JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

BLOCKCHAIN EXPANSION TECHNOLOGY-BASED DATA INTEGRITY AUDITING METHOD

[1]Syed Sabiha and [2]Dr. PVN Rajeswari

[1]M.Tech Research scholar, [2]Associate Professor,
Department of CSE,
Visvodaya Engineering College, Kavali, Nellore(Dt), AP, India -524201

Abstract : The growing trend of clients outsourcing data to cloud storage highlights critical concerns regarding information integrity. Blockchain technology, with its inherent decentralization and immutability, is increasingly being explored as an alternative to traditional third-party auditors. However, existing blockchain-based data integrity auditing schemes often face challenges related to the high overhead of network maintenance and the rapid expansion of blocks, which can hinder the creation of new blocks. This paper introduces a novel data integrity framework leveraging blockchain expansion technology to mitigate these issues. Our approach proposes the deployment of smart contracts by clients and Cloud Service Providers (CSPs) across a main chain and its associated sub-chains. To ensure transactional finality and efficient processing, sub-chains periodically, or as needed, submit the outcomes of their intensive computations to the main chain. Furthermore, to enhance the user experience by minimizing direct communication with the CSP during audits, a non-interactive auditing mechanism is integrated. A reward pool system is also implemented to fortify data security. The efficacy of this design is rigorously evaluated through comprehensive analysis, encompassing storage efficiency, batch evaluation capabilities, and data consistency. Experimental results obtained from the Ethereum blockchain platform demonstrate that this scheme can effectively reduce both storage capacity demands and computational overhead.

1. INTRODUCTION

The exponential growth of data and the widespread adoption of cloud computing necessitate robust solutions for data storage and management. In modern primary and backup storage systems, chunk-based deduplication is a widely employed technique to optimize storage space. This method works by storing only a single physical copy of duplicate data chunks, with all redundant chunks referencing this unique copy via small-sized pointers.

Deduplication has proven highly effective, capable of reducing storage space in primary storage by approximately 50% and in backup storage by up to 98%. This efficiency has led to its extensive adoption across various commercial distributed storage services, including Dropbox, Google Drive, Bitcasa, Mozy, and Memopal, primarily to reduce substantial storage costs. To maintain data confidentiality, encrypted deduplication integrates an encryption layer. In this scheme, each data chunk is deterministically encrypted using symmetric-key encryption, with the key often derived from the chunk's content (e.g., a cryptographic hash of the content), before being written to deduplicated storage. This allows deduplication to be applied even to encrypted chunks, as identical content will yield the same encrypted chunk, thus preserving space savings. Numerous studies have explored various encrypted deduplication schemes for effective outsourced data management in cloud environments.

Beyond storing non-duplicate data, a deduplicated storage system must also manage deduplication metadata. This metadata typically comprises two categories: a fingerprint index that tracks the fingerprints of all previously stored chunks to identify duplicates, and a file recipe that maps chunks within a file to their corresponding physical copies, enabling file reconstruction. Deduplication metadata is known to contribute significantly to storage overhead , particularly for highly repetitive workloads such as backups, where metadata storage can become dominant. In the context of encrypted deduplication, this work posits that the requirement to maintain additional key metadata, such as "key recipes" that keep track of the chunk-to-key mappings that make it possible to decrypt individual files, further exacerbates storage overhead. Key recipes must be managed independently from file recipes, encrypted using the master keys of file owners, and stored separately for each file owner because they contain sensitive key information. Such substantial metadata storage overhead can compromise the storage efficiency of encrypted deduplication in practical deployments.

This paper addresses the critical challenge of ensuring data integrity in cloud storage while mitigating the inherent overheads and scalability issues of traditional and existing blockchain-based solutions. We propose a novel blockchain-based approach that aims to offer a more efficient and scalable data integrity auditing mechanism for outsourced cloud data.

2. LITERATURE SURVEY

H. Yan, J. Li, and Y. Zhang, ‘‘Remote data checking with a designated verifier in cloud storage,’’ IEEE Syst. J., vol. 14, no. 2, pp. 1788–1797, Jun. 2020. Yan, Li, and Zhang (2020) proposed a remote data possession checking (RDPC) scheme designed for cloud storage, focusing on scenarios where a specific user is authorized to verify data integrity, rather than public or private verification alone. Their work addresses limitations in previous designated-verifier provable data possession (DV-PDP) protocols, specifically an insecurity against replay attacks from malicious cloud servers. The authors introduce a new RDPC scheme that allows a data owner to assign a unique verifier for data integrity checks. They provide a security proof based on the computational Diffie-Hellman assumption within a random oracle model. Their theoretical and experimental findings suggest that their scheme offers reduced communication, storage, and computation overhead while maintaining a high probability of error detection.

