer 1: Any used considered as a node in the cloud environment can start the big data transmission procedures. According to the properties of the big data, the data source of this model starts the reading procedures. Based on the above properties, the characteristics of the big data such as sensitivity are important to identify the data. Next layer will take over after the identification processed from the data source.

Layer 2: In this layer, we have to prepare security procedure because wireless channel expects to have evolving threats.

Layer 3: As we explained in Section 2, all operations and steps used in the authentication algorithms provide the security levels according to the nature of big data. Here, key generation is one of the operations depends on the policies of the authentication protocols and quantum key distribution (QKD). Also, GA minimizes the key searching procedures which influence with QC, size of the big data.

Layer 4: Authentication methods that may be either existing or proposed authentication protocol can be employed [9]. Here, KM looks after all the necessary security operations according to the properties of the big data. Although KM has many issues, KM issues the security key to the individual user. Here, KM protocols finalize the security levels of each service.

Layer 5: Despite many applications, managing correct security level of big data rely on the services used in the specific applications. Big data that are either sensitivity or confidential need to be secured and monitored to control the security levels of each service used by the organization.

3.1 Big data security in cloud computing

Using this theoretical model, security issues are addressed to improve the security levels while service providers and users are actively handling the required services. Security levels are used to maintain the services efficiently because each service may be interrupted by many reasons such as delay, traffic, etc. In this research, time complexity that influences with the key searching based on the proposed research is considered instead of delay and traffic. Security levels have measured using above-mentioned influences which provides considerable changes in each service used in cloud computing.

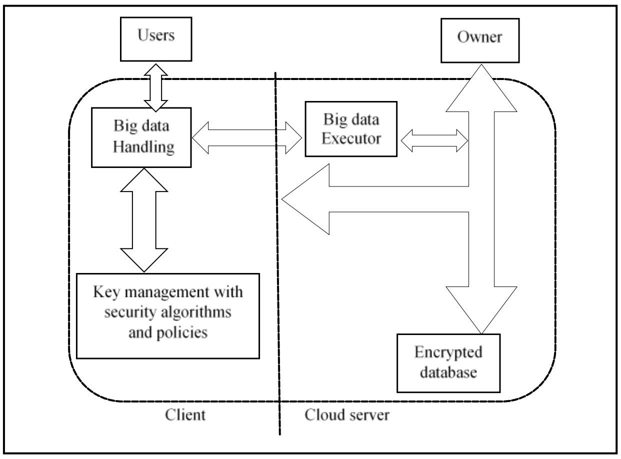

Figure 3: Security schemes of cloud system for big data

As shown in figure 3, a block of owner’s big data is encrypted from the cloud server, and it is sent to the client of the cloud system through the big data executer. In the client module, the user will be able to handle the big data for further processing through the big data handler where user retrieve either the encrypted or decrypted version of the big data.

Responsibilities of all resources used in cloud computing depend on security levels of applications. Table 1 summarizes the responsibilities of the service providers and users when security issues are considered for the big data in cloud computing.

Table 1. Responsibilities during the big data security

Service providers	**Service users**
Platform security: Message queues, storages	**Information security:** Encryption
Infrastructure security: Firewall, security monitoring, Network-level, OS-level, host-level	**User security & monitoring:** Identity and supporting services
Management .level security: All including security key	**Application level security:** Service connections, storage

Service providers protect their platforms using firewalls and security monitoring. Looking after the platforms including the network architecture is very expensive protection. Even though service users have fewer responsibilities, user, information and application level securities are very important. They not only protect their own data but also they look after the services influenced during the cloud computing [12-14].

3.2 Time complexity and big data security

The efficiency of the security in which speed of the processing depends on the time. When we have quick processing, protection will be stronger. Here time complexity plays an important role to improve the security level in the cloud computing. Further, time complexity depends on the number of nodes and active resources located in cloud computing. Although we use time complexity as an example, space and other complexities are involved in the design.

