VERIFIABLE SEARCHABLE ENCRYPTION FRAMEWORK AGAINST INSIDER KEYWORD-GUESSING ATTACK IN CLOUD STORAGE

K.Hari Krishna [1], M.Pranay Teja Reddy (21S15A6706) [2], K.Sathvika (21S15A6708) [3], Prajwal Gondi (21S15A6705) [4], M.Varshitha (20S11A6751) [5],
ASSISTANT PROFESSOR [1], UG STUDENTS [2,3,4,5,]
DEPARTMENT OF CSE(DATA SCIENCE)
MALLA REDDY INSTITUTE OF TECHNOLOGY & SCIENCE,
Maisammaguda, Medchal (M), Hyderabad-500100, T. S

ABSTRACT

Searchable encryption (SE) allows cloud tenants to retrieve encrypted data while preserving data confidentiality securely. Many SE solutions have been designed to improve efficiency and security, but most of them are still susceptible to insider Keyword-Guessing Attacks (KGA), which implies that the internal attackers can guess the candidate keywords successfully in an off-line manner. Also in existing SE solutions, a semi-honest but- curious cloud server may deliver incorrect search results by performing only a fraction of retrieval operations honestly (e.g., to save storage space). To address these two challenging issues, we first construct the basic Verifiable SE Framework (VSEF), which can withstand the inside KGA and achieve verifiable searchability. Based on the basic VSEF, we then present the enhanced VSEF to support multi-keyword search, multi-key encryption and dynamic updates (e.g., data modification, data insertion, and data deletion) at the same time, which highlights the importance of practicability and scalability of SE in realworld application scenarios. We conduct extensive experiments using the Enron email dataset to demonstrate that the enhanced VSEF achieves high efficiency while resisting to the inside KGA and supporting the verifiability of search results.

Introduction:

Cloud storage [1], [2] offers a cost-effective way to store big data and deal with the rapid growth of data variety, volume, and velocity. It is estimated that approximately 40% of big data will be kept or processed by the cloud in 2020 [3]. Although cloud storage has become entrenched in our digitalized society, there are challenges that have yet to be resolved. For example, both the cloud server and data owners are typically not in the same trusted domain, and hence security and privacy risks remain issues that need to be considered when utilizing cloud storage solutions. Some solutions have been presented for different security risks. For example, to ensure data utilization over encrypted cloud data, one typical solution is to utilize searchable encryption (SE) [4], [5], [6], [7], which enables the cloud to retrieve encrypted data based on a set of keywords on behalf of data users. Existing SE schemes generally fall into two divisions, namely: symmetric SE (SSE) and asymmetric SE (ASE) [3]. Although the former enjoys higher efficiency than the latter, costs associated with secret key management can be prohibitive. ASE, also known as public key Encryption with Keyword Search (PEKS) [5], is relatively effective in multiuser settings. Despite the versatile features (e.g., guaranteeing data security and utilization) provided by SE technique, risks such as those due to hardware failures, software bugs, network attacks, and so on, yet remain. As data owners lose direct physical control over their data, the malicious cloud (e.g., compromised Google cloud, Amazon cloud) [8], [9] may discard rarely or never accessed data, or even conceal data loss incidents [10]. In other words, the malicious cloud is a stronger adversary than the honest-but-curious adversary1, and may yield a fraction of inaccurate search results. To guarantee the exactness of search results and detect such misbehavior, it is arguably essential to have in-place a verification mechanism within SE. Such a scheme is referred to as verifiable keyword search in the literature [11], [12]. Cryptographic or Bloom Filter (BF) based verification mechanism has been used to ensure the correctness of retrieved results without accessing the raw data.

LITERATURE SURVEY

Zhihua Xia [1] received the BS degree in Hunan City University, China and PhD degree in computer science and technology from Hunan University, China, in 2006 and 2011, respectively. He works as an associate professor in School of Computer & Software, Nanjing University of Information Science & Technology. His research interests include digital forensic and encrypted image processing. He is a

member of the IEEE. Leqi Jiang [2] is a Ph.D. student in School of Computer and Software, Nanjing University of Information Science & Technology. He received his M.S. degree in Software Engineering from Nanchang Hang Kong University in 2016. His research interests include encrypted image processing, data security in cloud. Dandan Liu [3] received her BS in network engineering from Nanjing University of Information Science & Technology in 2015, China. She is currently pursuing her MS in computer science and technology at the College of Computer and Software, in Nanjing University of Information Science & Technology, China. Her research interest is encrypted image retrieval. Lihua Lu [4] received her BS in network engineering from Nanjing University of Information Science & Technology in 2016, China. She is currently pursuing her MS in computer science and technology at the College of Computer and Software, in Nanjing University of Information Science & Technology, China. Her research interest is encrypted image retrieval. Byeungwoo Jeon [5] received the B.S. degree (Magna Cum Laude) in 1985, the M.S. degree in 1987 in electronics engineering from Seoul National University, Seoul, Korea, and the Ph.D. degree in electrical engineering from Purdue university, West Lafayette, in 1992. Since September 1997, he has been with the faculty of the School of Electronic and Electrical Engineering, Sungkyunkwan University, Korea, where he is currently a full Professor. He has authored many papers in the areas of video compression, pre post processing, and pattern recognition. His research interests include multimedia signal processing, video compression, statistical pattern recognition, and remote sensing.

Existing System:

Verifiable SE: The first SSE scheme [4] is designed to allow the cloud server to implement searches without sacrificing functionality for security. Although this particular scheme has several crucial advantages (e.g., controlled searching and hidden query), it has high key management overhead in a symmetric setting. Thus, many researchers mainly pay attention to PEKS scheme in the public-key setting, as evidenced by the range of schemes proposed in the literature. Examples include conjunctive keyword search [26], fuzzy keyword search [27], ranked keyword search [28], [29], attribute-based keyword search [30], [31].

