IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124182

Collusion Resistant Secure Outsourcing of Sequence Comparison Using Cloud Computing Algorithm

Thenmozhi K[1], Pavithra S[2], Balaji A S[3], Maheswari M[4]

Student, Computer science and engineering, Anand Institute of Higher Technology, Chennai, India1,2

Assistant Professor, Computer Science and Engineering, Anand Institute of Higher Technology,Chennai, India[3]

Assistant Professor, Computer Science and Engineering, Anand Institute of Higher Technology,Chennai, India[4]

Abstract: Any large-scale data sharing system needs this quality more than others because if a user accidentally discloses sensitive information, the owner of the data will find it harder to keep the information secure.For data owners, sharing their data on servers or in the cloud presents many difficulties. These methods are crucial for managing keys shared by the data owner. The trustworthy authority will be introduced in this article in order to authenticate users who have access to cloud data. The trustworthy authority generates the key using the SHA algorithm, and both the owner and the user will have access to it. The trusted authority module computes a hash value using the MD-5 method after receiving an AES-encrypted file from the data owner. a file is sent to the CSP module by a trusted authority for cloud storage.

Keywords: SHA algorithm, AES algorithm**,** MD-5 algorithm.

1. INTRODUCTION

Currently, public clouds offer primary sequence comparison algorithms as a universal outsourcing service. However at the same time, more and more security andprivacy concerns are being raised. Individual private information carried by character sequences could more or less be divulged or misused, and the outsourced data saved as plaintext could readily be made available to malevolent internal and external attackers in the CSP. In order to preserve the confidentiality of character sequences and to guarantee that the planned computing requests are typically handled by cloud servers, secure outsourcing was developed.

2.RELATED WORKS

Our strategy is straightforward to implement, effective to process, and overhead- controllable.The contributions of this paper mainly in the following four aspects. Based on the universal model of a public cloud outsourcing, we propose an overall architecture for E-SC. This architecture is built on the end user and the un modified CSP. Its overall system model, which has been demonstrated to be secure under the threat model, is user-friendly and implementation-friendly. To protect against statistical attacks, a salted hash technique is developed to hash character sequences and the indexes of cost matrices. An additive order preserving encryption algorithm is designed to encrypt the elements of cost matrices. Moreover, this approach can accomplish a linear time complexity additive ordered chosen-plaintext attack indistinguishability. For the first time, a single cloud server can deliver a privacy- preserving compute outsourcing service to successfully fend off cloud collusion attacks. In the non-interactive sequence comparison stage, there is no need to decrypt any outsourced data thanks to per-processing modules for padding, partitioning, and expansion. Simulation resultsshow that the overall execution performance of our E-SC is negatively correlated with its security.

3.EXISTING SYSTEM

Internet computing technologies, such grid computing, which enable the vast cooperative sharing of computational resources, arerevolutionising the way that large-scale issues in the physical and life sciences are currently solved.power, bandwidth, storage, and data. Once connected to one of these grids, a poor computational device is no longer constrained by its slow speed, scant local storage, and constrained bandwidth since it can access the wealth of these resources that are available elsewhere on the network. either one's own data or the result of a computation using one's own data.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124182

4.PROPOSED SYSTEM

We suggest a safe data sharing strategy that can ensure safe crucial sharing and datasharing for a dynamic group. We give a secure way for crucial distribution without any securecommunication channels. The druggies can securely gain their private keys from groupdirector without any Certificate Authorities due to the verification for the public key of the stoner.We propose a secure data participating scheme which can be defended from conspiracy attack. The abandoned druggies can't be suitable to get the original data lines once they're abandoned indeed if they conspirewith the un trusted pall.

1. IMPLEMEMTATION

5.1 ADVANCE ENCRYPTION ALGORITHM(AES)

The new encryption standard proposed by NIST to replace DES is called Advanced Encryption Standard. The only known attack that successfully breaks the encryption is a brute force attack, in which the attacker tries every possible character combination. Both AES and DES are block ciphers. A 128, 192, or 256 bit variable key length is available. Depending on the key size, it encrypts data blocks of 128 bits in 10, 12, and 14 rounds. Fast and adaptable AES encryption is available. It can be used on many platforms, particularly small smartphones. AES has also undergone thorough testing for a variety of security applications.

