Survival Analysis on Secured Data Communication in Cloud

S.L. Swapna
Research Scholar
Hindusthan College of Arts and Science
(Autonomous)
Coimbatore, India

V. Saravanan
Professor
Hindusthan College of Arts and Science
(Autonomous)
Coimbatore, India

ABSTRACT

Cloud computing is here to stay within Internet technology that presents user everything in terms of services like computing power to computing infrastructure, applications, business processes over the web. With fast growth of big data analytics, network security is an essential research field. Secure transmission is defined as the data transfer like confidential information through the secure channel. Secure communication is the process of communicating with two entities and not want third party to listen. The people share their information with diverse degrees of certainty. The data security is an essential role for guaranteeing the drastic development of number of cloud users. Many secured data storage mechanisms are designed by many researchers in past. But, the available mechanisms are time consuming process and not attained the necessary security. It failed to provide adequate security analytics performance with higher accuracy and minimal time consumption. In order to address these problems, existing secured data communication in cloud is reviewed.

General Terms

Cloud, Data Communication, Data Security

Keywords

Cloud Computing, Big Data Analytics, Network Security, Data Storage

1. INTRODUCTION

Cloud computing described the different types of computing ideas that involve large number of computers linked through real-time communication network. Cloud computing is used in large cluster of manageable resources to distribute optimum resource consumption. Cloud computing comprised different types of configurable distributed systems with connectivity and usage [1]. Data security is an essential problem in data communication. Many efficient and secure schemes are introduced to preserve the network from external attacks like the one in [2].

This article is structured as follows: Section 2 describes the review of various secured data communication techniques in cloud, Section 3 presents study and analysis of existing secured data communication techniques, Section 4 explains the comparison between them. Section 5 discusses the limitations of conventional secured data communication techniques.

2. RELATED WORKS

Security is one of the major concerns in handling big data [3] especially in cloud computing environment. An energy-efficient and big data-based secure framework was designed in [4] with Internet of Things (IoT) for green environment. The designed framework reduced the overhead and energy consumption for reliable transmission. Though the energy consumption and overhead was minimized, the accuracy level was not increased by energy-efficient and big data secure framework.

The big data security access control algorithm was introduced in [5] depending on memory index acceleration. Though the time complexity was reduced, the security level was not improved by big data security access control algorithm. Secure Authentication and Data Sharing in Cloud (SADS-Cloud) architecture was designed in [6] for solving the Big data security issues over the Cloud. The data users contribute in secure file retrieval. However, the computational complexity was not minimized by SADS Cloud architecture.

Incentive-based protection and recovery strategy was introduced in [7] to manage the virus spread and to increase the data security. A protection and recovery strategy (PRS) minimized the infected users. However, the time complexity was not reduced by incentive-based protection and recovery scheme. The network security and big data protection strategy was used in [8] to emphasize the factors affecting the network security platforms in Big data. But, the data security level was not enhanced through big data protection strategy.

Secured Map Reduce (SMR) Layer was built in [9] with the security and privacy layer between HDFS and MR Layer. A privacy and security guaranteed and preserved the privacy-utility trade-off for the data miners. However, the privacy level was not increased by SMR Layer. The data processing flow of access control strategy was performed in [10] with big data systems. The designed strategy reduced the time complexity and increased the system performance. But, the access control strategy failed to increase the security level by access control strategy.

A cloud-enabled IoT environment was employed in [11] through multifactor authentication. However, the data communication was not performed in secured manner with cloud-enabled IoT environment to preserve the big data system. A new deep learning-based data minimization algorithm was designed in [12] to minimize the dimensionality during the transfer over the carrier channels. Though dimensionality was reduced, the accuracy level was not increased through deep learning-based data minimization algorithm.

Ruidong Li et al. [13] put forth a data security framework called DCAuth for Big Data retrieval over the cloud. Also an identity-based signcryption technique was introduced in [14] by joining the encryption and signature for secure and authenticated communication in Big data. But, data confidentiality rate was not increased through identity-based signcryption technique in the above two frameworks. Besides, a secured User Authentication Protocol was been proposed in

[15] for Big Data retrieval through an IoT-based agent by without affecting data retrieval efficiency.

3. SECURED DATA COMMUNICATION IN CLOUD

Cloud computing is the huge development of on-demand networks that provide the access to configurable resources with minimal management effort. With fast development of the data sources, big data security in Cloud is a big challenge. Different problems were addressed in area of big data security like infrastructure security, data privacy, data management and data integrity. Secure data communication is an essential issue during message transmission over the networks. Security in data communication is an essential one for transferring the message between sender and receiver to maintain the message as confidential one.