Disadvantages of Existing System:

- The system is not implemented Decisional Bilinear Diffie-Hellman (DBDH) Assumption.
- The system is not implemented Public key Encryption with Keyword Search (PEKS).

Proposed System:

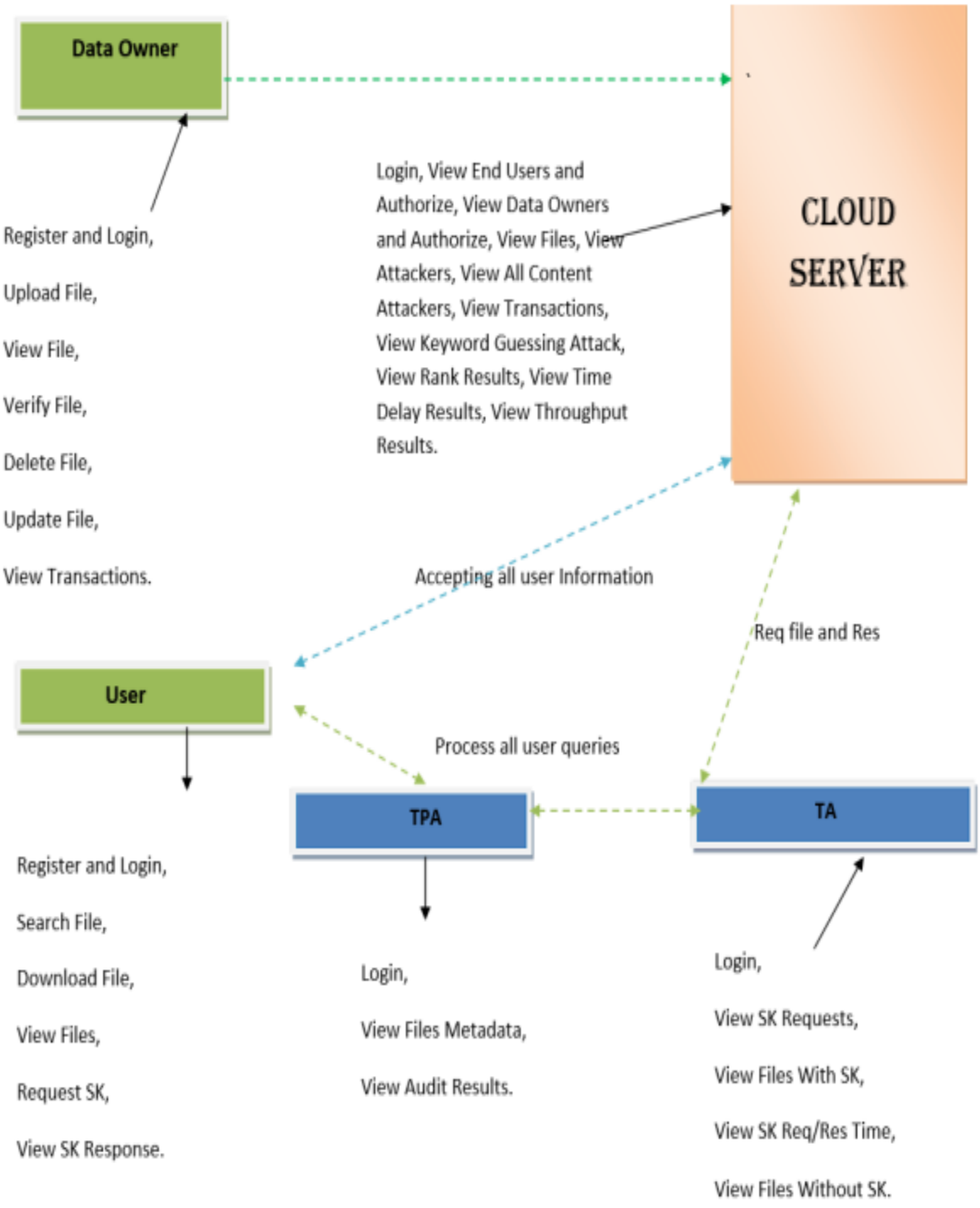
In this paper, we first introduce a basic Verifiable SE Framework against insider KGA (hereafter referred to as basic VSEF), by extending the public auditing technique [22] to SE scheme. In the basic VSEF, the costly correctness verification tasks are assigned to a fullytrusted third-party auditor, and in turn, the auditor honestly reports the auditing results to cloud clients.

- Verifiable keyword search.
- Resisting insider KGA.
- Multi-keyword search.
- Multi-key encryption.
- Dynamic update