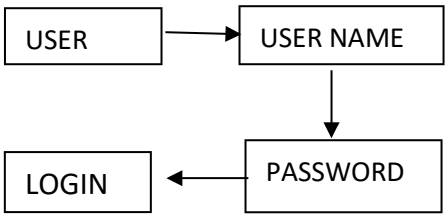
Major advanatges of AES over DES are

1. Data block size is 128 bit
2. Depending on the version, key size is 128/192/256 bits
3. AES is now generally supported by hardware, which makes it extremely quick
4. It uses substitution and permutations.
5. Among the potential keys are 2128, 2192, and 2256 [10
6. More secure than DES.
7. Most adopted symmetric encryption algorithm.

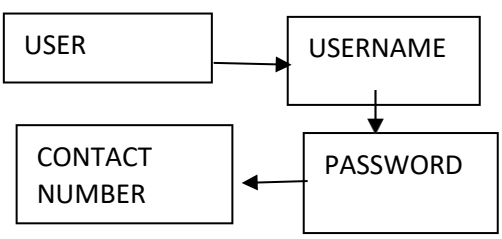
5.2 MD5 ALGORITHM: The MD5 function is a cryptographic technique that generates a message digest that is 128 bits long from an input of any length. The digest may alternatively be referred to as the input's "hash" or "fingerprint." In many circumstances when a potentially lengthy message needs to be swiftly processed and/or compared, MD5 is employed. The production of digital signatures and their verification is the most widespread application. Ronald Rivest, a renowned cryptographer, created MD5 in 1991. MD5 was revealed to have some significant weakness in 2004. It is still unknown how these problem would affect everything.

5.3 MODULES

Login Module



5.3.2 USER REGISTRATION



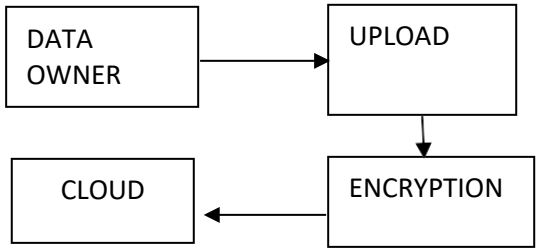
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023