J. Chang, H. Wang, F. Wang, A. Zhang, and Y. Ji, ‘‘RKA security for identity-based signature scheme,’’ IEEE Access, vol. 8, pp. 17833–17841, 2020. Chang et al. (2020) investigated the security of identity-based signature (IBS) schemes against related-key attacks (RKA), a form of side-channel attack relevant to various cryptographic primitives. They observe that RKA-security had not been previously considered for IBS, a fundamental primitive in identity-based cryptography. Their work introduces the concept of RKA security into IBS, defining a security model that accounts for RKA occurring in either users' signing keys or the master key of the key-generation center (KGC). The authors demonstrate that an efficient Schnorr-like IBS scheme by Galindo and Garcia is vulnerable to a simple RKA. They then propose a minor modification to this scheme, resulting in an RKA-secure IBS scheme, for which they provide detailed security proof in the random oracle model. Performance analysis indicates that their modified scheme retains high efficiency while offering enhanced security.

3. PROPOSED SYSTEM

This section outlines the architecture and key components of our proposed data integrity auditing system.

Hierarchical Blockchain for Auditing: The suggested framework introduces a data integrity auditing protocol built upon plasma-based smart contracts. This protocol leverages a hierarchical blockchain structure, implementing smart contracts on both a main chain and its associated plasma sub-chains. This design significantly reduces the storage burden on the main chain and manages the rate of block growth, enhancing overall scalability and efficiency.

Efficient Auditing Mechanisms: To improve the auditing process, a batch auditing technique is proposed, enabling the simultaneous processing of multiple audit tasks. This approach minimizes computational and communication overhead during the execution of the Third-Party Auditor (TPA) protocol. Additionally, the concept of a non-interactive review is introduced to reduce the impact of communication between the client and the Cloud Service Provider (CSP) during the audit cycle, thereby improving the user experience. To ensure the reliability and accuracy of the audit, a reward pool mechanism is utilized, providing appropriate incentives for confirmation hubs.

Security Analysis and Validation: Within the proposed framework, a thorough security analysis demonstrates how the design achieves common security objectives. Furthermore, the effectiveness and viability of the scheme have been validated through multiple experimental trials conducted on the Ethereum blockchain. Important Note for Proposed System: This section describes your proposed system. While it's acceptable to build upon existing concepts (like plasma chains), the specific way you integrate them and the unique features of your protocol should be clearly articulated and differentiated from prior work. If any elements are directly inspired by or adapted from existing research, ensure proper citation.

3.1 IMPLEMENTATION

3.1.1. Data Owner

The Data Owner module is responsible for managing and uploading encrypted data to the cloud server. To ensure confidentiality, data owners encrypt their files prior to storage. This module empowers data owners to interact with their encrypted data through various operations, including registration and login, file upload, viewing uploaded files, updating existing files, and initiating data integrity auditing for file blocks.

3.1.2. Cloud

The Cloud module serves as the primary data storage provider for Data Owners. Data owners encrypt their files before storing them on the server, facilitating secure sharing with data consumers. When data consumers wish to access shared files, they download the encrypted data from the server. The mechanism for decryption needs to be clearly defined: if the server decrypts, it implies the server holds the keys, which might be a security concern; if the consumer decrypts, the server would simply provide the encrypted file. (Please clarify this decryption process: Does the server decrypt, or does the consumer receive the encrypted file and decrypt it using their own keys or keys provided by the owner?). The server also manages access authorizations, generating aggregate keys upon end-user requests for file access. Key operations managed by the Cloud include: Login, viewing and authorizing users, viewing and authorizing owners, displaying files by blockchain record, viewing all transactions, processing search requests, managing download requests, identifying potential attackers, generating file rank charts, and presenting time delay and throughput results.

