International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

Cloud-Assisted Privacy-Preserving Spectral Clustering Algorithm Within A Multi-User Setting

T.Rama Krishna, K.Vasanth Rishi, S.Raju, S.Pragna

[1]Associate Professor, Department Of IT, Guru Nanak Institutions Technical Campus (Autonomous), India.

[2,3,4]B.Tech Students, Department Of IT, Guru Nanak Institutions Technical Campus (Autonomous), India

ABSTRACT

Spectral clustering, a powerful algorithm in the field of AI, holds a significant role despite its inherent high time complexity. For data owners grappling with limitations such as small datasets and restricted computational resources, harnessing the computational capabilities of cloud computing and aggregating data from multiple sources can yield precise spectral clustering results. However, explicit data uploading to cloud servers poses privacy risks. In response to this challenge, we explore the outsourcing dilemma of spectral clustering in a cloud and multi-user environment and propose a quantum-secure and efficient solution. Specifically, by employing the CKKS homomorphic encryption algorithm within a dual non-collusive server model, we formulate a comprehensive and multi-user spectral clustering outsourcing scheme. Our approach addresses privacy concerns by introducing secure computation protocols for L_2 norm, exponential function, and negative half power function. We elaborate on efficient computational algorithms for each stage of spectral clustering, ensuring accurate clustering outcomes without compromising dataset privacy. Moreover, in our scheme, users only need to upload their encrypted dataset without requiring direct interaction with each other or the cloud server until obtaining clustering results. Finally, we argue the IND-CPA security of our design and substantiate its accuracy and efficiency through theoretical comparison analysis and experimental evaluations.

1-INTRODUCTION

With the increasing prevalence of artificial intelligence (AI), various industries such as healthcare, ecommerce, and finance are incorporating AI into their operations. Clearly, user-friendly AI models necessitate training algorithms and large amounts of high-quality training datasets. The performance of the model is directly impacted by the quality of the training data. Clustering, as a machine learning technique for discovering patterns, relationships, and trends from large-scale datasets, can preprocess the training data for AI, providing the required high-quality training and learning data. However, for data owners with limited computational resources, clustering these large datasets can be a cumbersome task. Moreover, since most data owners often have limited storage resources, not all of them possess extensive datasets. Faced with multiple data owners, each possessing relatively small datasets, leveraging their data resources collectively becomes an effective approach to enhance clustering accuracy. Nevertheless, due to the limited data volume and insufficient computational resources of individual users, conducting large-scale data processing and clustering analysis locally is often impractical. To overcome this challenge, cloud environments offer an ideal solution by providing users with powerful computing resources and a platform for collaborative analysis [1]. In scenarios involving multiple users sharing data, uploading their data to cloud servers and utilizing the robust computational power of cloud computing can facilitate the integration and joint analysis of data from multiple sources. This collaborative clustering model holds the promise of improving the accuracy and robustness of clustering results, aiding data owners in better understanding the inherent structure and patterns within their data [2]-[6].

Nevertheless, the pressing concern of privacy leakage has spurred an urgent need for effective privacy protection solutions, particularly in the context of processing data within a cloud environment. The physical separation between data owners (DOs) and cloud servers (CS) introduces a risk of users losing control over their data. Simultaneously, potential disparities in the trusted domains of users and CS give rise to various security challenges [7][13]. On one hand, the datasets of DOs may encompass sensitive information like asset records, genetic data, and trade secrets. The unauthorized disclosure of such information can lead to immeasurable losses for users. On the other hand, unforeseen events, such as hacker intrusions and financial interests, may compel CS to passively or actively intrude upon users' sensitive data. In 2017, a significant incident took place involving Cloud flare, a prominent provider of cloud security services. During this period, user-encrypted data transmitted over HTTPS was unintentionally exposed for several months. The repercussions of this event were felt by a minimum of 2 million websites, encompassing various well-established internet companies such as Uber. Hence, it becomes paramount to explore secure clustering methods within the cloud-assisted computing paradigm to address these privacy and security challenges. Traditional clustering algorithms, exemplified by wellknown methods like k-means clustering [14]

and density based spatial clustering of applications with noise (DBSCAN) [15], are recognized for their simplicity and ease of comprehension. However, they often face challenges like poor sample tolerance and susceptibility to local optimal solutions. In response to these issues, spectral clustering emerged as a solution [16], [17], showcasing superiority over traditional clustering methods. Spectral clustering has been proven effective due to its low sensitivity to sample shape and robust support for high-dimensional data. With the evolution of big data, spectral clustering has assumed a more significant role in large-scale data clustering. Nevertheless, its drawback lies in high computational complexity. As a solution, outsourcing the intricate spectral clustering tasks to a cloud with robust computing capabilities becomes a viable option. However, directly utilizing user-uploaded sensitive datasets for spectral clustering raises privacy concerns. Hence, ensuring the secure and efficient execution of the spectral clustering algorithm in a cloud environment has become a prominent research focus and the driving motivation behind our study.

EXISTING SYSTEM

- Existing system, it becomes paramount to explore secure clustering methods within the cloud-assisted computing paradigm to address these privacy and security challenges.
- Nevertheless, due to the limited data volume and insufficient computational resources of individual users conducting large-scale data processing and clustering analysis locally is often impractical.
- To overcome this challenge, cloud environments offer an ideal solution by providing users with powerful computing resources and a platform for collaborative analysis.

PROPOSED SYSTEM

- He proposed system introduces a Cloud-Assisted Privacy-Preserving Spectral Clustering Algorithm that enables multiple users to collaboratively cluster their data without compromising individual privacy.
- We elaborate on efficient computational algorithms for each stage of spectral clustering, ensuring accurate clustering outcomes without compromising dataset privacy.
- Moreover, in our scheme, users only need to upload their encrypted dataset without requiring direct interaction with each other or the cloud server until obtaining clustering results.