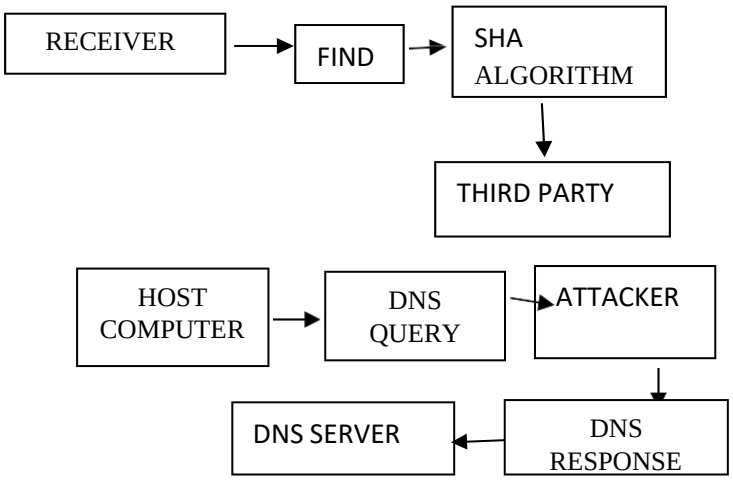
DOI: 10.17148/IJARCCE.2023.124182

5.3.3Key Enrollment



Collusion Detect

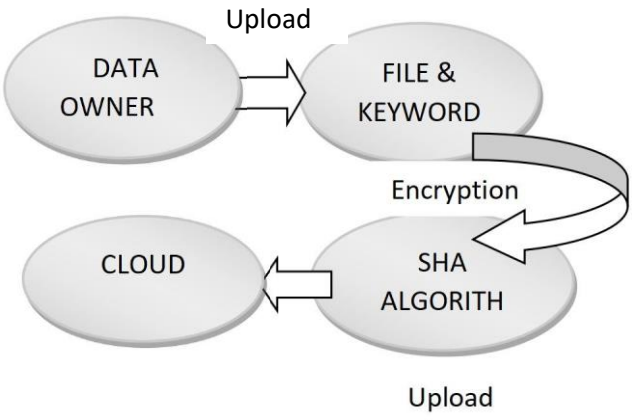
Third-Party Detection



DATA FLOW DIAGRAM

A data flow diagram (DFD) models the process elements of an information system by graphically representing the "flow" of data through it. They frequently serve as an initial step in the creation of an overview of the system that can then be developed. DFDs can be used to visualise data processing as well.

DFD-LEVEL 0:DATA OWNER



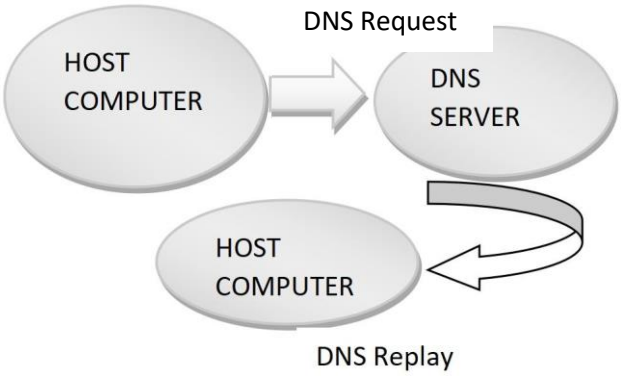
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023

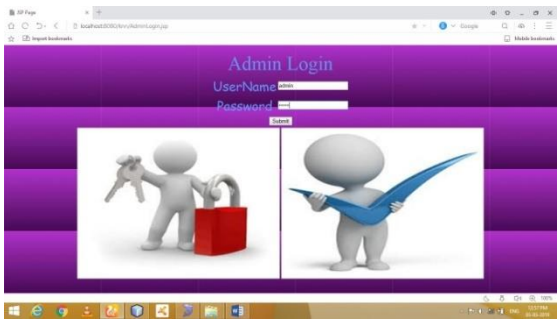
DOI: 10.17148/IJARCCE.2023.124182

DFD –Level 1:



6.RESULTS AND CONCLUSION

Admin Login



User Login

Registration

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

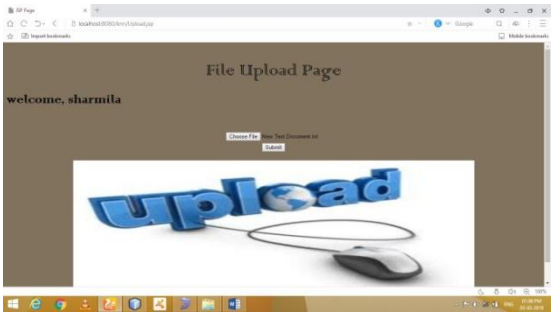
International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023