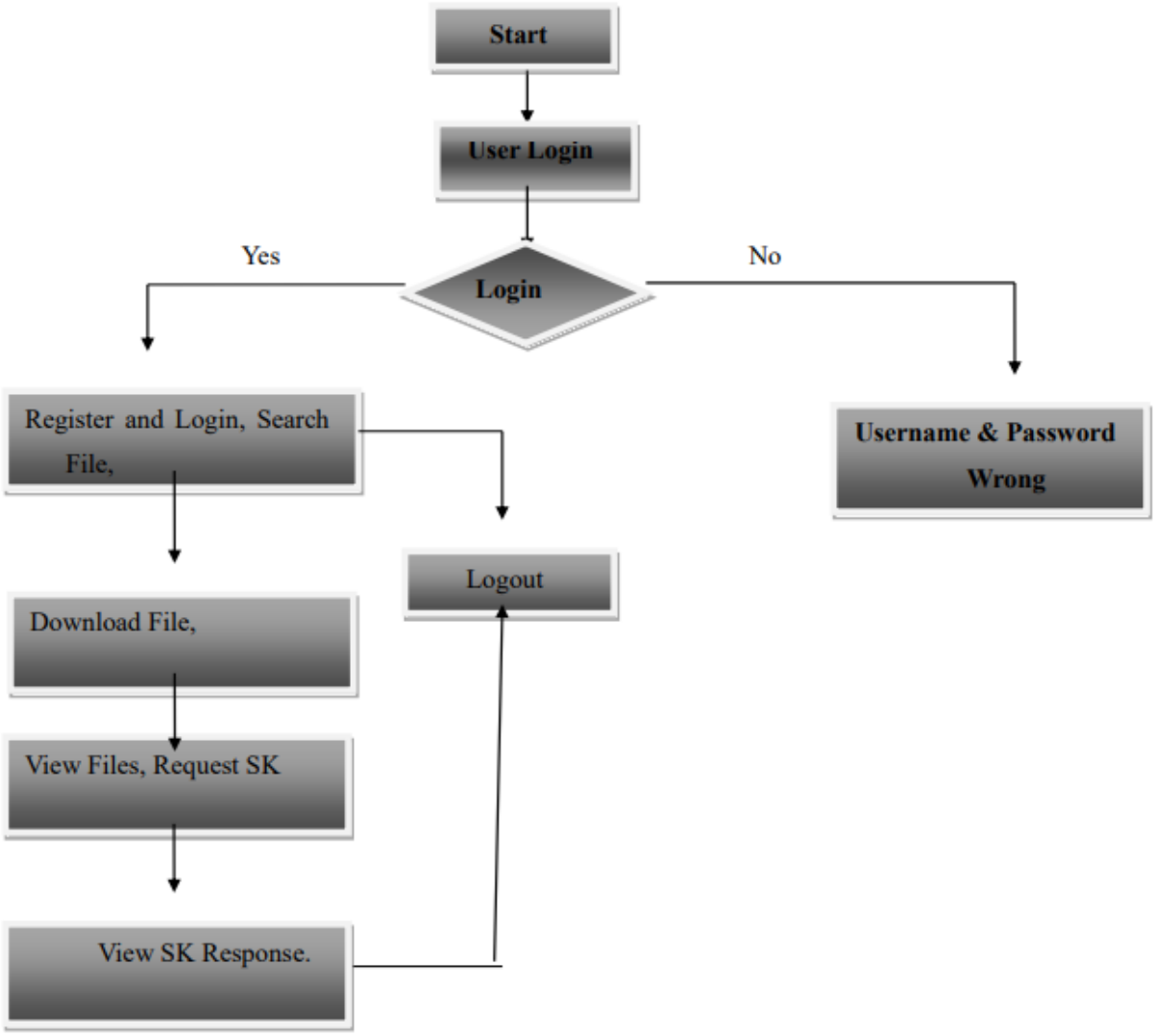
Advantages of Proposed System:

- Public auditing: Data auditing schemes, such as those presented in this system, allow cloud clients to examine the integrity of remote cloud data without downloading them locally.
- Inspired by public auditing schemes, the proposed system devises the basic VSEF by combining with the SE technique.

SYSTEM DESIGN



Flow Chart:



Hardware Requirements

- System : Intel Core i3
- Hard Disk : 250GB.
- Ram : 4 GB.

Software Requirements

- Operating System : windows 7 & above
- Coding Language : Java/J2EE (JSP, Servlet), MySQL
- Front End : J2EE
- Back End : MySQL

INPUT AND OUTPUT DESIGN

INPUT DESIGN

Input Design plays a vital role in the life cycle of software development, it requires very careful attention of developers. The input design is to feed data to the application as accurate as possible. So, inputs are supposed to be designed effectively so that the errors occurring while feeding are minimized. According to Software Engineering Concepts, the input forms or screens are designed to provide to have a validation control over the input limit, range and other related validations. This system has input screens in almost all the modules. Error messages are developed to alert the user whenever he commits some mistakes and guides him in the right way so that invalid entries are not made. Let us see deeply about this under module design. Input design is the process of converting the user created input into a computer-based format. The goal of the input design is to make the data entry logical and free from errors. The error is in the input are controlled by the input design. The application has been developed in user-friendly manner. The forms have been designed in such a way during the processing the cursor is placed in the position where must be entered. The user is also provided with in an option to select an appropriate input from various alternatives related to the field in certain cases. Validations are required for each data entered. Whenever a user enters an erroneous data, error message is displayed and the user can move on to the subsequent pages after completing all the entries in the current page.

RESULTS

OUTPUT DESIGN

The Output from the computer is required to mainly create an efficient method of communication within the company primarily among the project leader and his team members, in other words, the administrator and the clients. The output of VPN is the system which allows the project leader to manage his clients in terms of creating new clients and assigning new projects to them, maintaining a record of the project validity and providing folder level access to each client on the user side depending on the projects allotted to him. After completion of a project, a new project may be assigned to the client. User authentication procedures are maintained at the initial stages itself. A new user may be created by the administrator himself or a user can himself register as a new user but the task of assigning projects and validating a new user rest with the administrator only. The application starts running when it is executed for the first time. The server has to be started and then the internet explorer in used as the browser. The project will run on the local area network so the server machine will serve as the administrator while the other connected systems can act as the clients. The developed system is highly user friendly and can be easily understood by anyone using it even for the first time.