Despite many design factors, minimizing time complexity issues increases the efficiency of the big data security levels [22-24]. So, we assume that randomized big data security services of cloud computing design M have time complexity $T_c^M(n)$, and S is the summation of n data ($S = \sum^n$).

$$\mathrm{T}_\mathrm{c}^\mathrm{M}(n) = \max_{z\in S}\left(\mathrm{T}_\mathrm{c}^\mathrm{M}(z)\right) \tag{5}$$

Using (5), we can calculate $T_c^M(n)$ which, is the maximal expected running time of M overall big data. Here, the length n is obtained from the fragmented size of the big data.

When the secret key is established during the attack, following equation can be used to reduce the steps. Also, it reduces the complexity when big data security is considered with appropriate algorithm [8].

$$\text{complexity} = O\left(\sqrt{2^n}\right) \tag{6}$$

In the conventional KM scheme, many search operations depend on (6). Generally, many passive attacks, require more steps as in $\text{complexity} = O(2^n)$. Assume that key size is n in above cases. According to the theory, big data should be fragmented into a reasonable size which is suitable for the cloud system.

Definition of availability

Cloud services should be able to use whenever and wherever cloud users want to manage their big data. So, service providers should ensure that all of their offered cloud services are available during the time users need. Here, time complexity influences with the service time which is dependent on the up or downtime of the cloud services.

$$A = \frac{t_{cs} - t}{t_{cs}} \tag{7}$$

According to (7), availability (A) of the services used in cloud computing can be calculated. In this (7), t_{cs} is the cloud service time and t is the downtime of the service. Security levels depend on the CIA which is confidentiality, integrity, and availability. In the results, availability is not shown because it is a common mechanism for all the services. Instead, authentications are used because it is more important than the availability and it provides the necessary security to service users.

4. RESULTS

Security levels of each service used in cloud computing are shown in following figures.

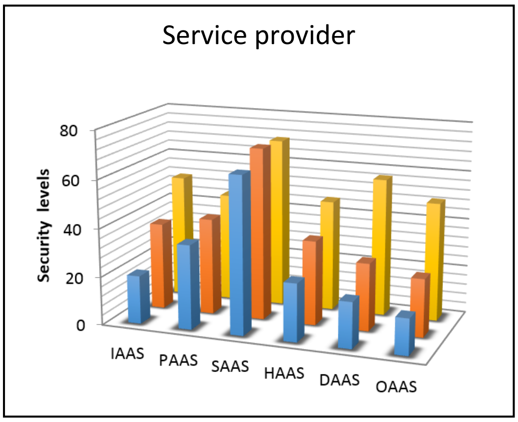

Figure 4: Cloud security management of the existing approach

Through the appropriate KM, protocols used existing approach, cloud service provider as shown in Figure 4 measures their services for big data security issues. When a big data block is in transit, cloud computing facilities and responsibilities must depend on the following services as defined in the previous cloud computing research studies. They are Infrastructure as a service (IaaS), Platform as a service (PaaS), Software as a service(SaaS), Hybrid as a service(HaaS), Data as a service (DaaS) and Other as a service (OaaS) respectively.

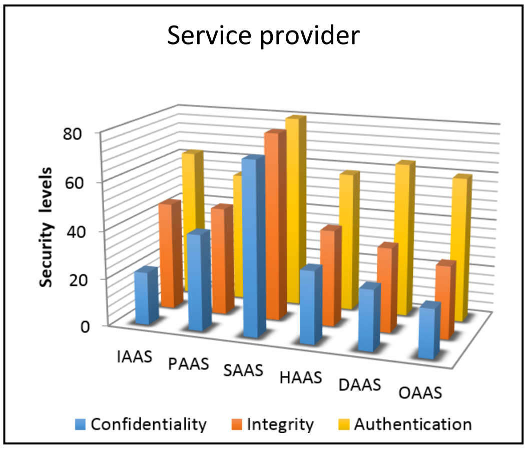

Figure 5: Cloud security management of the proposed approach

Through the appropriate KM protocols used in proposed approach, cloud service providers as shown in Figure 5 improve their services for big data security issues. In the proposed approach, authentication is the main concern because PairHand authentication protocol is used with QC which increases the speed of the key search. According to the result, the security level is not only increased in the authentication but also confidentiality and integrity are improved.

Here, users' responsibilities of how much security levels achieved during the services are shown in following figures.