2- LITERATURE SURVEY

Title: Distributed attribute based signature with attribute dynamic update for smart grid
Year: 2023
Author: Q. Su, R. Zhang, R. Xue, Y. Sun, and S. Gao

Description:
Smart grid is gaining more and more attention as one of the typical applications of Internet of Things. However, in a distributed environment, how to guarantee the privacy of users in electricity trading while ensuring the efficiency of the transactions is one of the urgent issues to be solved. In this article, we propose a distributed attribute-based signature (DABS) scheme for distributed electricity trading, which can support users' free choice of trade objects without revealing their real identities. We construct a signature generation and verification method by taking advantage of the open and hard-to-tamper properties of blockchain to achieve signature verifiability independent of dynamic changes in attributes. To improve the update efficiency, we propose an improved scheme that enables the update complexity to be reduced from O(n) to O(logn) , where n is the number of users. Finally, performance analysis and simulation experiments demonstrate the security and practicality of the DABS.

With the rapid development of Internet of Things (IoT) technology and the popularity of IoT devices [1], [2], an era of the interconnection of everything has arrived. A prominent example of the IoT is the smart grid [3], which can be seen as the next generation of the electricity grid [4]. Smart grids consist of multiple smart components that sense measurements and send this data over the network, thus enabling real-time feedback on energy consumption. The widespread use of smart grid technologies is due to the popularity of the use of heterogeneous energy sources such as renewable and nonrenewable resources and the need for simultaneous energy production and consumption. Smart grid has new features such as dynamic and transparent electricity prices, free electricity trading between energy producers and energy consumers [5], etc. Specifically, households, communities, or factories cannot only meet their own electricity needs by generating electricity by themselves, but also make profits by selling excess electricity to nearby users at a suitable price through the smart grid [6], [7].

Title: Privacy-Preserving Graph Matching Query Supporting Quick Subgraph Extraction
Year: 2023
Author: X. Ge, J.Yu, and R. Hao
Description:
Graph matching, as one of the most fundamental problems in graph database, has a wide range of applications. Due to the large scale of graph database and the hardness of graph matching, graph user tends to outsource the encrypted graphs to the cloud. The complex graph matching is performed by the cloud. Several schemes have been proposed to support graph matching query over encrypted graphs. However, none of them can realize efficient

subgraph extraction when the matched subgraph needs to be exactly located at the data graph. The graph user has to perform the complex subgraph isomorphism (NP-complete problem) operation to extract the isomorphic subgraph from the matched data graph in state-of-the-art schemes. In order to solve this problem, we propose a privacy-preserving graph matching query scheme supporting quick subgraph extraction in this paper. In our design, two non-colluding cloud servers are adopted to accomplish the matching operation jointly. Neither of them can infer the plaintexts of graphs. Two cloud servers jointly get a matched matrix to represent the matching relationship between vertices in data graph and query graph. Graph user can directly and quickly extract the subgraph isomorphic to query graph from data graph based on the matched matrix. No subgraph isomorphism operation is involved for graph user. The time complexity of subgraph extraction is O(m2) in our scheme, where m is the number of vertices in query graph. The extensive experiments with real-world database demonstrate the efficiency of the proposed privacy-preserving graph matching scheme.

With the advent of the Big Data era, massive multi-source heterogeneous data from the Internet is rapidly growing. There exist close affinities among these data. Graph, as a powerful and universal data structure, can be used to describe the relationship among these data. As one of the most fundamental problems in graph database, graph matching query plays an increasingly important role in a wide range of applications. There are total two types of graph matching queries. The first is to find all subgraphs isomorphic to a given query graph in a large-scale data graph. Typical application of this type is instance identification in Semantic Web.

Title: Secure auditing and deduplication with efficient ownership management for cloud storage
Year: 2023
Author: M.Wang, L. Xu, R. Hao, and M. Yang
Description:
More and more users opt to outsource their data to the cloud due to its popularity, which also brings some problems. One is that users cannot know whether the data stored in the cloud is intact since they no longer have physical control over it. The other is that the data stored in the cloud are often duplicated, which clearly wastes cloud storage resources. To solve the above problems, cloud auditing technology and secure deduplication technology have been widely studied. Furthermore, cloud auditing schemes supporting secure deduplication have also been proposed. However, data ownership management is rarely considered in these schemes. Once the data ownership changes, the revoked user should not be able to obtain the original ciphertext stored in the cloud to recover the original data. To solve this problem, we propose a cloud auditing scheme that supports cross-user secure deduplication with efficient data ownership management. We design a lazy update strategy to reduce update frequency and delegate the update task to the cloud, which greatly reduces the computing cost of ownership management. When the data ownership changes, the cloud does not need to update the original ciphertext immediately. It will decide whether to update the original ciphertext based on the value of a preset flag. Experimental results and the security analysis demonstrate the effectiveness and security of the proposed scheme.

With the continuous growth of outsourced data, it brings a big challenge to cloud storage services. According to the IDC, the total amount of data worldwide is expected to reach 175ZB by 2025. Actually, there is a large amount of data in the cloud is duplicated. According to the survey of EMC, the scale of the duplicate data in the cloud reaches 75%. How to eliminate duplicate data in the cloud and only store one copy of them has become an important problem for cloud service providers. In order to address this problem, data deduplication technology has been proposed. To ensure data security and privacy, data needs to be encrypted before being uploaded to the cloud. For the same data, if different users encrypt using their respective keys, it will produce different ciphertexts.