Fig 1: Output of Web Page

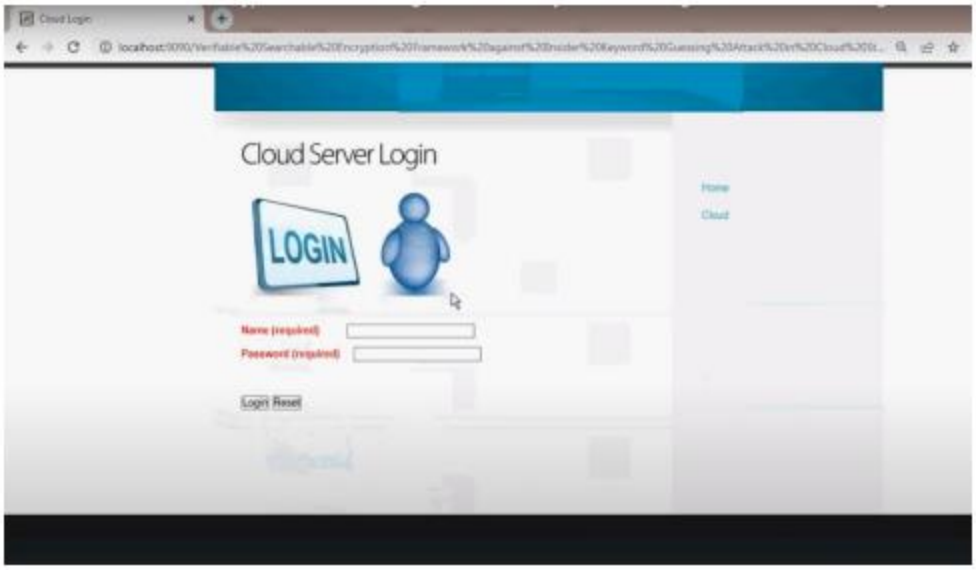


Fig 2: Cloud Server Login

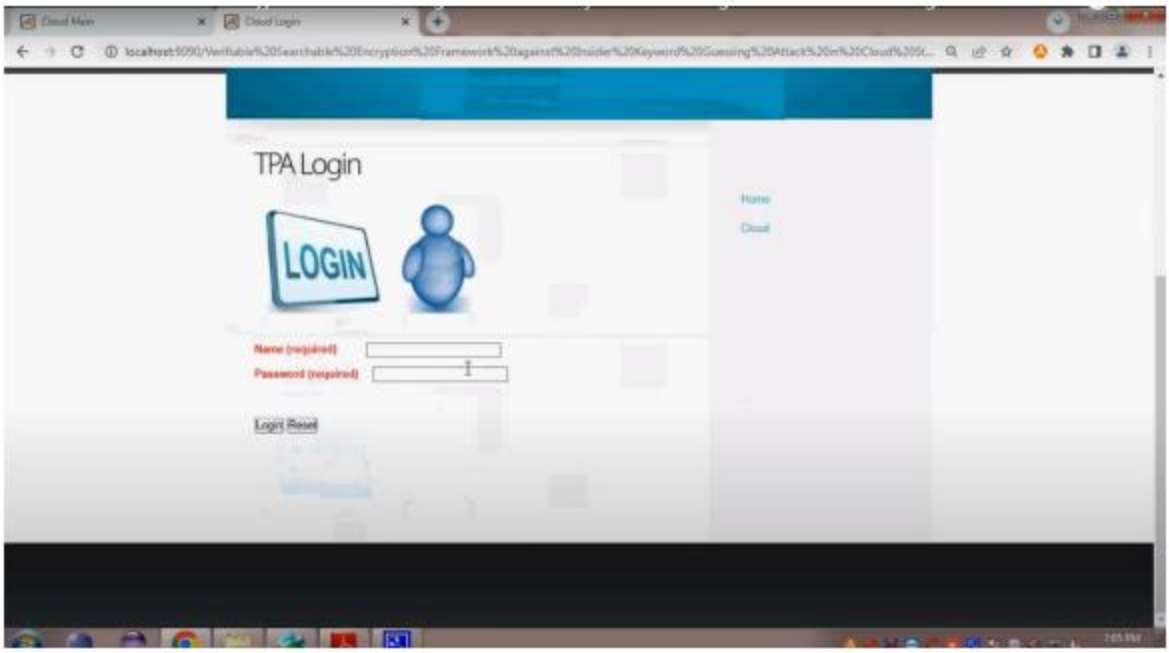


Fig 3: TPA Login

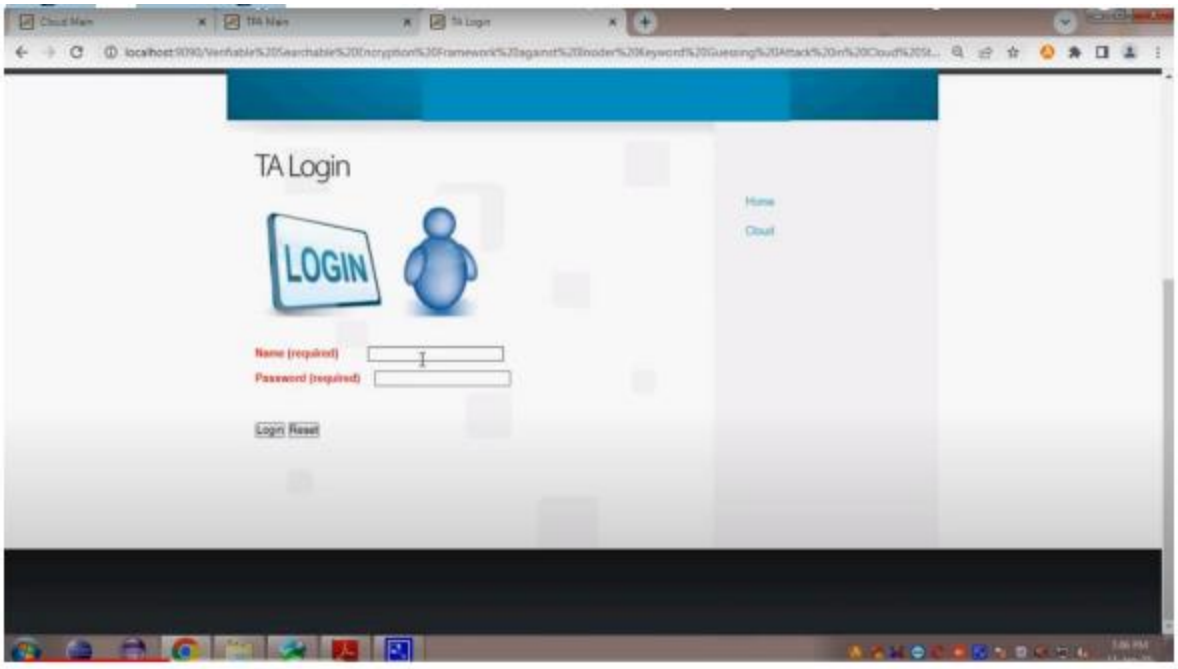


Fig 4: TA Login

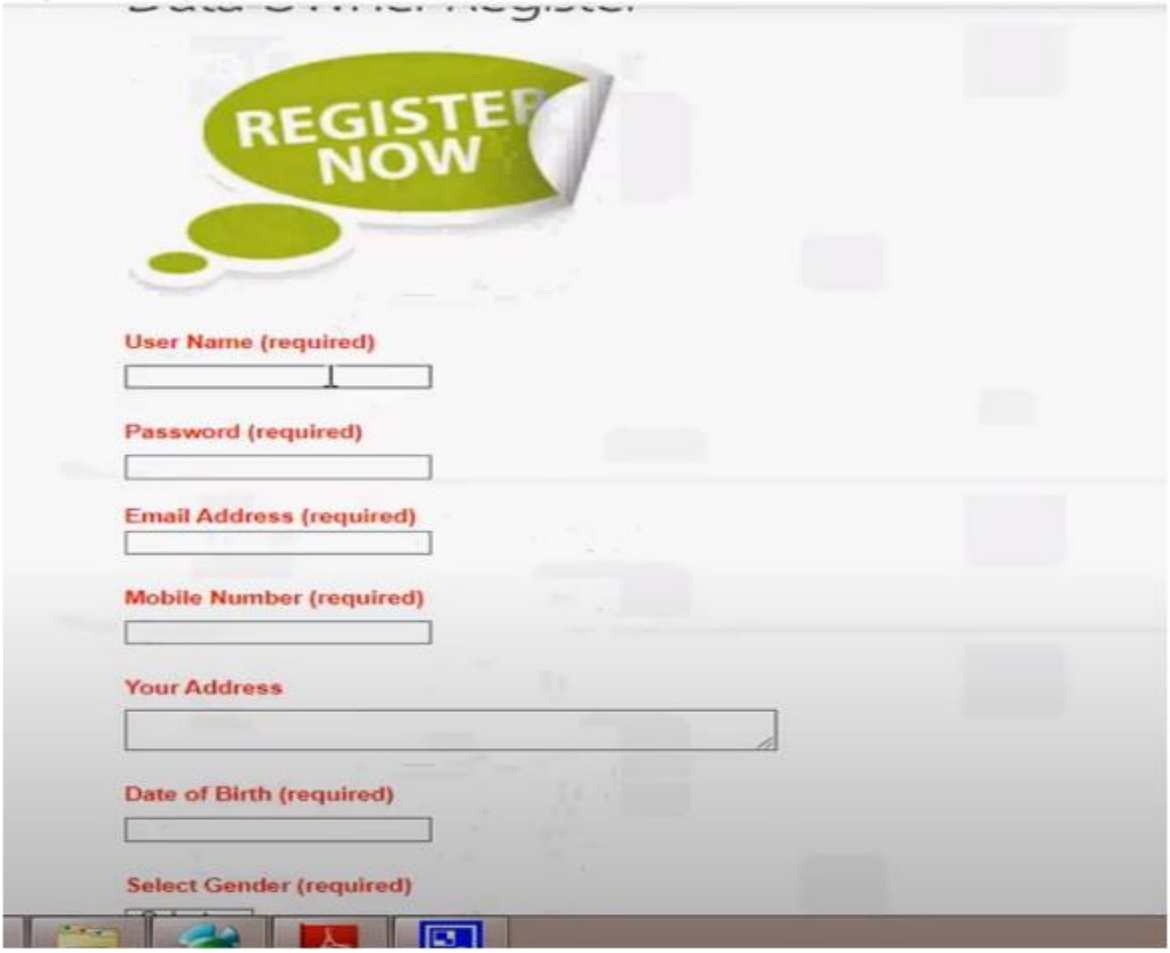


Fig 5: Register Form

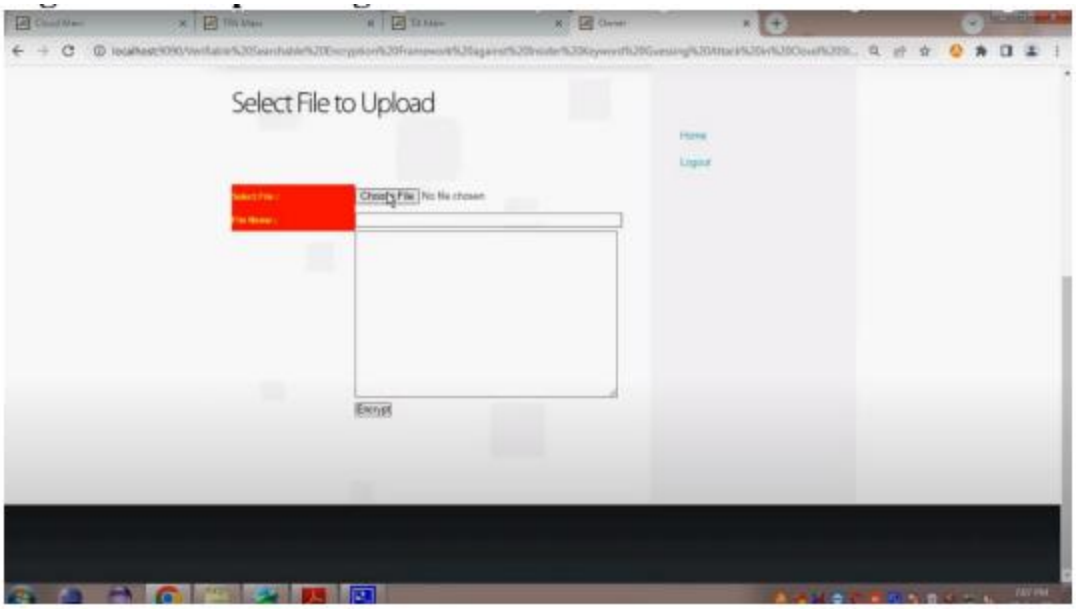


Fig 6: File Uploading

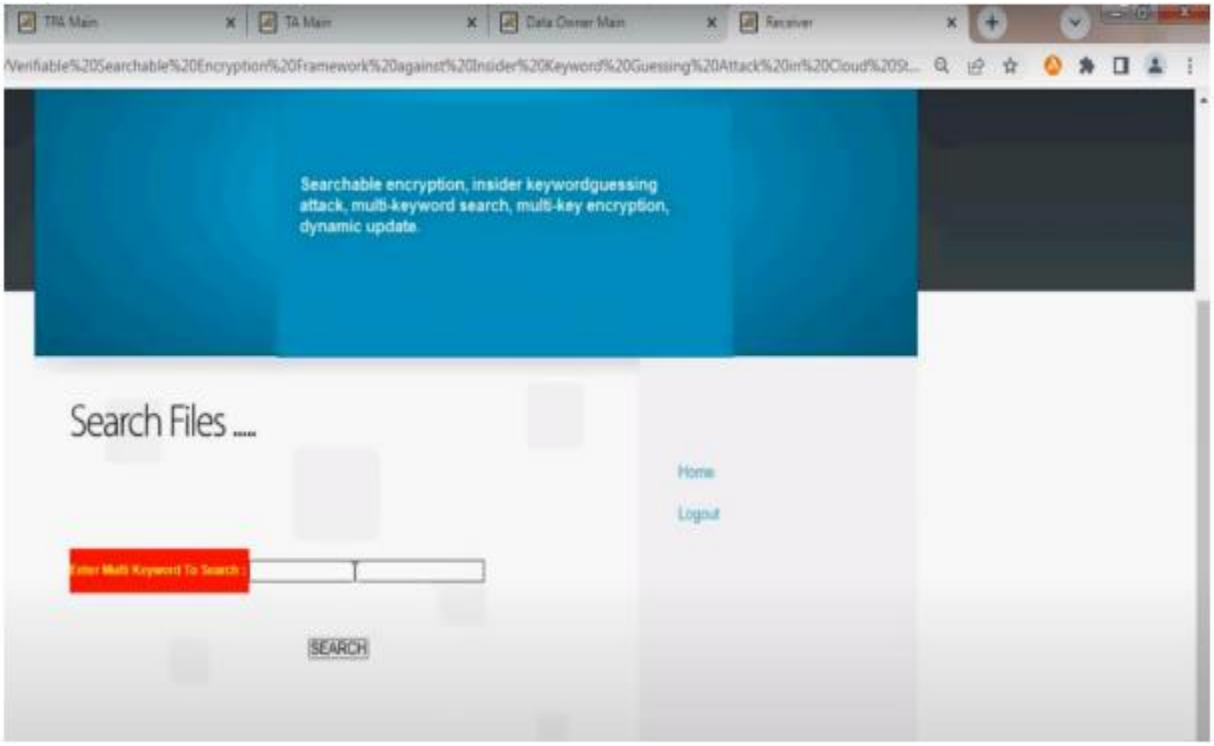