3.1.3 User

The User module allows authorized individuals to interact with the stored data. Users can access data files using a designated secret key. This module facilitates searching for files based on specified keywords, with the cloud server indexing relevant data and responding to user queries. The key operations available to users include: registration and login, searching for files, downloading files, viewing accessed files, submitting search requests, and managing download requests. Important Note for User: Ensure "secret key" is consistent with your overall security model (e.g., is it a symmetric key, a private key, etc.)

3.1.4 TPA

The Third-Party Auditor (TPA) module is entrusted with specific auditing responsibilities. Its primary functions include: Login, viewing file metadata, accessing files (likely for auditing purposes without necessarily decrypting the content), and

potentially generating secret keys (Please clarify what "Generate Secret Key" means in the context of the TPA. Is it for auditing challenges, or does the TPA generate decryption keys? This is a critical security detail.). The mention of 'CPU Speed' here seems ambiguous; if it refers to monitoring the TPA's computational performance, it should be rephrased for clarity. Important Note for TPA: The role of "Generate Secret Key" for the TPA is crucial and needs precise definition to avoid security ambiguities. "CPU Speed" requires clarification or removal if it's not a direct function

4. RESULTS AND DISCUSSION

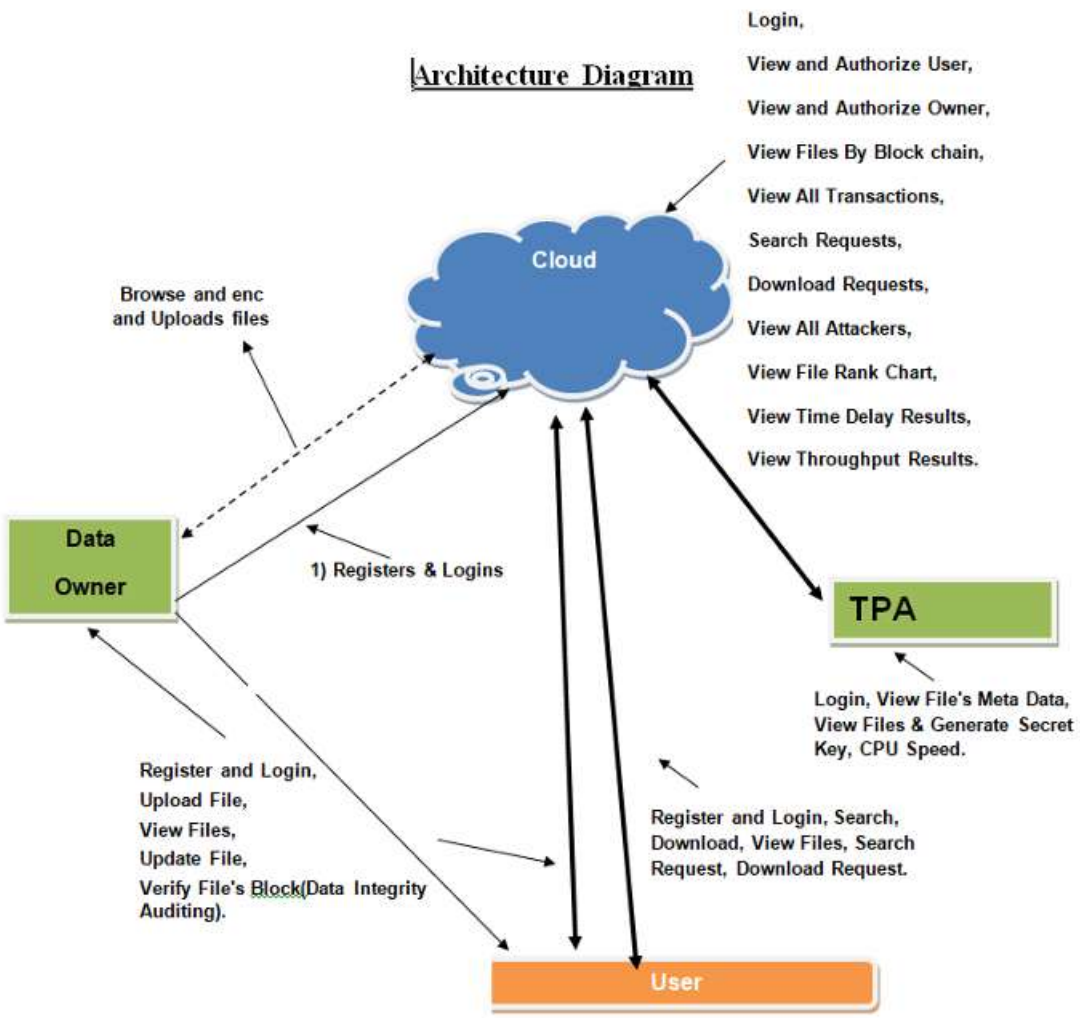

Fig 1:Architecture

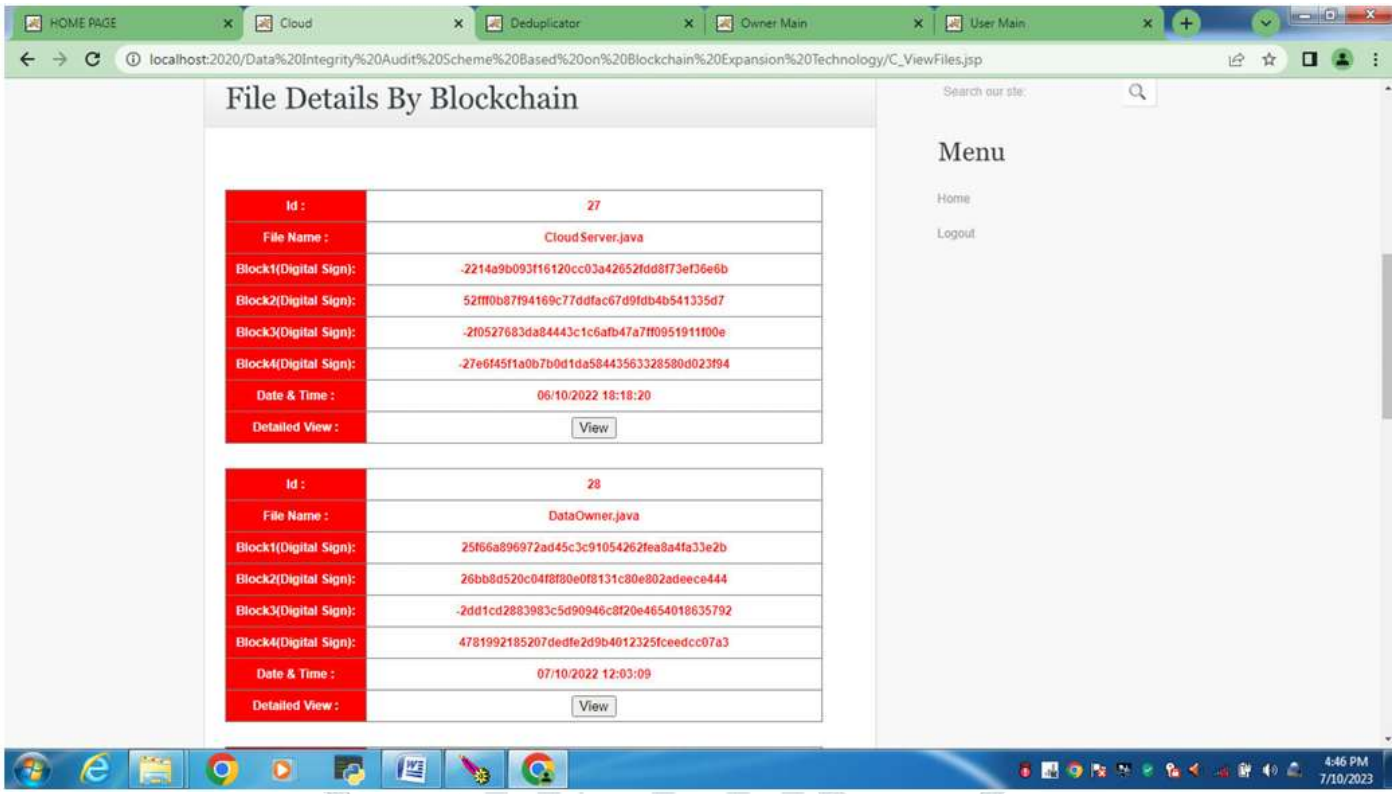

Fig 2: Uploaded data securely

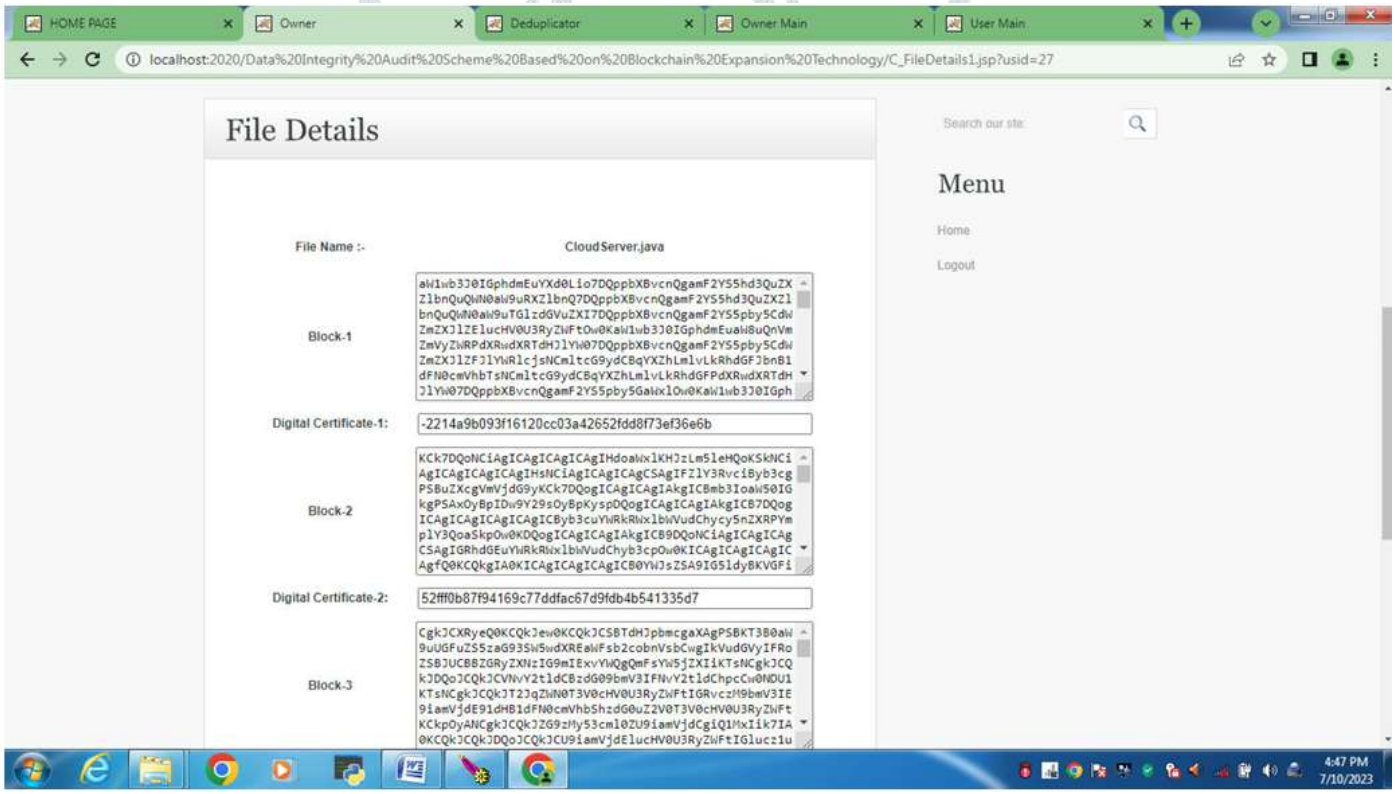

Fig 3:Encrypted data