Title: Verifiable fuzzy keyword search supporting sensitive information hiding for data sharing in cloud-assisted e-healthcare systems
Year: 2023
Author: Y. Zhang, R. Hao, X. Ge, and J. Yu.
Description:
Nowadays, cloud-assisted e-healthcare systems are playing a more and more important role in intelligent medical services. By uploading Electronic Medical Records (EMRs) to the cloud, users can facilitate data sharing and decrease local data management overhead. Before EMRs are outsourced to the cloud, the sensitive information in them is usually encrypted for protecting the privacy. Though the techniques of traditional searchable encryption could achieve the retrieval of ciphertext, it results in EMRs unable to be shared with external researchers. Moreover, the existing searchable encryption schemes for cloud-assisted e-healthcare systems rarely consider fault-tolerant search and result verification. To address these issues, this paper proposes a verifiable fuzzy keyword search scheme supporting sensitive information hiding for data sharing in cloud-assisted e-healthcare systems. We encrypt sensitive information and share the rest information in EMRs among all users. We improve the approaches of keyword transformation to support the type of keyword in EMRs and realize fuzzy keyword search. For purpose of

accomplishing the public verification, we adopt the BLS signature algorithm to generate two kinds of verification tags. We also generate the auxiliary match information to realize condition search. To enhance search efficiency, we build a secure index based on the balanced binary tree. In addition, the proposed scheme can also achieve dynamic update in real-time. We analyze the security and implement a series of experiments for evaluating the effectiveness of this scheme. The experimental results illustrate our scheme could reach high accuracy.
However, local e-healthcare systems are difficult to meet the requirements of growing storage and data sharing due to the problem of storage limitation and authority management. The cloud computing can offer enormous storage space, strong computing capability, and low-cost on-demand services. Therefore, the hospital uploads EMRs to the cloud for storage, which promotes the rise of cloud-assisted e-healthcare systems.

Title: Ppadt: Privacy-preserving identity based public auditing with efficient data transfer for cloud-based iot data.
Year: 2023
Author: C. Gai, W. Shen, M. Yang, and J. Yu.
Description: Public auditing is a significant technique in cloud-based Internet of Things (IoT) systems, which enables the verifier to check the integrity of IoT data stored in the cloud. Nowadays, data become a core property for owners. Once the data of one owner are sold to another one, the ownership of these data has to be transferred. However, the existing public auditing schemes with data transfer require all the authenticators corresponding to the transferred data to be transformed to the new ones for integrity auditing. It incurs significant computation cost because of recomputing the new authenticators for all transferred data, especially when a vast quantity of data is being transferred. In addition, the data privacy and the identity privacy of the data owner cannot be protected for the verifier in such schemes. Thus, how to achieve efficient data transfer and privacy protection are key challenges in public auditing with data transfer for cloud-based IoT data. In this article, we propose a privacy-preserving identity-based public auditing scheme with efficient data transfer for cloud-based IoT data (PPADT). In PPADT, not all the authenticators corresponding to the transferred data blocks need to be transformed. We only need to transform an aggregated authenticator in the integrity-auditing phase. It means that the computation cost of data transfer is independent of the number of transferred data blocks. Furthermore, the data owner's identity privacy can be ensured with the assistance of the private key generator. The data privacy can also be guaranteed by employing the random masking technique.

owever, once the IoT data is uploaded to the cloud, data owners will lose the physical control of their IoT data. It means that there are several potential factors that may cause the IoT data corruption or loss, such as software/hardware breakdown, the staff's wrong operation, hacker attacks, and the cloud's malicious deletion. To ensure the cloud correctly stores the IoT data, many public auditing schemes are proposed. In these public auditing schemes, the file is divided into plenty of data blocks. Before uploading the collected data to the CS, the data owner needs to utilize his private key to produce the authenticator for each data block.

Title: Blockchain-Aided Privacy-Preserving Outsourcing Algorithms of Bilinear Pairings for Internet of Things Devices.
Year: 2021
Author: H. Zhang, L. Tong, J. Yu, and J. Lin.
Description:
Bilinear pairing is a fundamental operation that is widely used in cryptographic algorithms (e.g., identity-based cryptographic algorithms) to secure IoT applications. Nonetheless, the time complexity of bilinear pairing is O(n3) , making it a very time-consuming operation, especially for resource-constrained IoT devices. Secure outsourcing of bilinear pairing has been studied in recent years to enable computationally weak devices to securely outsource the bilinear pairing to untrustworthy cloud servers. However, the state-of-art algorithms often require to precompute and store some values, which results in storage burden for devices. In the Internet of Things, devices are generally with very limited storage capacity. Thus, the existing algorithms do not fit the IoT well. In this article, we propose a secure outsourcing algorithm of bilinear pairings, which does not require precomputations. In the proposed algorithm, the outsourcer side's efficiency is significantly improved compared with executing the original bilinear pairing operation. At the same time, the privacy of the input and output is ensured. Also, we apply the Ethereum blockchain in our outsourcing algorithm to enable fair payments, which ensures that the cloud server gets paid only when he correctly accomplished the outsourced work. The theoretical analysis and experimental results show that the proposed algorithm is efficient and secure.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

Title: Efficient Identity-Based Data Integrity Auditing with Key-Exposure Resistance for Cloud Storage.
Year: 2022
Author: W. Shen, J. Yu, M. Yang, and J. Hu
Description:
The key exposure is a serious threat for the security of data integrity auditing. Once the user's private key for auditing is exposed, most of the existing data integrity auditing schemes would inevitably become unable to work. To deal with this problem, we construct a novel and efficient identity-based data integrity auditing scheme with key-exposure resilience for cloud storage. This is achieved by designing a novel key update technique, which is fully compatible with BLS signature used in identity-based data integrity auditing. In our design, the Third Party Auditor (TPA) is responsible for generating update information. The user can update his private key based on the private key in one previous time period and the update information from the TPA. Furthermore, the proposed scheme supports real lazy update, which greatly improves the efficiency and the feasibility of key update. Meanwhile, the proposed scheme relies on identity-based cryptography, which makes certificate management easy. The security proof and the performance analysis demonstrate that the proposed scheme achieves desirable security and efficiency.
In most of existing data integrity aiding shames, a pair of public key is used to verify the validity of the proof generated by the cloud. The private key is only utilized to calculate the authenticators for data blocks. The authenticators are used to verify whether the cloud correctly stores the user's data in the phase of data blocks. The authenticators are used to verify whether the cloud correctly stores the user's data in the phase of data auditing.