Fig 7: File Search

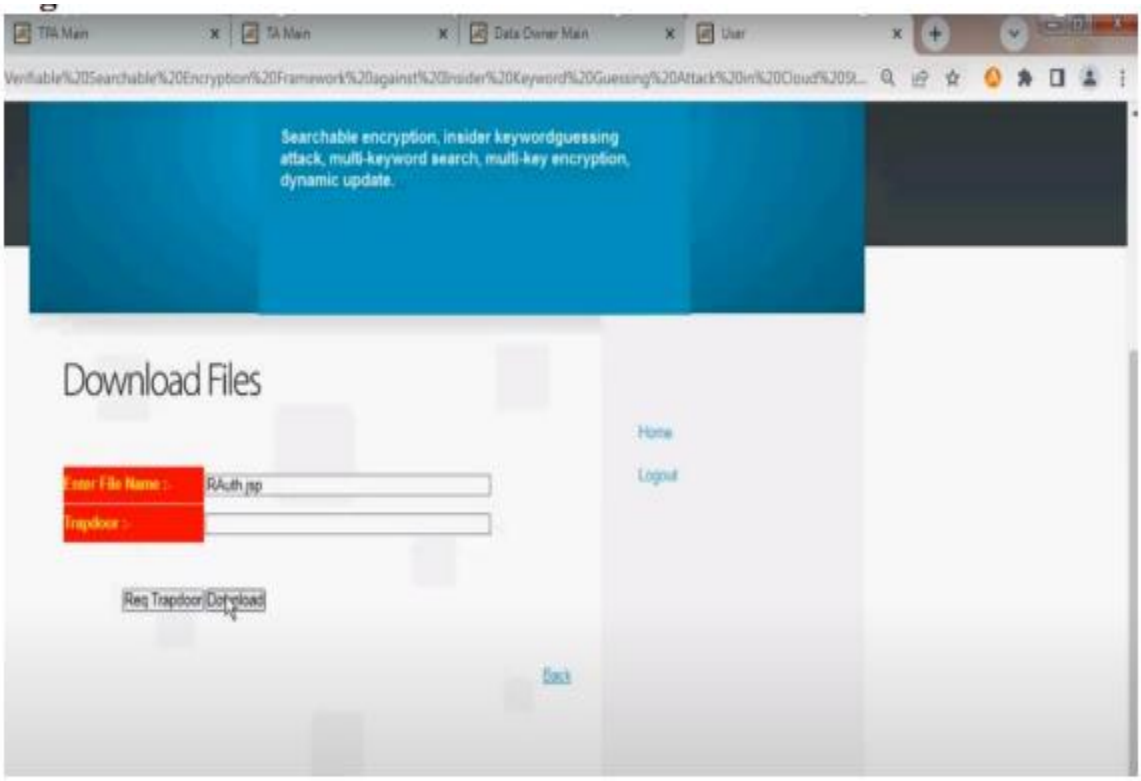


Fig 8: Download Files

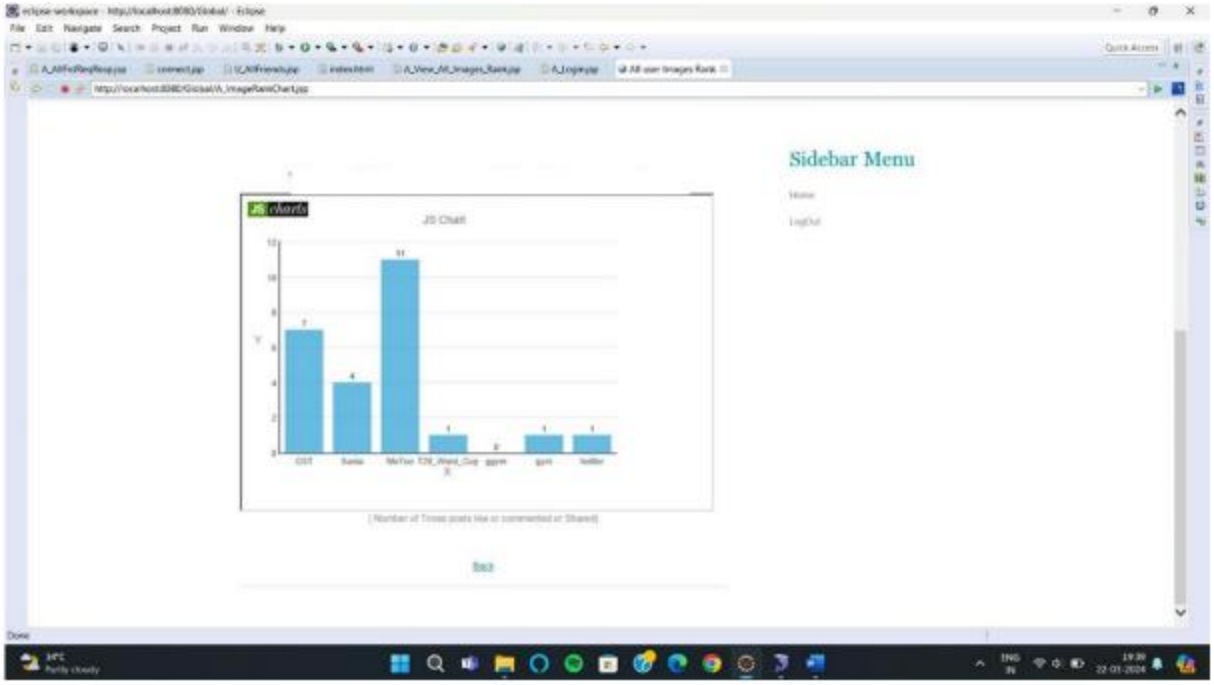


Fig 9: Rank Chart Results

Conclusion:

In this paper, a basic VSEF which mitigates the risk of inaccurate search results returned by the malicious CS and resists insider KGA was first proposed. Then, the basic VSEF was improved to support multi-keyword search, multi-key encryption and dynamic update at the same time in enhanced VKSF. We proved that basic or enhanced VSEF is semantically secure against the insider KGA, and achieves both correctness and soundness. We also evaluated the performance of the enhanced VSEF utilizing a real-world dataset to demonstrate its practicality.

Future Enhancement:

Future research concentrates on extending this work to consider a malicious DO, as well as collaborating with a real world cloud server to implement a prototype of the proposal for evaluation in a real-world setting.

BIBLIOGRAPHY

[1] Y. Zhang, C. Xu, N. Cheng, H. Li, H. Yang, and X. S. Shen, "Chronos+: An accurate blockchain-based time-stamping scheme for cloud storage," IEEE Transactions on Services Computing, 2019.