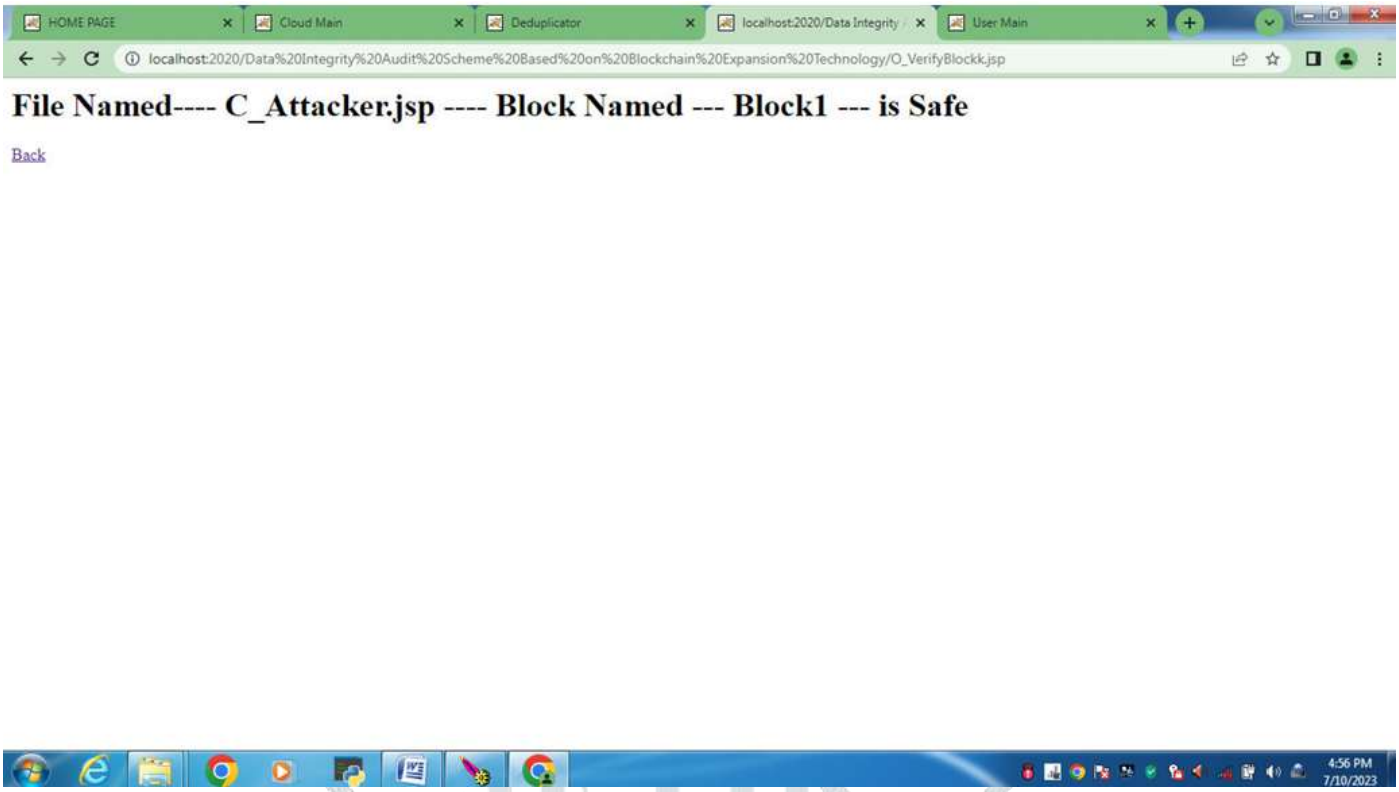

Fig 4:Data Integrity Checking

5. CONCLUSION

The continuous advancements in cloud computing and storage technologies, coupled with the escalating volume of data entrusted to cloud servers, underscore the critical importance of ensuring reliable user access and data integrity. This paper introduces a novel data integrity strategy built upon blockchain expansion technology. Our proposed scheme addresses several limitations inherent in conventional auditing methods by leveraging a blockchain network, thereby enhancing both effectiveness and security. Key to our approach is the deployment of smart contracts across both a main chain and its associated plasma sub chains. This architectural design significantly alleviates the capacity strain on the main chain, slows down its growth rate, and consequently reduces both storage and computational overhead, leading to improved Page 1 of 2 overall framework execution. Furthermore, to guarantee audit accuracy and minimize direct interaction between the smart contract platform and the Cloud Service Provider (CSP) during contract execution, we integrate a reward pool mechanism and the concept of a non-interactive audit. Through these mechanisms, our approach successfully achieves its predicted security objectives.

REFERENCES

[1] Ali, A. 2001.Macroeconomic variables as common pervasive risk factors and the empirical content of the Arbitrage Pricing Theory. Journal of Empirical finance, 5(3): 221–240.

[2] Basu, S. 1997. The Investment Performance of Common Stocks in Relation to their Price to Earnings Ratio: A Test of the Efficient Markets Hypothesis. Journal of Finance, 33(3): 663-682.