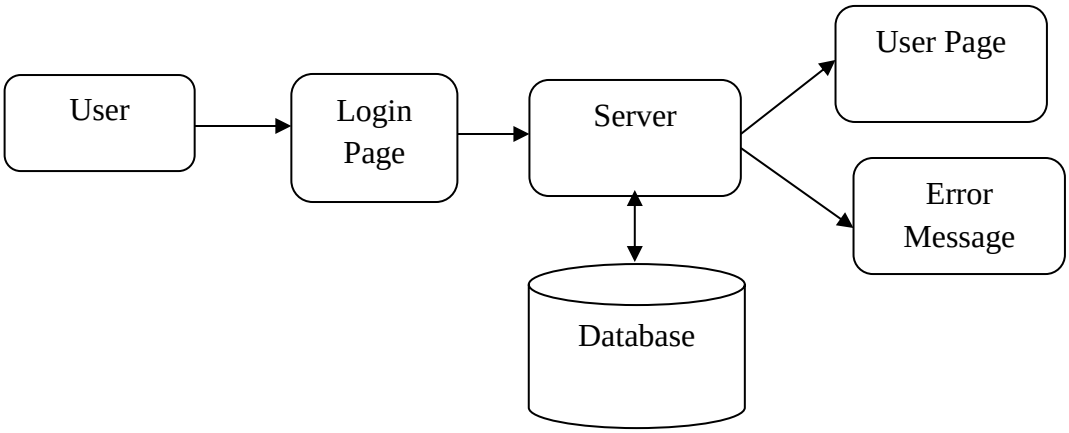
3-PROJECT DESCRIPTION

In our outsourcing system, DOs and KGC are generally trusted, the threats we considered are mainly from the untrusted cloud servers. Conforming to the potential threat behaviors in the real world, we assume that the cloud servers in our system are honest but curious. That is, to earn legitimate economic profit, they will perform the assigned task honestly while, to get extra benefits, they may be curious about the data of DOs and try to capture valuable information. Additionally, due to the conflict of interest between different cloud service providers, we assume that CS1 and CS2 do not collude with each other.

MODULES EXPLANATION AND DIAGRAM

- **User Interface Design**

To connect with server user must give their username and password then only they can able to connect the server. If the user already exits directly can login into the server else, user must register their details such as username, password, Email id, City and Country into the server. Database will create the account for the entire user to maintain upload and download rate. Name will be set as user id. Logging in is usually used to enter a specific page. It will search the query and display the query.



International Journal of Information Technology & Computer Engineering

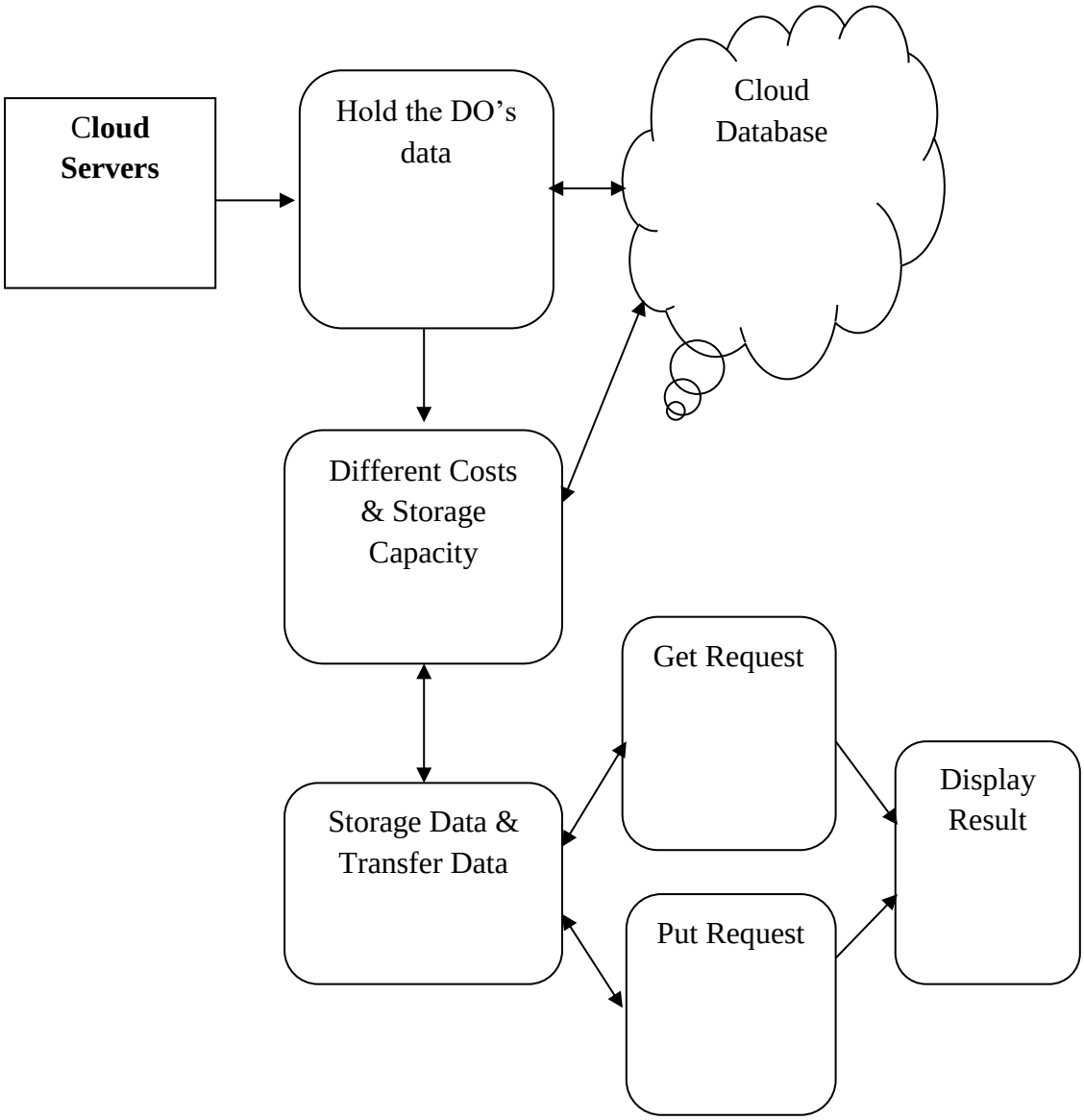
ISSN 2347–3657

Volume 13, Issue 2, 2025

Data Owners (DOs):
Data Owners (DOs): This party consists of m data owners DO_1, DO_2, … ,DO_m. Each DO owns some yet not enough data samples to obtain accurate clustering results through these samples. Also, each DO itself suffers from limited computing and storage resources and cannot afford complex clustering algorithms. Therefore, these DOs hope to obtain more accurate clustering results by jointly clustering their samples on resource-abundant cloud servers without leaking their respect data privacy.