[2] Y. Zhang, C. Xu, X. Lin, and X. S. Shen, "Blockchain-based public integrity verification for cloud storage against procrastinating auditors," IEEE Transactions on Cloud Computing, 2019.

[3] G. S. Poh, J.-J. Chin, W.-C. Yau, K.-K. R. Choo, and M. S. Mohamad, "Searchable symmetric encryption: Designs and challenges," ACM Computing Surveys, vol. 50, no. 3, pp. 40:1–40:37, 2017.

[4] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. IEEE Symposium on Security and Privacy (SP'00). IEEE, 2000, pp. 44–55.

[5] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. International conference on the theory and applications of cryptographic techniques (EUROCRYPT' 04). Springer, 2004, pp. 506–522.

[6] Y. Miao, J. Ma, X. Liu, X. Li, Q. Jiang, and J. Zhang, "Attributebased keyword search over hierarchical data in cloud computing," IEEE Transactions on Services Computing, vol. PP, pp. 1–14, 2017.

[7] Y. Miao, J. Ma, X. Liu, J. Weng, H. Li, and H. Li, "Lightweight finegrained search over encrypted data in fog computing," IEEE Transactions on Services Computing, vol. PP, pp. 1–14, 2018.

[8] S. O. Baror and H. Venter, "A taxonomy for cybercrime attack in the public cloud," in Proc. International Conference on Cyber Warfare and Security. Academic Conferences International Limited, 2019, pp. 505– X.

[9] D. Bove and T. M¨uller, "Investigating characteristics of attacks on public cloud systems," in Proc. IEEE International Conference on Cyber Security and Cloud Computing (CSCloud'19). IEEE, 2019, pp. 89–94.

[10] Q. Chai and G. Gong, "Verifiable symmetric searchable encryption for semi-honest-butcurious cloud servers," in Proc. IEEE International Conference on Communications (ICC'12). IEEE, 2012, pp. 917–922.

[11] W. Sun, X. Liu, W. Lou, Y. T. Hou, and H. Li, "Catch you if you lie to me: Efficient verifiable conjunctive keyword search over large 38 dynamic encrypted cloud data," in Proc. IEEE Conference on Computer Communications (INFOCOM'15). IEEE, 2015, pp. 2110–2118.

[12] Q. Zheng, S. Xu, and G. Ateniese, "Vabks: verifiable attribute-based keyword search over outsourced encrypted data," in Proc. IEEE Conference on Computer Communications (INFOCOM'14). IEEE, 2014, pp. 522–530.

[13] Z. Chen, F. Zhang, P. Zhang, J. K. Liu, J. Huang, H. Zhao, and J. Shen, "Verifiable keyword search for secure big data-based mobile healthcare networks with fine-grained authorization control," Future Generation Computer Systems, vol. 87, pp. 712–724, 2018.

[14] Y. Miao, J. Weng, X. Liu, K.-K. R. Choo, Z. Liu, and H. Li, "Enabling verifiable multiple keywords search over encrypted cloud data," Information Sciences, vol. 465, pp. 21–37, 2018.

[15] R. Chen, Y. Mu, G. Yang, F. Guo, X. Huang, X. Wang, and Y. Wang, "Server-aided public key encryption with keyword search," IEEE Transactions on Information Forensics and Security, vol. 11, no. 12, pp. 2833– 2842, 2016.

[16] D. He, M. Ma, S. Zeadally, N. Kumar, and K. Liang, "Certificateless public key authenticated encryption with keyword search for industrial internet of things," IEEE Transactions on Industrial Informatics, vol. 14, no. 8, pp. 3618–3627, 2018.

[17] Q. Huang and H. Li, "An efficient public-key searchable encryption scheme secure against inside keyword guessing attacks," Information Sciences, vol. 403, pp. 1–14, 2017.

[18] Y. Zhang, C. Xu, J. Ni, H. Li, and X. S. Shen, "Blockchain-assisted public-key encryption with keyword search against keyword guessing attacks for cloud storage," IEEE Transactions on Cloud Computing, 2019.

[19] P. Xu, H. Jin, Q. Wu, and W. Wang, "Public-key encryption with fuzzy keyword search: A provably secure scheme under keyword guessing attack," IEEE Transactions on computers, vol. 62, no. 11, pp. 2266– 2277, 2013.

[20] L. Fang, W. Susilo, C. Ge, and J. Wang, "Public key encryption with keyword search secure against keyword guessing attacks without random oracle," Information Sciences, vol. 238, pp. 221–241, 2013.

[21] Y. Yang and M. Ma, "Conjunctive keyword search with designated tester and timing enabled proxy re-encryption function for e-health clouds." IEEE Trans. Information Forensics and Security, vol. 11, no. 4, pp. 746–759, 2016.

[22] B. Wang, B. Li, and H. Li, "Panda: Public auditing for shared data withefficient user revocation in the cloud," IEEE Transactions on services computing, vol. 8, no. 1, pp. 92– 106, 2015.

[23] Y. Miao, J. Ma, X. Liu, Q. Jiang, J. Zhang, L. Shen, and Z. Liu, "Vcksm: Verifiable conjunctive keyword search over mobile e-health cloud in shared multi-owner settings," Pervasive and Mobile Computing, vol. 40, pp. 205–219, 2017.

[24] R. A. Popa and N. Zeldovich, "Multi-key searchable encryption," IACR Cryptology ePrint Archive, vol. 2013, pp. 508–525, 2013.

[25] X. Liu, G. Yang, Y. Mu, and R. Deng, "Multi-user verifiable searchable symmetric encryption for cloud storage," IEEE Transactions on Dependable and Secure Computing, vol.PP,pp.1–12,-2018