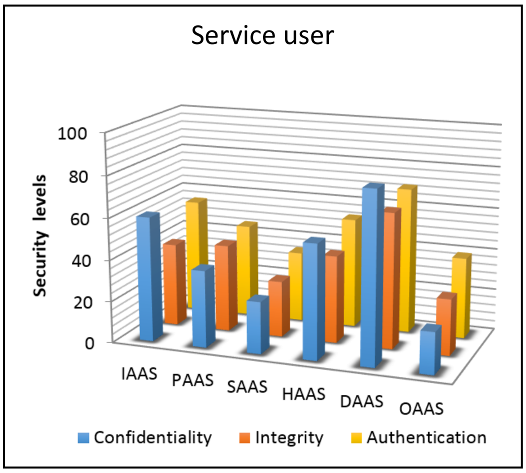

Figure 6: Cloud user security management of the existing approach

As shown in Figure 6, users' maximum security levels of services are measured when they handle the big data from their location which is a random point in the cloud environment.

Users access the cloud computing from anywhere they want, but services should be available at any time that may not be possible because security levels of services mentioned above depend on the availabilities. According to Figure 7, proposed approach reduces the complexity. Therefore availabilities of all services can be controlled through the security levels. In spite of strong KM controls the proposed approach, security levels of available services are still depends on the service providers. Cloud computing always provides necessary access to both service providers and users. Although the controlling responsibilities of the services and security levels are manageable by both parties, owner of the big data is still daunting about the overall cloud security that depends on other factors.

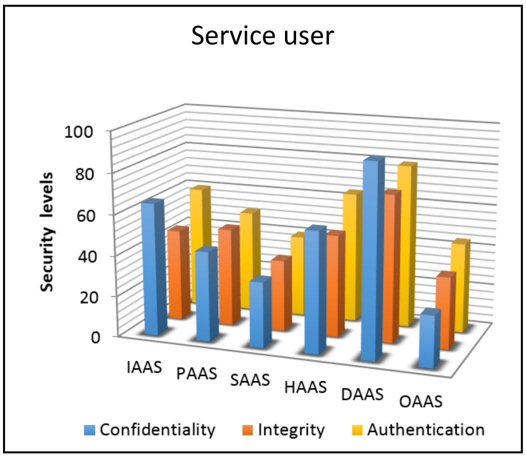

Figure 7: Cloud user security management of the proposed approach

4.1 Analysis of service providers and users

Comparing the results of existing and proposed approaches, responsibilities during the big data security provide the necessary enhancement to improve the cloud services with maximum safety. With the stronger confidentiality, integrity and authentication mechanism, service providers manage the security levels in all recommended services used in the cloud computing. In both service provider and user analyses, we have used our proposed model to improve the security levels of the authentication. In the analysis of service provider, overall all services used in the cloud computing are improved by 10%. Also in the service user, security levels of all services show some improvements. Despite the flexible management responsibilities, key management is very important. In this case, we have analyzed the key processing with our proposed model.

4.2 Complexity analysis of key generation

In this analysis, complexity is reduced because our proposed model uses GA, QC and PairHand protocol which provide lower complexity than the existing system. According to the key processing mentioned in the proposed model, when we increase the number of qbits (quantum bits) the number of operation is reduced. Therefore, algorithms we used in this research and analysis not only reduce the complexity but also improve the security levels in each service considered in the cloud computing.

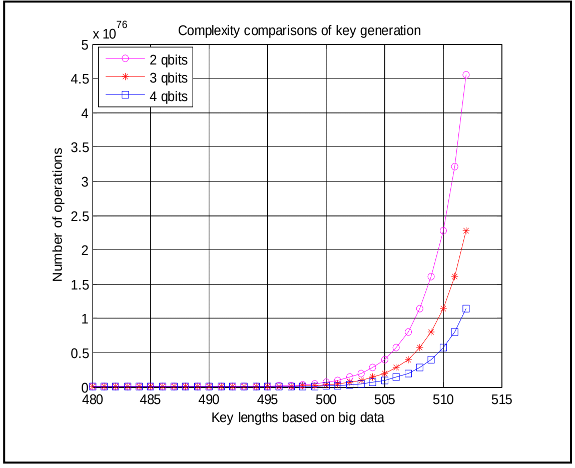

Figure 8: Complexity of key generation processing for big data

According to [8], complexity which is a number of operation can be analyzed for individual services. As shown in Figure 8, the complexity of specific service can be illustrated using QC approach. Again, KM decides the appropriate algorithms and policy to generate the keys which provide the security levels to service providers and users. In this research, GA is used to calculate the number of operations for the required key sizes used in the big data security. Details of GA and necessary procedures are in [8, 18, 19].