3.1 EBDS: An Energy-Efficient Big Data-Based Secure Framework Using Internet of Things for Green Environment

An energy-efficient and big data-based secure (EBDS) framework was introduced with Internet of Things for green environment. An IoT-based sensor was employed for data gathering and performed the data routing by Dijkstra-based optimal algorithm. The designed framework reduced overhead and energy consumption through finding reliable and least transmission distance routes. EBDS framework generated the big data from network attackers and maintained the consistency of green environment. The graph-based learning method was designed with optimal set of relay nodes for environmental data forwarding. The selection criterion algorithm was employed with node attributes between sensor nodes. The designed algorithm provided security level to preserve the big data from the malicious actions and attains network privacy with data reliability and certainty. The designed framework provided the route conservation method to reduce the chance of data failure and increased the fraction of routing reliability. EBDS framework comprised two main sub-components, namely data routing and security with integrity. The nodes were arranged in connected graph. The routing table origin depends on network information. The relay nodes were selected by highest weighted rank with optimal finding depending on Dijkstra algorithm and managed the multi-hop communication. The environmental data was preserved from network threats to guarantee the integrity and privacy.

3.2 Big Data Security Access Control Algorithm Based on Memory Index Acceleration in WSNs

Wireless sensor network (WSN) is an autonomous wireless communication system with large number of micro-sensor nodes. The big data security access control algorithm was introduced depending on memory index acceleration in WSNs. The access control was employed to guarantee the data security. In big data system, access control operation was carried out with policy, information, decision and implementation information for attaining right to the access data when the visitor accesses the data. The data storage structure was fundamental concern of access control determined by data characteristics of WSNs. The second-level cache was employed to minimize the number of cycles during policy extraction. A memory index was constructed to reduce access time of security policy. The second-level cache and index content gets updated to guarantee the efficiency of policy variations. The access control authority efficiency gets increased devoid of affecting access control security. The policy-based model was the fine-grained security control model to handle the data security on the Hadoop platform. When right resource was varied, the algorithm updated the indexes in secondary cache synchronously.

3.3 Novel System Architecture for Secure Authentication and Data Sharing in Cloud Enabled Big Data Environment

A new architecture termed Secure Authentication and Data Sharing in Cloud (SADS-Cloud) was introduced with three processes, namely big data outsourcing, big data sharing and big data management. In Big data outsourcing, data owners were registered to the Trust Center using SHA-3 hashing algorithm. MapReduce model was employed to divide the input file into fixed data block size and SALSA20 encryption algorithm was employed over every block. In Big data sharing, data users contribute in secure file retrieval. Trust Center was employed to register and monitor the data owner behavior. User privacy was considered for data sharing through authentication process. Cross product of User ID and Password was generated for registration. The data compression was carried out to minimize complexity for data encryption. Lempel Ziv Markow Algorithm (LZMA) compressor was the best compressor with Arithmetic coding and Huffman compression. Access control was carried out to all data as data owners encrypt data depending on the sensitivity level. Data owners categorize the files into two types, namely sensitive or non-sensitive. SALSA20 with MapReduce process was employed to encrypt the large volume of data. SALSA20 Encryption algorithm was introduced to address SDES encryption for large volume of data and applications. The clustering and indexing was carried out for efficient big data management over cloud. The files were stored in Cloud servers by Density-based Clustering of Applications with Noise (DBSCAN) algorithm. Indexing was carried out with Fractal Index Tree to perform insertions, deletions, and searching operation.

4. PERFORMANCE ANALYSIS ON DIFFERENT SECURED BIG DATA COMMUNICATION METHODS

In order to compare the different secured big data communication techniques, number of data points is taken as an input to conduct the experiments. Experimental evaluation of three methods namely energy-efficient and big data-based secure (EBDS) framework, big data security access control algorithm and Secure Authentication and Data Sharing in Cloud (SADS-Cloud) are implemented using Java language. In order to perform secure data communication, the smart healthcare application is taken using MHEALTH dataset. The dataset is taken from UCI machine learning repository https://archive.ics.uci.edu/ml/datasets/MHEALTH+Dataset.

The MHEALTH (Mobile Health) dataset is used for monitoring human behavior through the multimodal body sensing. The sensor is employed as IoT device positioned on subject chest, right wrist, and left ankle to determine motion experienced by different body parts like acceleration, rate of turn and magnetic field orientation. For each subject, thousands of data generated and stored in a different log file. Each file included the attributes for all sensors. Result analysis of existing techniques is estimated with certain parameters like Data Confidentiality Rate, Time Complexity and Accuracy.

