Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)

Using Role-based to Implement Certificate Authority Management for Big Data

Omar Salah F. Shareef [a] **,** ***Ali Makki Sagheer*** [b]

[a] *University of Fallujah, Iraq.Email: omar.alshareef@uofallujah.edu.iq*

[b] *Al-Qalam University College, Kirkuk, Iraq.Email: prof.ali@alqalam.edu.iq*

ARTICLEINFO

Article history:
Received: 06 /08/2019
Revised form: 25 /08/2019
Accepted: 02 /09/2019
Available online: 29 /09/2019

Keywords:

PKI management, Certificate Authority, Role-Based Access Control, Big data

DOI : 10.29304/jqcm.2019.11.4.626

ABSTRACT

One of the big issues for dynamic organizations is managing the confidential information of big data. Improving security is needed to secure exchange of confidential documents, protection against unauthorized accesses, handling changes in the permissions and roles of people and dealing with the dynamism that can happen if any person leaves or join the system. However, there are limitations in using traditional cryptographic systems and Public Key Infrastructure (PKI) concerning flexibility and manageability. In this paper, we propose a secure and verifiable access control system that implements a Certificate Authority coupled with Role-Based Access Control to provide the permissions to the user to access data. The digital certificate is certified, issued, and revoked by a central administrator; as a result, the certificate is sent based on the role of the user. The proposed scheme has been demonstrated on a big dataset. We believe that our work can be applied to organizations that rely heavily on big data.

1 . Introduction

The information becomes available and playing an important part in everyday life like internet, cloud computing and mobile appliances. Big data are created from social networks like Facebook, Twitter, and YouTube [1]. Moreover, there are 7.6 billion mobile subscribers currently in existence and they are anticipated to reach 9 billion by 2020 [2]. The number of internet users now is about 3.4 billion, and it's increasing rapidly with a rate of 150% each year [3]. Compared with 20 years ago, each second now has data more than which is on the whole internet [4]. However, big and complicated datasets groups are very difficult to manage using the classical system of database management. Thus, the term "Big data" has been appeared and using now in our lives. Big data is considered as developing term that is used to characterize a huge amount of data which can be semi-structured and unstructured that has the possibility to extract information from it. In particular, the Big Data includes three characteristics;

Corresponding author Omar Shareef

Email addresses: omar.alshareef@uofallujah.edu.iq

Communicated by Qusuay Hatim Egaar

Volume, Velocity, and Variety, which are indicated as the 3V's of Big Data. Then two more characteristics added (Value and Veracity) to make it the 5V's of Big Data [5][6].

However, Big Data has resulted in new problems regarding not just to the characteristics of the data but also to data security. The occurrence of Big Data has brought about new challenges regarding data security. The gathering, storage, manipulation and retention of huge quantities of data has led in critical security and privacy considerations [7]. Moreover, with the advent of cloud technology data storage has become accessible and affordable for more and more parties. It becomes harder to ensure security and protect privacy when data is being multiplied and stored on various servers around the world. That data can include sensitive information, such as the privacy of individuals, sensitive corporate data, or sensitive customer data. If the attackers gain access to the sensitive data, they are able to compromise the data in 60% of the cases before being discovered [8]. Further, there are laws that require regulatory compliance and keeping the data safe. This regulated data can contain health, personal, or payment data.

The access to the data needs to be controlled so that wrong entities have no possibility to tamper with or access the data. When improving the security and authentication to sensitive data it can give companies new business opportunities. Thus, high attention has been drawn to the significant to firmly secure the data from any unauthorized access. Sharing big data is a major open and crucial issue [9][10]. In particular, sharing big data can lead to illegal alteration, and impersonated publication and retrieval and another unauthorized accesses [11][12].

In order to prevent such attacks, an authentication scheme is very necessary. Enabling the identities of data users requires authentication with integrity check to inhibit any attacks or illegals changes of information. Therefore, Dynamic organizations need secure solutions to store, exchange confidential information and protect their privacy against unauthorized accesses or disclosures. These solutions need to be flexible enough to cope with changes in people's roles and permissions. To solve these problems, hybrid solutions are building that use certificate authority (CA) coupled with Role-Based Access Control (RBAC) to provide the permissions to the user on resources and it targets access control for computational and file-based storage resources. The user accessing their data by contacts the CA manager, which delegates rights to the data based on the request and the user's role within the system.