C**loud Servers (CSs)**

Cloud Servers (CSs): This party consists of two servers CS1 and CS2, both of which have powerful computing and storage resources. Its role is to help DOs with limited data samples and resources obtain accurate clustering results. Intermediate values privacy means that the intermediate values incurred during the interactions between CS1 and CS2 should not leak the characteristic information of the data points in the original datasets. Efficiency means that, under the premise of ensuring security, the design should reduce the time consumption of each participant as much as possible.



International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

4-ALGORITHM

Cloud-Assisted Privacy-Preserving Spectral Clustering Algorithm:

Spectral clustering is an algorithm for clustering based on the principles of spectral graph partitioning theory. The main idea is to transform the clustering problem into a graph partitioning problem. It has the characteristics of low sensitivity to sample shape, convergence to the global optimal solution, and good support for high-dimensional data. Detailedly, given a data set $D=\{(i,a_i) \mid i=1,\cdots, n\}$ and the number of clusters *k*. Let $A=[a_1,\cdots,a_n]$. Then, the spectral clustering goes as follows:

Step1: Construct an affinity matrix W that can describe the characteristics of each two samples. That is,

$$\boldsymbol{W}[i,j] = w_{ij} = \boldsymbol{s}(\boldsymbol{a}_i, \boldsymbol{a}_j),$$

where $s(a_i,a_j)$ is the similarity function, representing the similarity between ai and aj . Commonly used similarity functions include Euclidean distance, cosine similarity and Gaussian kernel function. Among them, the Gaussian kernel function

$$\boldsymbol{s}(\boldsymbol{a}_i, \boldsymbol{a}_j) = \boldsymbol{e}^{-\frac{\|\boldsymbol{a}_i - \boldsymbol{a}_j\|_2^2}{2\sigma^2}}$$

Step2: Calculate the normalized Laplacian matrix

$$\boldsymbol{L} = \boldsymbol{D}^{-\frac{1}{2}}(\boldsymbol{D} - \boldsymbol{W})\boldsymbol{D}^{-\frac{1}{2}},$$

where $D=\mathrm{diag}\{d_1,d_2,\cdots,d_n\}$ is a diagonal matrix with the diagonal elements

$$d_i = \sum_{j=1}^{n} w_{ij}, i = 1, \cdots, n .$$

Step3: Calculate the eigenvalues σ1,σ2,⋯,σn and eigenvectors u1,u2,⋯,un of the matrix L . Take the eigenvectors uj1,uj2,⋯,ujk corresponding to the first *k* smallest eigenvalues after sorting and arrange them as column vectors to form a new solution space Y .

Step4. Use the classic *k*-means clustering algorithm to cluster the new solution space and finally map the clustering results back to the original solution space A . Specifically, let $yi \in Rk$ be the *i*-th row vector of Y . Then cluster the points $\{(i,yi) \mid i=1,\cdots,n\}$ with *k*-means clustering algorithm into *k* clusters $C'1,\cdots,C'k$. Finally, the cluster result is $\{C1,C2,\cdots,Ck\}$ with

The Fully Homomorphic Encryption Scheme CKKS

The celebrated CKKS encryption scheme is proposed by Cheon et al. [46] in 2017. Before reviewing the CKKS, we introduce some necessary notations. For a power-of-two integer *N*, R=Z[X]/(XN+1) denotes the ring of integers of the (2N) -th cyclotomic field, and RqL=R/(qLR) is the residue ring of R modulo an integer qL . For a real σ>0 , let DG(σ2) denote the discrete distribution over ZN with mean value 0 and variance σ2 . For a positive integer *h*, let HWT(h) be the uniform distribution over the set {s:s∈{0,1,−1}N∧the hamming weight of s is h} . For a real 0≤κ≤1 , ZO(κ) refers to the distribution over the set {0,1,−1}N that draws each entry in the vector with probability κ/2 for each of -1 and +1, and probability 1−κ for 0.

Precisely, the CKKS is a four-tuple (Setup(1λ) , Keygen(1λ) , Encpk(m) , Decsk(ct)) defined as follows.

- Setup(1λ) . Given a security parameter λ , choose a power-of-two integer *N*, an integer *h* and an integer *P* For a base integer *p* and the number of levels *L*, set the chain of ciphertext moduli qℓ=pℓ for 1≤ℓ≤L .
- Keygen(1λ) .
  Sample s←HWT(h) , a←U(RqL) , a′←U(RP·q L) , e←DG(σ2) , and e′←DG(σ2) , and set the secret key sk←(1,s) , the public key pk←(b,a)∈R2qL for b=−a·s+e(modqL) , and the evaluation key evk←(b′,a′)∈R2P·qL for b′=−a′·s+e′+Ps2(m odP·qL) .
- Encpk(m) .
  Sample r←ZO(0.5) and e0,e1←DG(σ2) .
  Output the ciphertext ct=r·pk+(m+e0,e1) (modqL) .
- Decsk(ct) . For an input ciphertext ct=(c0,c1) of level ℓ , recover the plaintext c0+c1·s

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

5-REQUIREMENTS ENGINEERING

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis, the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

- ECONOMICAL FEASIBILITY
- TECHNICAL FEASIBILITY
- SOCIAL FEASIBILITY

ECONOMICAL FEASIBILITY: This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus, the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

TECHNICAL FEASIBILITY: This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

SOCIAL FEASIBILITY: The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

Activity Diagram

These are the requirements for doing the project. Without using these tools & software's we cannot do the project. Therefore, we have two requirements to do the project.