4.1 Data Confidentiality Rate

Data confidentiality rate is defined as the ratio of number of data accessed by authorized user to the total number of data. The data confidentiality rate (*DCR*) is determined as,

$$DCR = \left(\frac{n}{N}\right) \text{ x } 100 \qquad (1)$$

Where *n* represents the number of data accessed by the authorized user and *N* denotes the total number of data accessed. It is computed in terms of percentage (%).

Table 1. Tabulation for Data Confidentiality Rate

Number of Data points (Number)	Data Confidentiality Rate (%)		
	EBDS framework	Big Data Security Access Control Algorithm	SADS-Cloud
100	90	85	78
200	92	88	80
300	95	90	82
400	93	87	79
500	91	84	76
600	87	82	73
700	85	80	71
800	82	78	69
900	80	75	66
1000	83	78	70

Table 1 describes the data confidentiality rate with respect to number of data points varying from 100 to 1000. Data confidentiality rate comparison takes place on existing EBDS framework, Big Data Security Access Control Algorithm and SADS-Cloud. The graphical representation of data confidentiality rate is illustrated in the figure 1.

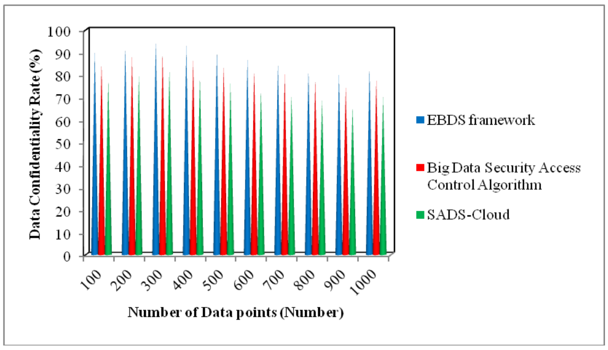

Fig 1: Measurement of Data Confidentiality Rate

From above figure 1, data confidentiality rate depending on different number of data points is described. The blue color cone denotes the data confidentiality rate of EBDS framework. The red color cone and green color cone symbolizes the data confidentiality rate of big data security access control algorithm and SADS-Cloud correspondingly. It is clear that data confidentiality rate using EBDS framework is higher when compared to big data security access control algorithm and SADS-Cloud. This is because of applying the graph-based learning method with optimal set of relay nodes for environmental data forwarding. The relay nodes were chosen through highest weighted rank depending on Dijkstra algorithm and managed multi-hop communication. Consequently, data confidentiality rate of EBDS framework is increased by 6% when compared to the big data security access control algorithm and 18% when compared to SADS-Cloud respectively.

4.2 Time Complexity

Time complexity is defined as the product of number of data and amount of time consumed by one data to perform the secure data communication. The time complexity (*TC*) is estimated as,

$$TC = n \text{ x } T \qquad (2)$$

In Eq.2, *n* denotes the number of data and *T* represents the time consumed to perform a single data communication. The time complexity is measured in terms of milliseconds (ms).

Table 2. Tabulation for Time Complexity

Number of Data points (Number)	Time Complexity (ms)		
	EBDS framework	Big Data Security Access Control Algorithm	SADS-Cloud
100	36	25	31
200	39	28	33
300	42	31	35
400	45	34	38
500	48	37	40
600	51	40	42
700	54	43	45
800	57	46	48
900	60	49	51
1000	63	52	55

Table 2 describes the time complexity with respect to number of data points ranging from 100 to 1000. Time complexity comparison takes place on existing EBDS framework, Big Data Security Access Control Algorithm and SADS-Cloud. The graphical representation of time complexity is explained in figure 2.

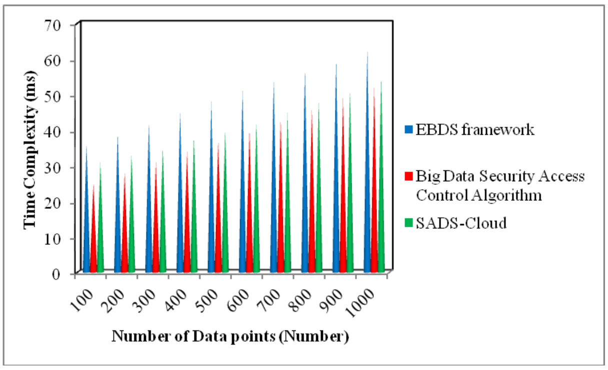

Fig 2: Measurement of Time Complexity

From above figure 2, time complexity depending on different number of data points is illustrated. The blue color cone in figure represents the time complexity of EBDS framework. The red color bar and green color cone represents the time complexity of Big Data Security Access Control Algorithm and SADS-Cloud correspondingly. It is observed that time complexity using big data security access control algorithm is lesser when compared to EBDS framework and SADS-Cloud. This is because of applying second-level cache to minimize the number of cycles during policy extraction. A memory index was constructed to reduce access time of security policy. Therefore, time complexity of big data security access control algorithm is reduced by 23% when compared to the EBDS framework and 9% when compared to SADS-Cloud respectively.