The remaining parts of this paper are organized as follows. In Section 2, we discuss related work. In Section 3, we provide preliminaries about the RBAC and CA. Sections 4 provides the proposed model. In Section 5, we provide performance analysis. Finally, we conclude our work in Section 6.

2.Related Work

There are numerous researches concerning the security of big data in both industry and academic environment [13]. Most notably developing protocols and tools for anonymization or encryption of data for confidentiality purposes. This section presents a literature review only on existing work based on the concepts of the certificate authority and big data access control.

A. Access Control

Access control techniques are introduced to protect the privacy, prevent unauthorized access and to selectively share contents in big data. While methods of access control are simple and may merely require comparing credentials, big data have specific access control requirements, due to their particular characteristics.

The widest technique used to analyze the big data is Hadoop, in which numerous methods of access control were presented [27][28][29]. Currently, a number of approaches of attribute-based access control (ABAC) in big data [30][31] have been presented, where the data owners characterize the access policies according to the attributes that are needed by the data and encrypt the data according to the access mechanisms.

Several approaches have emphasized on the resolution of security problems that have been raised by the usage of big data in environments of health-care by applying access control mechanisms to ensure the security of patients' healthcare data [32][33][34][35].

The update of policy is a key task for the management of access control in big data. NTRU (open-source public-key cryptosystem) method has been proposed for storing the big data in clouds to overcome the failures of decryption of the original NTRU. Consequently, presenting a secure and verifiable model of access control [36].

Yang et al. [37] propose a sufficient approach of access control which will allow data owners to dynamically updating policy for big data in the Cloud; they focused on the ciphertext-policy attributed-based encryption (CP-

ABE) to realize their scheme. Besides, some of the researches based on CP-ABE scheme have presented to accomplish access control for Big Data [38][39][40].

Sara et al. [41] proposed an innovative hybrid approach which improves the privacy and security of big data which is shared in a cloud with the use of a method of access control based on the Key-Policy Attribute-Based Encryption (KP-ABE) and authentication system, the ciphered data might be accessed only by legitimate users.

Furthermore, for the solution of problems of security and privacy which limit big data improvement, blockchain-based access control for improving big data platforms security [42][43][44] has been proposed.

B. Certificate Authority

Certificate Authority (CA) is a powerful authentication method which offers scalable secure communication solutions in open networks which are relying on an asymmetric pair of keys defined as public (shared with all the parties) and private (owned only by the user) keys [14].

The blockchain is a linear set of data components, in which every component is referred to as a block. Every block is connected in consecutive order for the sake of composing a chain and secured using cryptography [21]. However, many studies recently propose implementation of blockchain technology to build secure PKI systems [6][22][23][24][25].

Sangram et al. [15][16] have presented health information exchange solutions to fulfill the Health Insurance Portability and Accountability Act (HIPAA) privacy and security regulations with a different cryptographic solution. The proposed schemes demand smart card for every patient. The health care centers have special devices to authenticate patients and deliver the patient information in a secret way. The primary disadvantage of these methods is the smart card can be damaged, lost or stolen.

Fulare et al. [17] propose a new approach based on the key management system to supply secure authentication, that utilizes the Virtual Certificate Authority (VCA). VCA ensured secure data communication between the WSN and securing the networks.

A fully and dynamic distributed certificate authority (DCA) scheme on the basis of a polynomial over the elliptic curve (ECC) and based on trust graphs and threshold cryptography has been used for mobile ad hoc networks (MANET) in [18]. DCA is one of the major approaches which can be used to issue, revoke and handle the certificate in MANET. This scheme takes less computational overhead and provides the same level of security as RSA.

Pahl et al. [19] apply site-local certificates to secure the internet of things microservices. Their solution enables all entities to confirm the right operation of the previous entities in the processing chain.