5. SECURITY ANALYSIS

In this section, big data security issues based on QC is considered, but we focus on responsibilities of security levels recommended by cloud service providers and users. From the results based on the proposed approach, following cases need to be analyzed for securing big data in cloud computing. They are confidentiality, integrity, authentication, and availability. Although service users do not consider the availability as a security, the service provider must consider the availability as security.

5.1 Confidentiality

In big data security, confidentiality is measured for both service providers and users under different service models used in cloud computing. In IaaS service model, service provider's responsibilities are about 20% but users responsibilities very high. Using proposed techniques, users are able to manage their security responsibilities according to the requirements. Through this research, users' will be able to handle their data with minimum responsibilities of service providers. In PaaS service model, both service providers and users should consider almost same levels of confidentiality. It means that allocated services used in the platform of cloud computing are same from both. Other service models are also compared for measuring security levels.

5.2 Integrity

The integrity requirements of all service models are useful in big data security analysis. It involves maintaining accuracy, atomicity, etc. of big data over its entire lifecycle. The complexity of individual service or interactions of multiple services may be caused to alter the big data. Users' access control should be monitored to improve the security services in cloud computing.

6. CONCLUSIONS

In this research, we have investigated the big data security based on QC which provides better services in cloud computing. Through these studies, we designed the theoretical model which uses PairHand authentication protocol, Grover's algorithm, and QC concepts. This model provides not only the secure communication between cloud users and service providers but also it increase security levels of the cloud environment. Secure communication of cloud environments included with the secure storage and transmission needs appropriate KM with minimum commitment and complexity. Throughout this research, we have analyzed that PairHand protocol employed in the proposed theoretical model that can be implemented for user authentication. The PairHand protocol uses only two handshakes which reduces the complexity and unnecessary commitments such as delay, traffic, etc. Regarding the future scope, the algorithms used in the proposed approach will help us to develop the secure and low-cost 5G based cloud computing with the emerging quantum technologies.

In future work, securing big data framework is one of the challenges. From the proposed theoretical model used in this research should be implemented according to the current resources, architectures, and technology used in the cloud computing. Security intelligence on big data access by users is another point where services are protected automatically and intelligently according to the users' preferences.

7. ACKNOWLEDGMENTS

We are glad that we completed this work successfully. This work would be impossible without the help of Dr. Vijey Thayananthan. I would like to thank him for his expert advice and usual support.

8. REFERENCES

[1] H. A. Elbaz, M. H. Abd-elaziz, & T. Nazmy, "Trusting Identity Based Authentication on Hybrid Cloud Computing," In Cloud Computing, Springer International Publishing, pp. 179-188, 2014.

[2] M. Maurera,∗, I. Brandica, R. Sakellariou, "Adaptive resource configuration for Cloud infrastructure management," Future Generation Computer Systems 29 472–487, 2013.

[3] N. Couch and B. Robins, "Big Data for Defence and Security," report, Royal United Services Institute (RUSI), pp. 2 -36, 2013.

[4] D. He, M. Ma, Y. Zhang, C. Chen and J. Bu, "A Strong User Authentication Scheme with Smart Cards for Wireless Communications," Computer Comm., vol. 34, no. 3, pp. 367-374, 2011.

[5] D. He, J. Bu, S. Chan and C. Chen, "Handauth: Efficient Handover Authentication with Conditional Privacy for Wireless Networks," IEEE Transactions on Computers, VOL. 62, NO. 3, MARCH 2013.