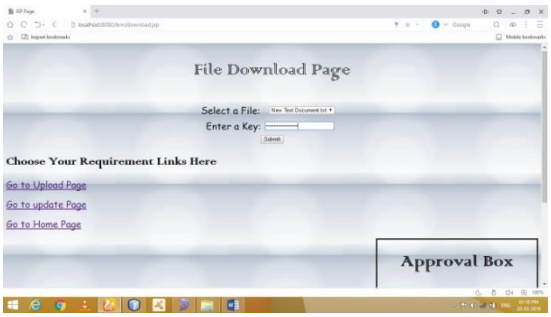
DOI: 10.17148/IJARCCE.2023.124182

Create New Group

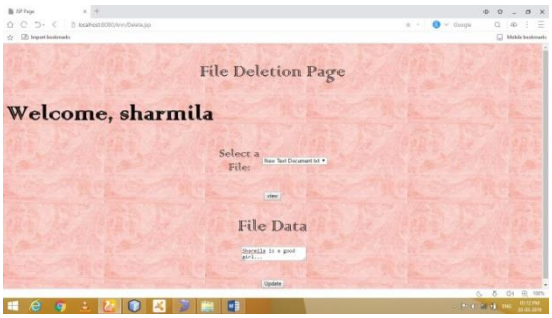
Data Upload Page



File Download Page



Update Page



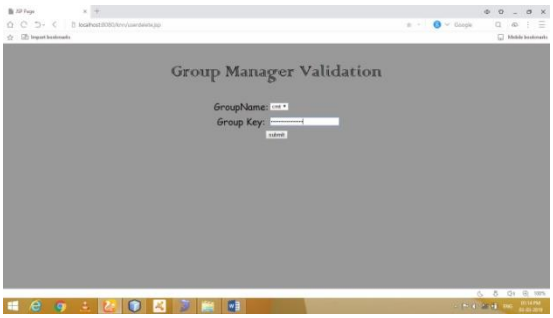
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

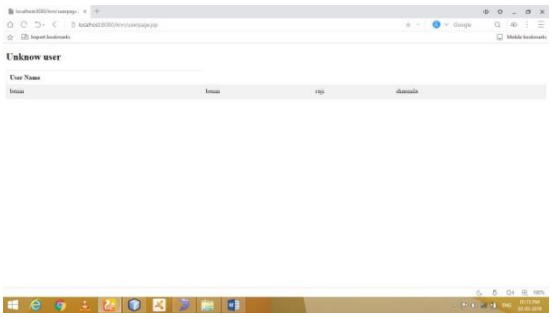
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124182

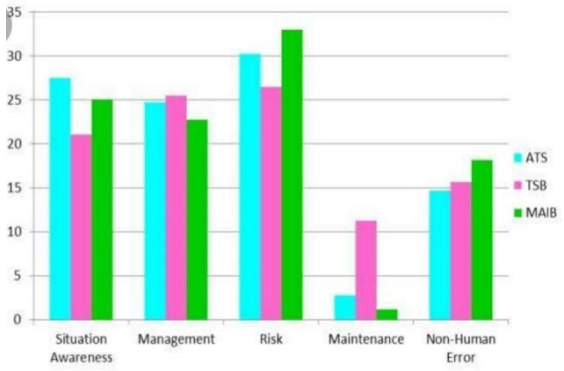
User Revocation Details



Find Third Party



Performance Analysis



7.CONCLUSION

Through the problems about the collusion attacks that are widespread in thesecure outsourcing of sequence comparison algorithms, this content will introduce the trusted authority to authenticate user those who have the access to the data on cloud. SHAalgorithm is used by the trusted authority to generate the key and that key will get share to user as well as the owner. The trusted authority module receives encrypted file using AES Algorithm from the data owner and computes hash value using MD-5 algorithm. It stores key in its database which will be used during the dynamic operations and to determine the cheating party in the system.Trusted authority send file to CSP module to store on cloud. The resulting key sets are shown to have a number of desirable properties that ensure the confidentiality of communication sessions against collusion attacks by other network nodes.

FUTURE ENHANCEMENT

It is somewhat hard to extend the work in our paper to certain applications with multi- data source. Firstly, two character sequences from different sources should be encrypted respectively with different keys. Secondly, three cost

DOI: 10.17148/IJARCCE.2023.124182

matrices should be encrypted together after being constructed by the negotiation between both sides. The security target is to complete sequence comparison on a single cloud server in the way of privacy preservation and to ensure that the string typed data of the end user on any side will not be arbitrarily stolen by the other user or the CSP.