Zhang et al. [20] introduce a new virtual bridge certificate authority to trust the model and also an effective approach of implementation that can get the cross data-center authentication in the distributed collaborative systems of manufacturing.

Ibrahim et al. have proposed a powerful system based on cryptographic to use for securing the process of data shared in the cloud environment to overcome all the challenges in honest-but-curious cloud environments. The challenges include preserving the confident data, enforcing fine-grained data access control, the application of sufficient approach of user revocation, and forbidding collusion amongst the users of the system [26]. However, getting PKI public/private keys with the related right digital certificate is treated outside the scope of the presented research. Moreover, the secure storage and administration of PKI secret key and the related verification procedures of digital certification are treated outside the scope of this research as well.

3. Preliminaries

To facilitate our proposed scheme, the following entities are briefly introduced.

1. Role-Based Access Control

Access control is an essential function of data security. Its aim is ensuring that authorized individuals alone can get access to particular resources of information. Role-based systems of access control determine several roles and give every one of the roles a group of rights. Each user is then assigned to one or more roles and inherits the permissions assigned to the roles.

Privileges which a user gets are given according to the roles of that user in the organization, and security mechanisms are conveyed based on the job tasks that are related to those tasks in an organization. In advanced RBAC approaches, administrators can put restrictions. For instance, there is a possibility for a maximum number of individuals that have the authority of simultaneously activating sessions of duty for the same role; which is the restriction of cardinality which is put on the role.

2. Certificates Authority

Certificates are electronic files that use digital signature for associating a public key with an identification (both of an individual or a business enterprise), therefore the certification that the public key belongs to a person. The signing of certificates executed through both a certificate authority (CA) or the individual, or maybe other individuals. A CA is an entity which is responsible for signing certificates. A person who would like to verify their identification predicated on certificates authority for signing their certificate. The certificates handed to some other person that could check it towards the issuing certificate authority for validity. Each party needs to accept as true with the certificates authority that verifies their identity with any other certificate, which is signed by highest-rank certificates. Certainly, there ought to be one certificating authority which does not have higher authority. This certificate authority signs their own certificates that is a self-signed certificate also called as the root certificate [46]. The main operations of the certificate authority regarding our scheme are:

a) Certificate generation

X.509 certificate is a standard that is widely used for defining digital certificates. X.509 makes use of PKI (Public Key Infrastructure) to confirm the identification of a user with the public key. X.509 certificate includes information concerning the identity which a certificate is issued to and the identity which has issued it. Typical information in X.509 certificates includes the following:

- *Subject*: the recognized entity (an individual or an organization).
- *Serial number*: which is utilized for uniquely identifying the certificate.
- *Key-usage*: the objective of the public key (signature confirmation, encryption, or both).
- *Signature algorithm*: the approaches which are utilized to create the signature.
- *Issuer*: identifies which issued the certificate.
- *Public key*: which is utilized for the encryption of a message to the subject or for the verification of the signature from the subject.
- *Thumbprint algorithm*: which is the approach which is utilized for hashing.
- *Valid from*: the date from which the certificate is valid.
- *Thumbprint*: the hash for the certificate which is utilized for verifying that the certificate has not been changed.
- *Valid to*: the date until the certificate is valid.

b) Certificate Signature

It is a certificate authority class for issuing user certificates to clients, with the use of their own secret key and their subject information. The pair of keys and the subject information from the request of the purchaser joined for generating an X.509 digital certificate which has been signed with the use of the secret key of the certificate authority using SHA256WithRSA as the algorithm of signature and provided to the user (which means the issuer has RSA public key inside the certificate, and the hash algorithm that used for signing the leaf certificate is SHA-256). Certificate path (the root certificate subject is the sub-certificate issuer) is made up of the certificate authority and the certificate of the user.

c) Certificate Validation

It is utilized for the validation of the certificate of the client for the certificate authority. It has the aim to check the validity of the digital certificate that has been delivered by the client. In this sense, the certificate authority reads the certificates of the client, which include the path of the certificates, and checks the certificates in the following manner:

- Checking the validity duration, or else, it is not valid;
- Checking the certificate signature with the use of the public key of root certificate, or else it is not authentic;

- Checking the certificate’s serial number whether or not it's far within the latest Revocation List (RL), the certificate is not valid in the case where it is true;
- Recursively checking the certificate along the path of the certificate one by one until the certificate of the root, the certificate is valid if it’s root certificate.