FUNCTIONAL REQUIREMENTS

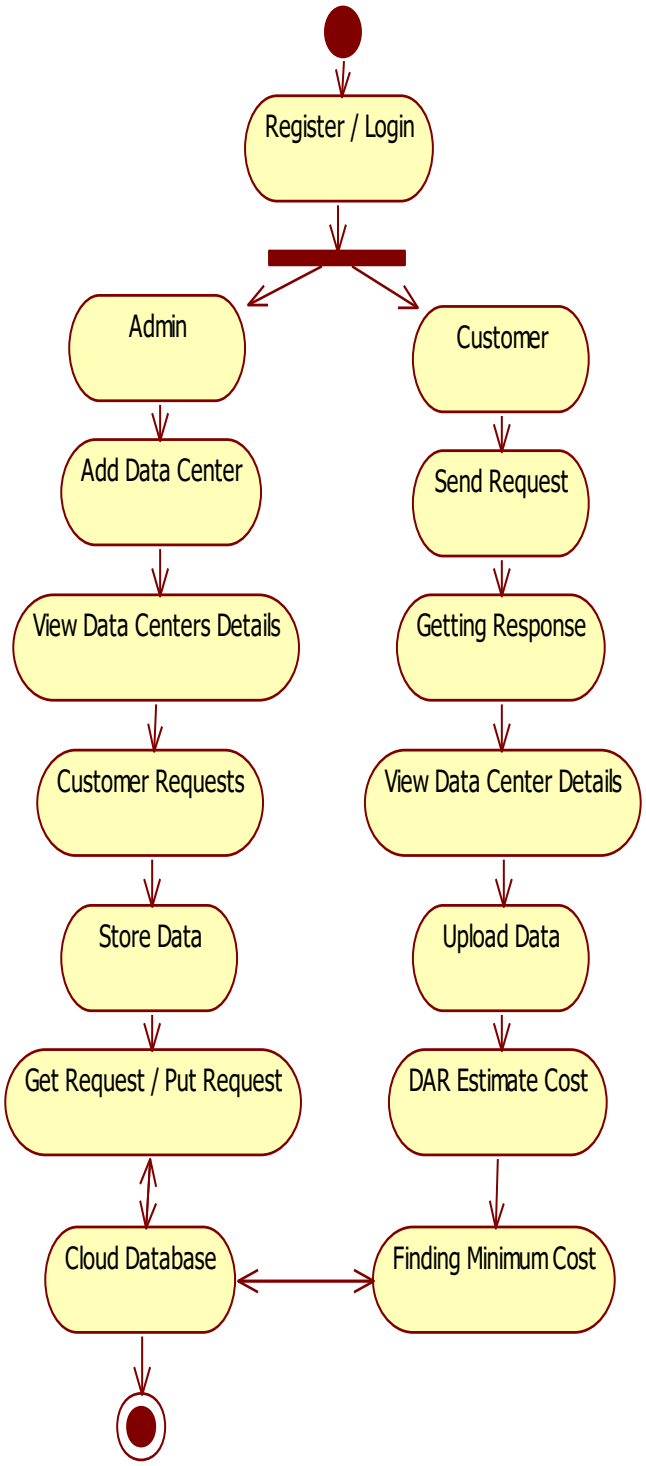
A functional requirement defines a function of a software-system or its component. A function is described as a set of inputs, the behaviour, and outputs. The outsourced computation is data is more secured. Briefly, our primary goal is to design a correct, secure and efficient spectral clustering outsourcing protocol. Now, we illustrate the detailed objectives item by item.

- **Correctness:** Correctness means that, if the cloud servers perform the assigned task in the protocol honestly, Dos should obtain the same clustering result as that obtained with the spectral clustering calculated on the plaintext dataset by themselves.
- **Dataset Privacy:** Dataset privacy means that, even if the cloud servers are curious, the data in the dataset of each DO should be CPA-secure. That is, even the clouds can adaptively collect several plaintext-ciphertext data pairs, they still can not distinguish the two challenge ciphertexts.
- **Intermediate Values Privacy:** Intermediate values privacy means that the intermediate values incurred during the interactions between CS1 and CS2 should not leak the characteristic information of the data points in the original datasets.
- **Efficiency:** Efficiency means that, under the premise of ensuring security, the design should reduce the time consumption of each participant as much as possible.

6-DESIGN ENGINEERING

Design Engineering deals with the various UML [Unified Modelling language] diagrams for the implementation of project. Design is a meaningful engineering representation of a thing that is to be built. Software design is a process through which the requirements are translated into representation of the software. Design is the place where quality is rendered in software engineering. Design is the means to accurately translate customer requirements into finished product.

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 2, 2025



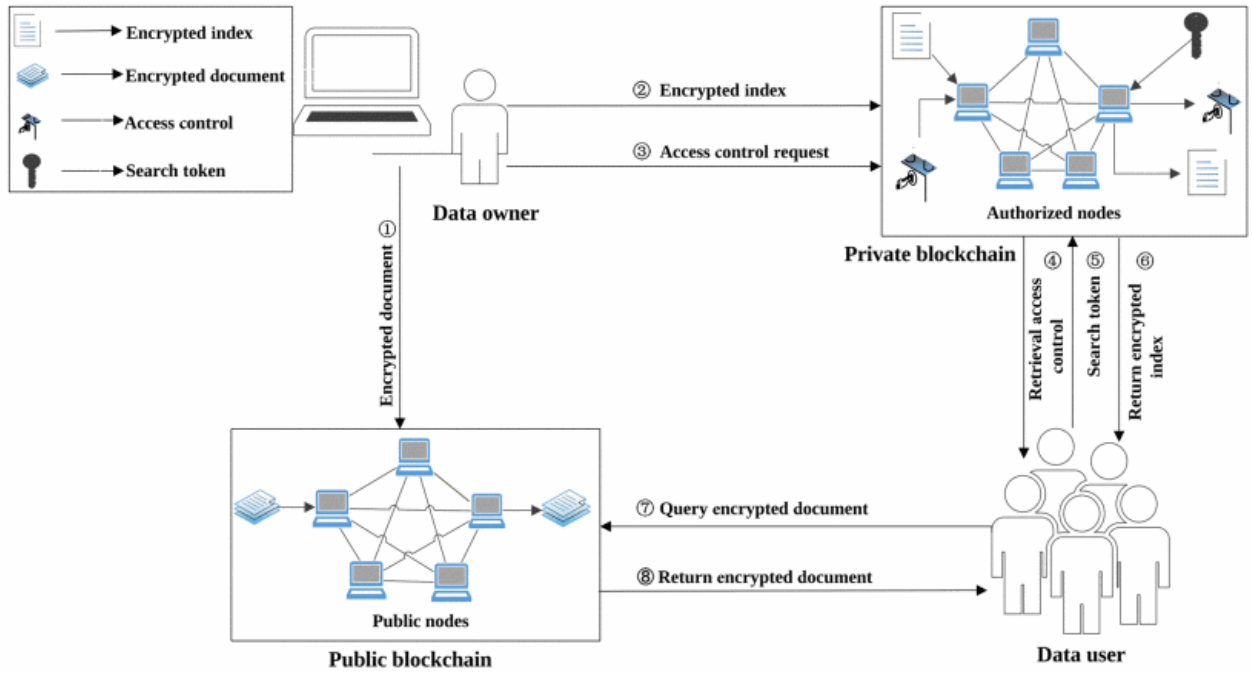
Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 2, 2025