REFERENCES

[1]Y.Feng,H.Ma,andX.Chen,‘‘Efficientandveri fiableoutsourcingscheme of sequence comparisons,’’ Intell. Autom. Soft Comput., vol. 21, no. 1, pp. 51–63, Jan. 2015.
[2]M. J. Atallah and J. Li, "Secure outsourcing of sequence comparisons," Proc.Toronto, Ontario, Canada: International Workshop Privacy Enhancing Technologies (PET), 2004, pp. 63–78
[3]Secure and private sequence comparisons by M. J. Atallah, F. Kerschbaum, and W. Du were published in Proc. ACM Workshop Privacy Electron. Soc. Washington, DC, USA: (WPES), 2003, pp. 39–44.
[4]At Proc. Netw. Distrib. Syst. Secur. Symp. (NDSS), San Diego, CA, USA, 2006, pp. 253- 265, D. Szajda, M. Pohl, J. Owen, and B. Lawson a distributed version of the Smith- Waterman genome sequence comparison technique, discuss their work on "Toward a realistic data privacy approach.
[5]X. Chen, J. Li, J. Ma, Q. Tang, and W. Lou, "New algorithms for secure outsourcing of modular exponentiations," 2014 September; IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 9, pp. 2386–2396.
[6]R. Akimana, O. Markowitch, and Y. Roggeman, "Secure outsourcing of DNA sequences comparisons in a Grid context," WSEAS Trans. Comput. Res., vol. 2, no. 2, February 2007, pp. 262-269.
[7]M. Blanton, M. J. Atallah, K. B. Frikken, and"Secure and efficient outsourcing of sequence comparisons," by Q. Malluhi, published in Proc. Eur. Symp. Res. Comput. Secur. (ESORICS), Pisa, Italy, 2012, pp. 505– 522.
[8]Y. Feng, H. Ma, X. Chen, andSecure and verifiable outsourcing of sequence comparisons by H. Zhu, Proc.Int. Conf. Inf. Commun. Technol.(ICT-EurAsia), Yogyakarta, Indonesia, 2013, pp. 243–252.
[9]X. Chen, J. Li, P. Li, and S. Salinas, "A tutorial on safe outsourcing of large-scale computations for big data," 2016 April, IEEE Access, vol. 4, pp. 1406–1416.
[10]X. Chen, J. Li, J. Weng, J. Ma, and W. Lou, "Verifiable computation over huge databases with incremental updates," IBM Trans. Comput., vol. 65, no. 10, pp. 3184– 3195, Oct. 2016
[11]]B.Berger,N.M.Daniels,andY.W.Yu,‘‘Com putationalbiologyinthe21st century: Scaling with compressive algorithms,’’ Commun. ACM, vol. 59, no. 8, pp. 72–80, Aug. 2016.
[12]E. Ayday, J. L. Raisaro, J.-P. Hubaux, and J. Rougemont,"Protecting and assessing genetic privacy in medical diagnostics and individualised medicine, "in Proc.ACM Workshop Privacy Electron.Soc.(WPES), Berlin,Germany,2013, pp.95-106.
[13]Secure genomic testing with size- and position-hiding private substring matching was described by E. D. Cristofaro, S. Faber, and G. Tsudik in Proc. ACM Workshop Privacy Electron. Soc. (WPES), Berlin,Germany,2013, pp. 107–117.
[14]At Proc. Int.Conf. Financial Cryptography and Data Security (FC), Puerto Rico, 2015, pp. 194–212, J. H. Cheon, M. Kim, and K. Lauter published a paper titled "Homomorphic computation of edit distance."
[15]K. Hua, Q. Yu, and R. Zhang, ‘‘A guaranteed similarity metric learning framework for biological sequence comparison,’’ IEEE/ACM Trans. Comput. Biol. Bioinf., vol. 13, no. 5, pp. 868–877, Sep.2016.
[16]N. S. Vo, Q. Tran, N. Niraula, and V. Phan, ‘‘RandAL: A randomized approach to aligning DNA sequences to reference genomes,’’ BMC Genomics, vol. 15, p. S2, Jul. 2014.
[17]Pattern Recognit. Lett., vol. 62, pp. 1-7, September 2015; V. Palazón-González and Marzal, "Increasing the cyclic edit distance with early abandon using LAESA.".
[18]K. Lauter, A. López-Alt, and In Proc. Naehrig, "Private computation on encrypted genetic data," Int. Conf. Cryptol. Inf. Secur. Latin Amer. (LATINCRYPT), Florianópolis, Brazil, 2015, pp. 3–27.
[19]K.Shimizu,K.Nuid and G.Rätsch, ‘‘Efficient privacy-preserving string search and an application in genomics,’’ Bioinformatics, vol. 32, no. 11, pp. 1652–1661, Jun. 2016.
[20]J.KatzandY.Lindell,Introduction to Modern Cryptography. BocaRaton, FL, USA: CRC Press, 2014, pp. 241–284.