4.3 Authentication Accuracy

Authentication accuracy (AA) is defined as the number of cloud users that are correctly authenticated to the total number of cloud users. It is computed as,

$$AA = \left(\frac{nAU}{NU}\right) \text{x } 100 \qquad (3)$$

where 'nAU' denotes the number of correctly authenticated cloud users and 'NU' represents total number of cloud users. It is measured in terms of percentage (%).

Table 3. Tabulation for Authentication Accuracy

Number of Data points (Number)	Authentication Accuracy (%)		
	EBDS framework	Big Data Security Access Control Algorithm	SADS-Cloud
100	78	82	87
200	80	84	90
300	83	87	92
400	81	85	91
500	79	82	88
600	77	80	86
700	80	83	89
800	82	86	91
900	84	89	93
1000	87	91	94

Table 1 explains the authentication accuracy with respect to number of data points ranging from 100 to 1000. Authentication accuracy comparison takes place on existing EBDS framework, Big Data Security Access Control Algorithm and SADS-Cloud. The graphical representation of authentication accuracy is described in figure 3.

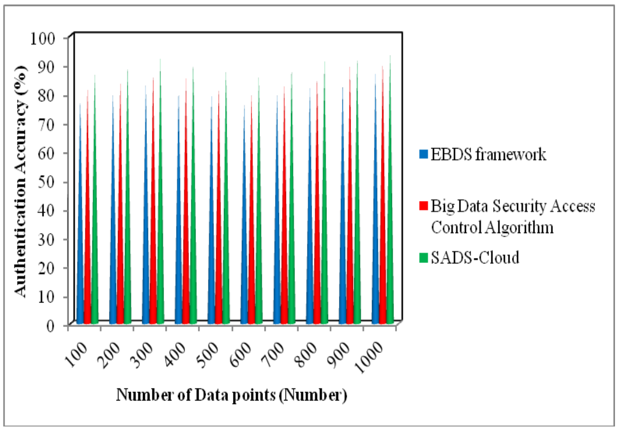

Fig 3: Measurement of Authentication Accuracy

From above figure 3, authentication accuracy with different number of data points is explained. The blue color cone represents the authentication accuracy of EBDS framework. The red color bar and green color cone represents the authentication accuracy of Big Data Security Access Control Algorithm and SADS-Cloud respectively. It is observed that authentication accuracy using SADS-Cloud is higher when compared to EBDS framework and big data security access control algorithm. This is due to the application of trust centre to register and monitor the data owner behavior. The second-level cache and index content guaranteed the efficiency of policy variations. Therefore, authentication accuracy of SADS-Cloud is increased by 11% when compared to the EBDS framework and 6% when compared to big data security access control algorithm respectively.

5. DISCUSSION

An energy-efficient and big data-based secure framework was designed with IoT for green environment. IoT-based sensors were linked for data gathering and data routing using Dijkstra-based optimal algorithm. Dijkstra-based optimal algorithm minimized the overhead and energy consumption. Though energy consumption and overhead was minimized, the accuracy was not increased by energy-efficient and big data-based secure framework. Secure Authentication and Data Sharing in Cloud (SADS-Cloud) was introduced in Big Data environment. The data owners were registered to Trust Center by SHA-3 hashing algorithm. MapReduce model split the input file into fixed-size of data blocks. But, the computational complexity was not reduced by SADS-Cloud architecture.

The big data security access control algorithm operated depending on memory index acceleration in WSN. The access control authority efficiency was increased without reducing the security level. A second-level cache minimized the number of cycles during the policy extraction. A memory index minimized the access time of the security policy. Though the time complexity was minimized, the security level was not increased by big data security access control algorithm.