4. The Proposed Model

The proposed approach brings the benefit of integrating the two functions of identity authentication and access control. Role-based access control has been coupled to the Certificate authority management to build a secure system.

The CA administrator assigns each role a certificate based on the organizational position such as secretary, manager, and employee.

The digital certificate, which contains the client role and its public key; is certified, issued, and revoked to the client by a centralized administrator. The onus is at the administrator for ensuring the safety of confidential data. Besides, the touchy data items continuously reported to the administrator. The primary characteristic of the certificate is the authentication of the identification of the certificate's owner to others. A certificate includes the owner’s public key, at the same time as the owner retains the secret key. The public key, which utilized for encrypting the messages, is transferred to the certificate owner who is the only one who gets the right of the entry to the secret key, therefore, only the owner is capable of decrypting these messages.

The most significant of the proposed system is the using of the role characteristic in generating the certificate. This will determine the accessibility of the user with the certificate. The following section will introduce more details concerning the design of the certificate authority management, as shown in Figure 1.

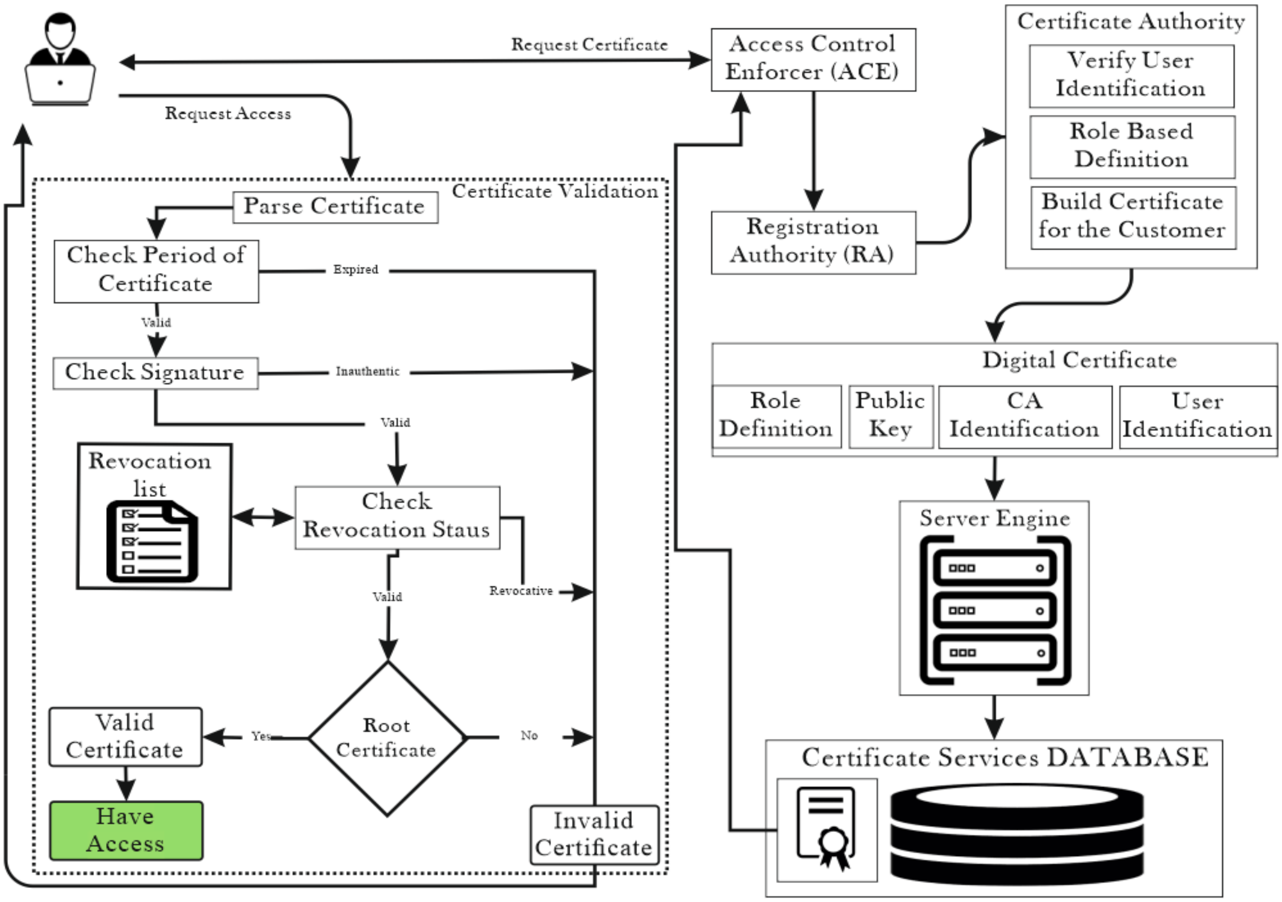

Fig. 1. The Proposed System

However, the proposed scheme contains five basic operations:

- At first, we generate the X.509 root certificate and the PKCS#12 (Public Key Cryptography Standard, PKCS) in addition to the corresponding CA's private key. A root certificate is a certificate that contains the public key of a CA. Clients can trust a CA only if a copy of the CA root certificate is in its trusted root certificate store. Moreover, the CA public key included in the CA root certificate is needed to verify the validity of any certificate that the CA issues. PKCS is a set of standard protocols to exchange secure information on the Internet using public key infrastructure. However, PKCS#12 is a standard that specifies a portable format for storing a user's private key.
- The CA class issues the certificate to the client based on the role given to the user by the system manager. The role of the user is defined during the creation of the certificate, which will determine the access provided by the granted certificate.
- The validation of the certificate comes from the CA, which also based on the role of the client. In addition, to check the validation of the client digital certificate as shown in Figure 1.