System Architecture



System Architecture Model

7-CONCLUSION

In this paper, we introduce a secure spectral clustering scheme tailored for multiple users in a cloud environment, utilizing a non-colluding two-server model to ensure the privacy of both datasets and intermediate values. Our approach leverages the lattice-based homomorphic CKKS scheme, enhancing resistance to quantum attacks. We have substantiated the accuracy and efficiency of our scheme through theoretical analysis and extensive experimental results. Looking ahead, our future work will explore the security of our scheme under a malicious server model and aim to enhance its verifiability. Meanwhile, we will consider more efficient and secure models, such as the single-server model.

REFERENCE

[1] ``The NIST Definition of Cloud Computing," StandardSP800-145, National Institute of Science and Technology, 2011.

[2] Q. Su, R. Zhang, R. Xue, Y. Sun, and S. Gao, ``Distributed attribute based signature with attribute dynamic update for smart grid," IEEE Trans. Industr. Inform., vol. 19, no. 9, pp. 94249435, 2023.

[3] X. Ge, J.Yu, and R. Hao, ``Privacy-Preserving Graph Matching Query Supporting Quick Subgraph Extraction," IEEE Trans. Dependable Security Comput., pp. 115, 2023.

[4] M. Wang, L. Xu, R. Hao, and M. Yang, ``Secure auditing and deduplication with efficient ownership management for cloud storage," J. Syst. Archit., vol. 142, p. 102953, 2023.

[5] Y. Zhang, R. Hao, X. Ge, and J. Yu, ``Verifiable fuzzy keyword search supporting sensitive information hiding for data sharing in cloud-assisted e-healthcare systems," J. Syst. Archit., vol. 142, p. 102940, 2023.

[6] C. Gai, W. Shen, M. Yang, and J. Yu, ``Ppadt: Privacy-preserving identity based public auditing with efficient data transfer for cloud-based iot data," IEEE Internet Things J., pp. 11, 2023.

[7] M. Ali, S. U. Khan, and A. V. Vasilakos, ``Security in cloud computing: Opportunities and challenges," Inf. Sci., vol. 305, pp. 357383, Jun. 2015.

[8] S. Singh, Y.-S. Jeong, and J. H. Park, ``A survey on cloud computing security: Issues, threats, and solutions," J. Netw. Comput. Appl., vol. 75, pp. 200222, Nov. 2016.

[9] H. Zhang, L. Tong, J. Yu, and J. Lin, ``Blockchain-Aided Privacy- Preserving Outsourcing Algorithms of Bilinear Pairings for Internet of Things Devices," IEEE Internet Things J., vol. 8, no. 20, pp. 15 596 15 607, Oct. 2021.

[10] W. Shen, J. Yu, M. Yang, and J. Hu, ``Efficient Identity-Based Data Integrity Auditing with Key-Exposure Resistance for Cloud Storage," IEEE Trans. Dependable Security Comput., pp. 115, 2022.