6. CONCLUSION

A comparative study of different secured data communication techniques with Big Data is carried out. From the survival study it was observed that the security level was not increased by big data security access control algorithm. In addition, the computational complexity was not reduced by SADS-Cloud architecture. Though energy consumption and overhead was minimized, the accuracy was not increased by energy-efficient and big data-based secure framework. The wide experiment on conventional techniques evaluates the results of different secured data communication techniques and discusses its issues. From the result analysis it is concluded that an efficient research framework need to be carried out and the same with the support of machine learning and ensemble learning techniques for secured data communication with higher accuracy and lesser time consumption. While various machine learning frameworks of secured data communications emerging, most of the works are in compromising nature with time. Hence, the upcoming research works of the secured cloud communication has to be aimed at improving the data confidentiality rate and to reduce the time consumption by adopting machine learning and deep learning methods.

7. REFERENCES

[1] Medhavi, S. Shriwas., Neetesh, Gupta., and Amit, Sinhal.

2012. Comparative Study of Cloud Computing and Mobile Cloud Computing. In IJCA Proceedings on National Conference. 13-19.

[2] Hatem, Hamad., and Mahmoud, Al-hoby. 2012. Managing Intrusion Detection as a Service in Cloud Networks. International Journal of Computer Applications. 41(1), 35-40.

[3] Swapna, S. L., and Saravanan, V. 2021. Big Data Challenges and Learning Paradigms: A Review. Journal of University of Shanghai for Science & Technology. 23(12),36-45.

[4] Khalid, Haseeb., Soojeong, Lee., and Gwanggil, Jeon. 2020. EBDS: An energy-efficient big data-based secure framework using Internet of Things for green environment. Environmental Technology & Innovation. Elsevier. 20, 1-19.

[5] Jianhua, Peng., Hui, Zhou., Qingjie, Meng., and Jingli, Yang. 2020. Big data security access control algorithm based on memory index acceleration in WSNs. EURASIP Journal on Wireless Communications and Networking. Springer. 90, 1-17.

[6] Uma, Narayanan., Varghese, Paul., and Shelbi, Joseph. 2020. A novel system architecture for secure authentication and data sharing in cloud enabled Big Data Environment. Journal of King Saud University - Computer and Information Sciences. 1-15.

[7] Youke, Wu., Haiyang, Huang,, Ningyun, Wu., Yue, Wang,, Md, Zakirul, Alam, Bhuiyan., and Tian, Wang. 2020. An incentive-based protection and recovery strategy for secure big data in social networks. Information Sciences. Elsevier. 508, 79-91.

[8] Imane, El, Alaoui., and Youssef, Gahi. 2020. Network Security Strategies in Big Data Context. Procedia Computer Science. Elsevier. 175, 730-736.

[9] Priyank, Jain., Manasi, Gyanchandani., and Nilay, Khare. 2019. Enhanced Secured Map Reduce layer for Big Data privacy and security. Journal of Big Data. 6 (30), 1-15.

[10] Jianhua, Peng., Hui, Zhou., Qingjie, Meng., and Jingli, Yang. Big data security access control algorithm based on memory index acceleration in WSNs. EURASIP Journal on Wireless Communications and Networking. Springer. 90, 1-17.

[11] Saleh, Atiewi., Amer, Al-Rahayfeh., Muder, Almiani., Salman, Yussof., Omar, Alfandi., Ahed, Abugabah., and Yaser, Jararweh. 2020. Scalable and Secure Big Data IoT System Based on Multifactor Authentication and Lightweight Cryptography. IEEE Access. 8, 113498-113511.

[12] Mohammed, Aledhari., Marianne, Di, Pierro., Mohamed, Hefeida., and Fahad, Saeed. 2021. A Deep Learning-Based Data Minimization Algorithm for Fast and Secure Transfer of Big Genomic Datasets. IEEE Transactions on Big Data. 7(2), 271 – 284.

[13] Ruidong, Li., Hitoshi, Asaeda., and Jie, Wu. 2020. DCAuth: Data-Centric Authentication for Secure In-Network Big-Data Retrieval. IEEE Transactions on Network Science and Engineering. 7(1), 15 – 27.

[14] Dharminder, Dharminder., Mohammad, S. Obaidat., Dheerendra, Mishra., and Ashok, Kumar, Das. 2021. SFEEC: Provably Secure Signcryption-Based Big Data Security Framework for Energy-Efficient Computing Environment. IEEE Systems Journal. 15(1), 598-606.

[15] Srinivas, Jangirala., Ashok, Kumar, Das., Mohammad, Wazid., and Athanasios, V. Vasilakos. 2021. Designing Secure User Authentication Protocol for Big Data Collection in IoT-Based Intelligent Transportation System. IEEE Internet of Things Journal. 8(9), 7727 – 7744.

.