- The given certificate is valid for a limited time, the client has to renew the certificate before expiry. The client requests a certificate renewal from the issuer, the issuer checks whether the certificate has not expired and not been revoked. The issuer ignores the request as long as the certificate has been revoked; or else, the request is granted. Then, the issuer produces a new certificate with limited time.
- CA administrator can revoke any issued certificate to any users in the instance of suspicion in the public key/identity binding. Also, can revoke any certificate if the private key of the certificate has been compromised.

The algorithm of our proposed scheme is shown in the following table.

TABLE 1. The Algorithm of the Proposed Model

```
Step1   The client request access
Step2   Check the client certificate
            o   Parse the client certificate
                    ▪   If the client has certificate go to the next condition
                    ▪   Else the client does not have certificate and go to step3
            o   Check the period of the client certificate
                    ▪   If valid go to the next condition
                    ▪   Else expired and go to Step3
            o   Check the certificate signature
                    ▪   If valid go to the next condition
                    ▪   Else inauthentic and go to Step3
            o   Check the revocation list
                    ▪   If valid go to the next condition
                    ▪   Else outmoded and go to Step3
            o   Check the certificate path until the root certificate
                    ▪   If it's root certificate it is valid and the client have access
                    ▪   Else invalid certificate and go to Step3
Step3   The client request new certificate
Step4   The ACE receives the user requests and returns the results to the users.
Step5   Registration authority (RA) verifies user requests for a digital certificate and tells the certificate authority (CA) to issue it.
Step6   Certificate Authority (CA) verify user id and define the client role to issues digital certificates.
Step7   Generate digital certificate that contain the client role and it is public key
Step8   Generated certificate exported to the server engine
Step9   Store the generated certificate at the certificate services database and inform the (ACE)
Step10  ACE returns the results to the users.
Step11  End
```

The dataset used for our model named Chicago Business Licenses and Owners [45]. This data-set includes a big number of data records/rows. This data-set includes the information of the owner for each account that is listed in the Business License Data-set and sorted according to the Number of the Account. C# has been chosen as the programming language because of the independence of its platform. Moreover, the open code library Bouncy Castle has been used for writing the program, which is an implementation of cryptographic algorithms of C# and Java and developed by the Legion of Bouncy Castle.