[6] M. S. Farash, M. A. Attari, R. E. Atani, & M. A. Jami, "New efficient authenticated multiple-key exchange protocol from bilinear pairings," Computers & Electrical Engineering, 39(2), 530-541, 2013.

[7] M. S. Farash, & M. A. Attari, "An Enhanced and Secure Three-Party Password-based Authenticated Key Exchange Protocol without Using Server's Public-Keys and Symmetric Cryptosystems," Information Technology And Control, 43(2), 143-150, 2014.

[8] V. Thayananthan and A. Albeshri, "Big data security issues based on quantum cryptography and privacy with authentication for mobile data center," 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15) Elsevier, India, Volume 50, Pages 149–156 2015.

[9] Y. Li, L. Du, G. Zhao, & J. Guo, "A lightweight identity-based authentication protocol," In Signal Processing, Communication and Computing (ICSPCC), 2013 IEEE International Conference on IEEE, pp. 1-4, August 2013.

[10] P. Mehrotra, L. H. Pryor, F. R. Bailey and M. Cotnoir, "Supporting "Big Data" Analysis and Analytics at the NASA Advanced Supercomputing (NAS) Facility," NAS Technical Report: NAS-2014.

[11] K. Lauther and A. Mityagin. Security Analysis of KEA Authenticated Key Exchange Protocol. PKC 2006, volume 3958 of Lecture Notes in Computer Science, pages 378-394. Springer-Verlag, 2006.

[12] B. Grobauer, T. Walloschek, and E. Stöcker, "Understanding Cloud Computing Vulnerabilities," IEEE Security and Privacy, vol. 99, 2010.

[13] S. Subashini, and V. Kavitha, "A survey on security issues in service delivery models of cloud computing," J Network Comput Appl. 2010, doi:10.1016/j.jnca.2010.07.006.

[14] S. Ramgovind, M. M. Eloff, E. Smith. "The Management of Security in Cloud Computing," In PROC 2010 IEEE International Conference on Cloud Computing 2010.

[15] J. Cao, H. Li, M. Ma, et al. "A Simple and Robust Handover Authentication between HeNB and eNB in LTE Networks," Journal Computer Networks: The International Journal of Computer and Telecommunications Networking, 56, pp. 2119-2131, 2012.

[16] W. Stallings , "Cryptography and Network Security: Principles and Practice," Prentice Hall; 5 edition, January 24, 2010.

[17] S. A. Goorden, M. Horstmann, A. P. Mosk, B. Škori, and P. W. H. Pinkse, "Quantum-Secure Authentication Of A Physical Unclonable Key," Optica, Vol. 1, No. 6, December 2014.

[18] V. Thayananthan, A. Alzahrani and M. S. Qureshi, "Efficient techniques of key management and quantum

cryptography in RFID networks," Security and Communication Networks, USA, 2014.

[19] V. Thayananthan and A. Alzahrani, "Analysis of Key Management and Quantum Cryptography in Wireless Sensors Networks," IJCA Special Issue on "Network Security and Cryptograph. (NSC 2011), International Journal of Computer Applications (IJCA), USA, pp. 45-49, 2011.

[20] R. K. Balachandra, P. V. Ramakrishna and A. Rakshit. "Cloud Security Issues," In PROC '09 IEEE International Conference on Services Computing, pp. 517-520, 2009.

[21] P. R. Bryant, R. H. Katz & E. Lazowska, "Big-Data Computing: Creating revolutionary breakthroughs in commerce, science and society," Washington, DC: Computing Community Consortium, 2008.

[22] John E. Hopcroft, Rajeev Motwani, Jeffrey D. Ullman, "Introduction to Automata Theory Languages and Computation". Ch-1, 27-29, 3rd Edition Cornell University, Stanford University, California, USA, 2001.

[23] J. Holub and S. ˇStekr. On parallel implementations of deterministic finite automata. In Implementation and Application of Automata, CIAA '09, pages 54–64, 2009.

[24] V Thayananthan, A. Alrehily and R. Fallatah, "Automata Design with Time Complexity for Intelligent Vending Machine based on Visual Automata Simulator," IEEE 17th International Conference on Computer Modelling and Simulation, UK, University of Cambridge, pages 159-164, 2015.