[11] ``How to securely outsource the extended euclidean algorithm for largescale polynomials over finite fields," Information Sciences, vol. 512, pp. 641 660, 2020.
[12] Y. Huang, W. Shen, J. Qin, and H. Hou, ``Privacy-preserving certificate less public auditing supporting different auditing frequencies," Comput. Secur., vol. 128, p. 103181, May 2023.
[13] Y. He, D. Yan, and F. Chen, ``Hierarchical federated learning with local model embedding," Eng. Appl. Artif. Intell., vol. 123, p. 106148, 2023.
[14] J. MacQueen et al., ``Some methods for classification and analysis of multivariate observations," in Proc. 5th Berkeley Symp. Math. Statist. Probab., vol. 1, Oakland, CA, USA, 1967, pp. 281297.
[15] M. Ester, H.-P. Kriegel, J. Sander, X. Xu et al., ``A density-based algorithm for discovering clusters in large spatial databases with noise," in Proc. 2nd Int. Conf. Knowl. Discovery Data Mining, Aug. 1996, pp. 226231.
[16] A. Ng, M. Jordan, and Y. Weiss, ``On Spectral Clustering: Analysis and an algorithm," in Adv Neural. Inf. Process. Syst., vol. 14. MIT Press, 2001.
[17] U. Von Luxburg, ``A tutorial on spectral clustering," Stat. Comput., vol. 17, no. 4, pp. 395416, Dec. 2007.
[18] G. A. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren, ``Secure, privacy-preserving and federated machine learning in medical imaging," Nature Mach. Intell., vol. 2, no. 6, pp. 305311, Jun. 2020.
[19] C. Tian, J. Yu, H. Zhang, H. Xue, C. Wang, and K. Ren, ``Novel secure outsourcing of modular inversion for arbitrary and variable modulus," IEEE Trans. Serv. Comput., vol. 15, no. 1, pp. 241253, 2022.
[20] Y. Shao, C. Tian, L. Han, H. Xian, and J. Yu, ``Privacy-preserving and verifiable cloud-aided disease diagnosis and prediction with hyperplane decision-based classier," IEEE Internet Things J., vol. 9, no. 21, pp. 21 64821 661, 2022.
[21] P. Bunn and R. Ostrovsky, ``Secure two-party k-means clustering," in Proc. 14th ACM Conf. Comput. Commun. Secur., 2007, pp. 486497.
[22] Y. Li, Z. Hao, W. Wen, and G. Xie, ``Research on differential privacy preserving k-means clustering," Comput. Sci., vol. 40, no. 3, pp. 287290, 2013.
[23] D. Liu, E. Bertino, and X. Yi, ``Privacy of outsourced k-means clustering," in Proc. 9th ACM Symp. Inf., Comput. Commun. Secur., 2014, pp. 123134.
[24] F.-Y. Rao, B. K. Samanthula, E. Bertino, X. Yi, and D. Liu, ``Privacy preserving and outsourced multi-user k-means clustering," in Proc. IEEE Conf. Collaboration Internet Comput. IEEE, 2015, pp. 8089.
[25] J. Ren, J. Xiong, Z. Yao, R. Ma, and M. Lin, ``Dplk-means: A novel differential privacy k-means mechanism," in Proc. IEEE 2nd Int. Conf. Data Sci. Cyberspace. IEEE, 2017, pp. 133139.
[26] Y. Zhang, N. Liu, and S. Wang, ``A differential privacy protecting k-means clustering algorithm based on contour coefficients," PLoS One, vol. 13, no. 11, p. e0206832, 2018.
[27] A. Theodouli, K. A. Draziotis, and A. Gounaris, ``Implementing private k-means clustering using a lwe-based cryptosystem," in Proc. IEEE Symp. Comput. Commun. IEEE, 2017, pp. 8893.
[28] H.-J. Kim and J.-W. Chang, ``A privacy-preserving k-means clustering algorithm using secure comparison protocol and density-based center point selection," in Proc. IEEE 11th Int. Conf. Cloud Comput. IEEE, 2018, pp. 928931.
[29] K. Xing, C. Hu, J. Yu, X. Cheng, and F. Zhang, ``Mutual privacy preserving k-means clustering in social participatory sensing," IEEE Trans. Ind. Informat., vol. 13, no. 4, pp. 20662076, 2017.
[30] G. Sakellariou and A. Gounaris, ``Homomorphically encrypted k-means on cloud-hosted servers with low client-side load," Computing, vol. 101, pp. 18131836, 2019.
[31] W. Wu, J. Liu, H. Wang, J. Hao, and M. Xian, ``Secure and efficient outsourced k-means clustering using fully homomorphic encryption with ciphertext packing technique," IEEE Trans. Knowl. Data Eng., vol. 33, no. 10, pp. 34243437, 2020.
[32] Y. Fan, J. Bai, X. Lei, W. Lin, Q. Hu, G. Wu, J. Guo, and G. Tan, ``Ppmck: Privacy-preserving multi-party computing for k-means clustering," J. Parallel Distrib. Comput., vol. 154, pp. 5463, 2021.
[33] L. Chen, Z. Zhang, and X. Wang, ``Batched Multi-hop Multi-key FHE from Ring-LWE with Compact Ciphertext Extension," in Theory of Cryptography.
[34] P. Zhang, T. Huang, X. Sun, W. Zhao, H. Liu, S. Lai, and J. K. Liu, ``Privacy-Preserving and Outsourced Multi-Party K-Means Clustering Based on Multi-Key Fully Homomorphic Encryption," IEEE Trans. Dependable Security Comput., pp. 112, 2022.
[35] A. Blum, C. Dwork, F. McSherry, and K. Nissim, ``Practical privacy: the sulq framework," in Proc. ACM PODS, Baltimore, Maryland, Jun. 2005, pp. 128138.
[36] C. Dwork, K. Talwar, A. Thakurta, and L. Zhang, ``Analyze gauss: optimal bounds for privacy-preserving principal component analysis," in Proc. 46th Annu. ACM Symp. Theory Comput. (SIGMOD), New York, NY, USA, 2014, pp. 1120.
[37] W. Jiang, C. Xie, and Z. Zhang, ``Wish art mechanism for differentially private principal

components analysis," in Proc. 13th AAAI Conf. Artif. Intell. (AAAI). Palo Alto, vol. 30, no. 1, CA, USA: AAAI Press, 2016.
[38] J. Li, J. Wei, M. Ye, W. Liu, and X. Hu, ``Privacy-preserving constrained spectral clustering algorithm for large-scale data sets," IET Infor. Secur., vol. 14, no. 3, pp. 321331, 2020.
[39] B. Liu, L. Chen, X. Zhu, and W. Qiu, ``Encrypted data indexing for the secure outsourcing of spectral clustering," Knowl. Inf. Syst., vol. 60, no. 3, pp. 13071328, Sep. 2019.
[40] S. Sharma, J. Powers, and K. Chen, ``Private graph: Privacy-preserving spectral analysis of encrypted graphs in the cloud," IEEE Trans. Knowl. Data Eng., vol. 31, no. 5, pp. 981995, May 2019.
[41] S. Wang, Y. Zheng, X. Jia, and X. Yi, ``Privacy-Preserving Analytics on Decentralized Social Graphs: The Case of Eigen decomposition," IEEE Trans. Knowl. Data Eng., pp. 115, 2022.
[42] L. Zhou and C. Li, ``Outsourcing Eigen-Decomposition and Singular Value Decomposition of Large Matrix to a Public Cloud," IEEE Access, vol. 4, pp. 869879, 2016.
[43] Y. Zhang, X. Xiao, L.-X. Yang, Y. Xiang, and S. Zhong, ``Secure and Efficient Outsourcing of PCA-Based Face Recognition," IEEE Trans. Inf. Forensics Security, vol. 15, pp. 16831695, 2020.
[44] Y. Ren, Z. Song, S. Sun, J. Liu, and G. Feng, ``Outsourcing LDA-Based Face Recognition to an Untrusted Cloud," IEEE Trans. Dependable Security Comput., pp. 11, 2022.