5. Performance Analysis

In order to avoid the forgery and cyber fraud of certificates, creating own CA makes securely stored certificate authority administrator's secret key. In that case, it can be of higher security compared to relying on an untrusted third party.

The proposed solution has advantages concerning manageability and flexibility against solutions that are entirely based on traditional CA. In the proposed scheme, access control data is managed centrally based on the role of the user. Particularly provides the capabilities that allow the administrator to issue, renew, and revoke certificates upon requesting access. As follows, the administrative infrastructure that provides the registration and initialization services leads to no end-user involvement required.

In our work, the owner of the data does not have to do a thing; that will save the time that will spend in the documentation of policies of data sharing. The administrator is the only one that is required to identify the certificate and keep updating it whenever necessary. The users are allowed to access data according to its sensitivity which is computed from user roles and the data itself. Suppose the individuals that are authorized to access certain data in an organization could vary by month, week, or even by day, based on their role (such as administrator, employee ...etc.). Contrary to that, certificates are usually designed for being effective for a considerably longer period (for example, one or two years). If it is necessary frequently revoking and reissuing certificates, this could severely impact on the performance properties of the resultant system of certificate management. Therefore, our model can minimize the workload of data owners, as well as one of the cloud servers.

6. Conclusions

Big data access control is an effective method for ensuring big data security. In the presented research, we have proposed an innovative model of access control, according to certificate authority management. Particularly, role-based access control has been considered from an organizational point of view. Which raises the function of authentication with the use of X.509 certificates to the authorization management level.

The most important advantage of assigning certificates using an employee's role is that sessions of duty may be done on every server with no need for connecting again to the servers of authentication and authorization unless the certificate expired. Which is why the needed effort of communication is minimized. The performance analysis of this research shows that the proposed certificate authority management with the use of RBAC provides a scalable way of big data access control. We believe that the contribution of this study is a significant step in the direction of providing efficient and strong security management based on the roles of users in the big data platform.

References

[1] H. V. Jagadish et al., “2,” Commun. ACM, vol. 57, no. 7, pp. 86–94, 2014.

[2] J. L. M. Shafiullah Khan, Green Networking and Communications, 1st ed. 2014.

[3] C. V. Konstantinos Samdanis, Peter Rost, Andreas Maeder, Michela Meo, Green Communications: Principles, Concepts, and Practice, 1st ed. Wiley, 2015.

[4] E. Brynjolfsson and A. McAfee, “Big Data : The Management Review,” Harv. Bus. Rev., no. October, pp. 1–12, 2012.

[5] J. Moura and C. Serrao, “Security and Privacy Issues of Big Data,” vol. 2, 2016.

[6] W. Wang, N. Hu, and X. Liu, “BlockCAM: A Blockchain-Based Cross-Domain Authentication Model,” in 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), 2018, pp. 896–901.

[7] B. Thuraisingham, “Bldg Ddwd Shfxulw\ Dqg Pulydf\,” pp. 279–280.

[8] L. Cheng, F. Liu, and D. D. Yao, “Enterprise data breach: causes, challenges, prevention, and future directions,” Wiley Interdiscip. Rev. Data Min. Knowl. Discov., vol. 7, no. 5, pp. 1–14, 2017.

[9] D. S. Terzi, R. Terzi, and S. Sagiroglu, “A survey on security and privacy issues in big data,” in 2015 10th International Conference for Internet Technology and Secured Transactions (ICITST), 2015, pp. 202–207.

[10] D. Wu, B. Yang, and R. Wang, “Scalable privacy-preserving big data aggregation mechanism,” Digit. Commun. Networks, vol. 2, no. 3, pp. 122–129, 2016.