[3] Bhatti, U. and Hanif. M. 2010. Validity of Capital Assets Pricing Model.Evidence from KSE-Pakistan.European Journal of Economics, Finance and Administrative Science, 3 (20).

[1] K. Hao, J. Xin, Z. Wang, and G. Wang, ``Outsourced data integrity veri_cation based on blockchain in untrusted environment," World Wide Web, vol. 23, no. 4, pp. 2215_2238, Jul. 2020.

[2] Y. Fan, X. Lin, G. Tan, Y. Zhang, W. Dong, and J. Lei, ``One secure data integrity veri_cation scheme for cloud jespublication.com Page 659 Journal of Engineering Sciences Vol 14 Issue 07,2023 storage," Future Gener. Comput. Syst., vol. 96, pp. 376_385, Jul. 2019.

[3] H. Wang and J. Zhang, ``Blockchain based data integrity veri_cation for large-scale IoT data," IEEE Access, vol. 7, pp. 164996_165006, 2019.

[4] Z. Miao, C. Ye, P. Yang, R. Liu, B. Liu, and Y. Chen, ``A scheme for electronic evidence sharing based on blockchain and proxy re-encryption," in Proc. 4th Int. Conf. Blockchain Technol. Appl., Dec. 2021, pp. 11_16.

[5] F. Kefeng, L. Fei, Y. Haiyang, and Y. Zhen, ``A blockchain-based _exible data auditing scheme for the cloud service," Chin. J. Electron., vol. 30, no. 6, pp. 1159_1166, Nov. 2021.

[6] K. He, J. Shi, C. Huang, and X. Hu, ``Blockchain based data integrity veri_cation for cloud storage with T Merkle tree," in Proc. Int. Conf. Algo- rithms Archit. Parallel Process. Cham, Switzerland: Springer, Oct. 2020, pp. 65_80.

[7] Y. Lei, Z. Jia, Y. Yang, Y. Cheng, and J. Fu, ``A cloud data access authorization update scheme based on blockchain," in Proc. 3rd Int. Conf. Smart BlockChain (SmartBlock), Oct. 2020, pp. 33_38.

[8] Y. Yuan, J. Zhang, W. Xu, and Z. Li, ``Identity-based public data integrity veri_cation scheme in cloud storage system via blockchain," J. ISSN:0377-9254 Supercomput., vol. 78, pp. 8509_8530, Jan. 2022.

[9] S. Wang, D. Zhang, and Y. Zhang, ``Blockchain-based personal health records sharing scheme with data integrity veri_able," IEEE Access, vol. 7, pp. 102887_102901, 2019.

[10] A. Liu, Y. Wang, and X. Wang, ``Blockchain-based data-driven smart customization," in Data-Driven Engineering Design. Cham, Switzerland: Springer, 2022, pp. 89_107.

[11] K. Dhyani, J. Mishra, S. Paladhi, and I. S. Thaseen, ``A blockchain-based document veri_cation system for employers," in Proc. Int. Conf. Comput. Intell. Data Eng. Singapore: Springer, 2022, pp. 123_137.

[12] K. Xu, W. Chen, and Y. Zhang, ``Blockchain-based integrity veri_cation of data migration in multi-cloud storage," J. Phys., Conf. Ser., vol. 2132, no. 1, Dec. 2021, Art. no. 012031.

[13] G. Xu, S. Han, Y. Bai, X. Feng, and Y. Gan, ``Data tag replacement algorithm for data integrity veri_cation in cloud storage," Comput. Secur., vol. 103, Apr. 2021, Art. no. 102205.

[14] G. Xie, Y. Liu, G. Xin, and Q. Yang, ``Blockchain-based cloud data integrity veri_cation scheme with high ef_ciency," Secur. Commun. Netw., vol. 2021, pp. 1_15, Apr. 2021.

[15] U. Arjun and S. Vinay, ``Outsourced jespublication.com Page 660 Journal of Engineering Sciences Vol 14 Issue 07,2023 auditing with data integrity veri_cation scheme (OA-DIV) and dynamic operations for cloud data with multi-copies," EAI Endorsed Trans. Cloud Syst., vol. 7, no. 20, Jul. 2018, Art. no. 169423