[11] R. Li, H. Asaeda, J. Li, and X. Fu, “A verifiable and flexible data sharing mechanism for information-centric IoT,” in 2017 IEEE International Conference on Communications (ICC), 2017, pp. 1–7.

[12] R. Li, H. Asaeda, J. Li, and X. Fu, “A distributed authentication and authorization scheme for in-network big data sharing,” Digit. Commun. Networks, vol. 3, no. 4, pp. 226–235, 2017.

[13] J. Moreno, M. A. Serrano, and E. Fernández-Medina, “Main issues in Big Data security,” Futur. Internet, vol. 8, no. 3, 2016.

[14] H. Gomes, J. P. Cunha, and A. Zúquete, “Authentication architecture for ehealth professionals,” in OTM Confederated International Conferences" On the Move to Meaningful Internet Systems", 2007, pp. 1583–1600.

[15] S. Ray and G. P. Biswas, “A Certificate Authority (CA)-based cryptographic solution for HIPAA privacy/security regulations,” J. King Saud Univ. - Comput. Inf. Sci., 2014.

[16] S. Ray and G. P. Biswas, “Design of RSA-CA Based E-Health System for Supporting HIPAA Privacy-Security Regulations,” Procedia Technol., vol. 6, pp. 954–961, 2012.

[17] R. P. Fulare and A. V. Sakhare, “Efficient sensor node authentication in wireless integrated sensor networks using virtual certificate authority,” Proc. - 2014 4th Int. Conf. Commun. Syst. Netw. Technol. CSNT 2014, pp. 724–728, 2014.

[18] A. Alomari, “Fully distributed certificate authority based on polynomial over elliptic curve for MANET,” SNPD 2013 - 14th ACIS Int. Conf. Softw. Eng. Artif. Intell. Netw. Parallel/Distributed Comput., pp. 96–100, 2013.

[19] M. O. Pahl and L. Donini, “Securing IoT microservices with certificates,” IEEE/IFIP Netw. Oper. Manag. Symp. Cogn. Manag. a Cyber World, NOMS 2018, pp. 1–5, 2018.

[20] W. Zhang, X. Wang, and M. K. Khan, "A virtual bridge certificate authority-based cross-domain authentication mechanism for distributed collaborative manufacturing systems," Secur. Commun. Networks, vol. 8, no. 6, pp. 937–951, 2015.

[21] Y. Zhang, X. Lin, and C. Xu, "Blockchain-Based Secure Data Provenance for Cloud Storage," in International Conference on Information and Communications Security, 2018, pp. 3–19.

[22] L. M. Axon and M. Goldsmith, "PB-PKI: A privacy-aware blockchain-based PKI," 2016.

[23] A. Yakubov, W. Shbair, A. Wallbom, and D. Sanda, "A blockchain-based PKI management framework," in The First IEEE/IFIP International Workshop on Managing and Managed by Blockchain (Man2Block) colocated with IEEE/IFIP NOMS 2018, Tapei, Tawain 23-27 April 2018, 2018.

[24] Z. Qikun, G. Yong, Z. Quanxin, W. Ruifang, and T. Yu-An, "A dynamic and cross-domain authentication asymmetric group key agreement in telemedicine application," IEEE Access, vol. 6, pp. 24064–24074, 2018.

[25] C. Fromknecht, D. Velicanu, and S. Yakoubov, "Certcoin: A namecoin based decentralized authentication system 6.857 class project," Unpubl. Cl. Proj., 2014.

[26] I. M. Mahmoud, S. H. Nour El-Din, R. Elgohary, H. Faheem, and M. G. M. Mostafa, "A robust cryptographic-based system for secure data sharing in cloud environments," Secure. Commun. Networks, vol. 9, no. 18, pp. 6248–6265, 2016.

[27] T. K. Ashwin Kumar, H. Liu, J. P. Thomas, and X. Hou, "Content sensitivity based access control framework for Hadoop," Digit. Commun. Networks, vol. 3, no. 4, pp. 213–225, 2017.

[28] M. Gupta, F. Patwa, and R. Sandhu, "POSTER: Access control model for the Hadoop ecosystem," Proc. ACM Symp. Access Control Model. Technol. SACMAT, vol. Part F1286, pp. 125–127, 2017.

[29] A. Gupta, K. Pandhi, P. V. Bindu, and P. S. Thilagam, "Role and Access Based data Segregator for Security of Big Data," Procedia Technol., vol. 24, pp. 1550–1557, 2016.

[30] M. Gupta, F. Patwa, and R. Sandhu, "An Attribute-Based Access Control Model for Secure Big Data Processing in Hadoop Ecosystem," pp. 13–24, 2018.

[31] J. Longstaff and J. Noble, "Attribute-Based Access Control for Big Data Applications by Query Modification," Proc. - 2016 IEEE 2nd Int. Conf. Big Data Comput. Serv. Appl. BigDataService 2016, pp. 58–65, 2016.

[32] A. E. Youssef, "A FRAMEWORK FOR SECURE HEALTHCARE SYSTEMS B ASED O N BIG DATA ANALYTICS I N MOBILE CLOUD," vol. 2, no. 2, pp. 1–11, 2014.

[33] L. Qiu, F. Cai, and G. Xu, "Quantum digital signature for the access control of sensitive data in the big data era," Futur. Gener. Comput. Syst., vol. 86, pp. 372–379, 2018.

[34] N. Lu and R. Jiang, "An Adaptive Access Control Model Based on Trust and Risk for Medical Big Data," in 2018 IEEE 3rd International Conference on Communication and Information Systems (ICCIS), 2018, pp. 232–236.

[35] Y. Yang, X. Zheng, W. Guo, X. Liu, and V. Chang, "Privacy-preserving smart IoT-based healthcare big data storage and self-adaptive access control system," Inf. Sci. (Ny)., vol. 479, pp. 567–592, 2019.

[36] C. Hu, W. Li, X. Cheng, J. Yu, S. Wang, and R. Bie, "A Secure and Verifiable Access Control Scheme for Big Data Storage in Clouds," IEEE Trans. Big Data, vol. 4, no. 3, pp. 341–355, 2017.

[37] K. Yang, X. Jia, K. Ren, R. Xie, and L. Huang, "Enabling efficient access control with dynamic policy updating for big data in the cloud," Proc. - IEEE INFOCOM, pp. 2013–2021, 2014.

[38] Q. Yuan, C. Ma, and J. Lin, “Fine-Grained Access Control for Big Data Based on CP-ABE in Cloud Computing,” pp. 344–352, 2014.

[39] K. Yang, Q. Han, H. Li, K. Zheng, Z. Su, and X. Shen, “An Efficient and Fine-Grained Big Data Access Control Scheme with Privacy-Preserving Policy,” IEEE Internet Things J., vol. 4, no. 2, pp. 563–571, 2017.

[40] S. Khuntia and P. S. Kumar, “New Hidden Policy CP-ABE for Big Data Access Control with Privacy-preserving Policy in Cloud Computing,” 2018 9th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2018, pp. 1–7, 2018.

[41] A. Sara, T. Yassine, and M. Abdellatif, “Secure confidential big data sharing in cloud computing using KP-ABE,” pp. 1–4, 2017.

[42] H. Es-Samaali, A. Outchakoucht, and J. P. Leroy, “A Blockchain-based Access Control for Big Data,” Int. J. Comput. Networks Commun. Secur., vol. 5, no. 7, pp. 137–147, 2017.

[43] S. Wang, Y. Zhang, and Y. Zhang, “A blockchain-based framework for data sharing with fine-grained access control in decentralized storage systems,” IEEE Access, vol. 6, pp. 38437–38450, 2018.

[44] E. B. Sifah et al., “Chain-based big data access control infrastructure,” J. Supercomput., vol. 74, no. 10, pp. 4945–4964, 2018.

[45] https://www.kaggle.com/chicago/chicago-business-licenses-and-owners

[46] A. Rensburg and B. Solms, “A comparison of schemes for certification authorities/Trusted Third Parties,” Inf. Secure. Res. Bus., pp. 222–240, 1997.