On Picture Fuzzy Information-Based Hybrid Cryptographic TOPSIS Approach for Best Suitable Cloud Storage and Security Level

Himanshu Dhumras, Rakesh Kumar Bajaj, and Varun Shukla

1 Introduction

Even while cloud computing has gained a lot of popularity, the lack of adequate security measures prevents the majority of businesses or customers from implementing it. Because of flawed plans, programming or configurations made by designing agencies and service provider firms at various architectural layers, such as infrastructural ground and applications that can compromise the evaluation of the contracted quality of service (QoS), the cloud may suffer from a number of vulnerabilities. Additionally, attackers choose the cloud as a favorite target since it allows them to engage in offensive behavior. Therefore, in order to protect the large cloud market, suitable and more protective cloud security is essential. Security issues are the most important problem that cloud computing is currently experiencing, according to research from the International Data Corporation (IDC) [1]. For securing extremely sensitive data/information and limiting unauthorized/unauthenticated accesses in the cloud or elsewhere, data owners may need to encrypt data before outsourcing it to the commercial public cloud [2]. The conventional plain text keyword search-based data usage services may be rendered useless as a result. The enormous cost of data transfer/exchange capacity/capability in cloud-scale frameworks would make it even more unfeasible to download all the data and decode it locally. Finding a good search engine and protecting the privacy of cloud data that has been encrypted is therefore of utmost importance. Due to various unavoidable security/protection boundaries with various stringent requirements like the "index privacy", "data protection", "keyword

H. Dhumras · R. K. Bajaj (✉)
Jaypee University of Information Technology, Waknaghat, Solan PIN 173 234, Himachal Pradesh, India
e-mail: rakesh.bajaj@juitsolan.in

V. Shukla
PSIT, Kanpur, Uttar Pradesh, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

B. K. Roy et al. (eds.), *Cryptology and Network Security with Machine Learning*, Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-99-2229-1_12

privacy", and many more [3, 4], the "encrypted cloud data search framework" is still a challenging task for recent distributed/parallel computing systems.

Over the years, a lot of researchers have put forth several "cloud storage and security frameworks". However, due to the steadily growing strengths of "cloud users and security concerns", extensive research for improvement is still ongoing. The first study on cryptography-based cloud storage was published by Kamara et al. [5]. By utilizing the advantages of non-standard/non-traditional cryptographic approaches, such as "attribute-based encryption and searchable encryption", they created secure cloud storage systems for both consumer and enterprise applications. Their earlier efforts introduced other features like integrity, searchability, and verifiability, which corrected their earlier works' overemphasis on achieving confidentiality [6]. In a distributed storage system called the cloud, simulation of the data security issue in cloud data storage is proposed by Wang et al. [7], and their research suggested a useful plan with excellent data support and other features like block update, remove, and append. Here, the file distribution system is set up using the erasure-correcting code approach with the goal of ensuring data dependability. Data error localization and storage correctness insurance are integrated into the process. Additionally, their plan is strong, incredibly effective, and resistant to many failures and attacks like "Byzantine failure". The researchers also offer an improved "public auditing system" with a protocol that strongly supervises all the operations involving dynamic data [8].

As a result, by impersonating the fundamental Markle hash tree, the provable data possession (PDP) or proof of retrievability (PoR) scheme's existing soundness was enhanced. With the aid of an effective bilinear aggregate signature mechanism and third-party auditing (TPA), this approach was capable of handling numerous auditing tasks. The Boneh–Lynn–Shacham (BLS) algorithm's high computational cost, however, was a significant problem in this case. Additionally, their model was not capable to support correctness for both dynamic data and public verification. A group of storage servers that are often installed with the aid of hundreds to thousands of servers offers processing power in the cloud computing environment [9]. The authors of this article modeled a typical four-layered cloud-based dataset. Massive physical resources (such as storage and application servers) made up the lowest tier and assisted in strengthening the storage servers. These servers directly handled the next-level virtualization tools and services that allowed sharing of capabilities among the server virtual instances. The virtual occurrences, however, were isolated from one another, creating a fault-tolerant behavior and an isolated security context [10]. For maintaining the unique feature of the "encryption and decryption" keys is crucial since cloud computing relies heavily on replication. This problem has recently become a hurdle for Amazon's cloud computing platforms. However, the lack of foresight in cryptography might lead to regrettable outcomes [11]. This application places a lot of emphasis on key distribution and encryption. Takabi et al. [12] established a thorough security strategy for cloud computing systems. Their model suggested a few methods for addressing security issues. The model was made up of many security-related elements. Issues including identity management, access control, policy integration across many clouds, and trust management across several

clouds as well as between a cloud and its users were covered in the modules. Guleria et al. [13] presented a parameterized information measure for the Pythagorean fuzzy set with monotonicity and maximality feature along with an algorithm for solving a decision-making problem. The cryptographic assessment and enhancement is certainly a crucial component in the process of cloud computing for security reasons. Maintaining the uniqueness of the encryption and decryption keys is crucial since cloud computing relies heavily on replication. This problem has recently become a hurdle for Amazon's cloud computing platforms. By routinely verifying the hash estimation of the files kept in the huge data storage, Venkatesan et al. [14] suggested an effective multi-agent-based static and dynamic data integrity protection. The multi-agent system was necessary for their proposed model (MAS). The agent in this situation was capable of self-rule, cunning, social aptitude, and other things. Three entities (the client, service provider, and data owner) are included in the suggested architecture, and several agents are used to screen and maintain the data integrity.

The present paper has been structured as follows: Sect. 2 briefly presents very important preliminary definitions and fundamental notions which are available in the literature. Section 3 describes the security classification using the picture fuzzy information with the incorporation of various security parameters. The decision-making algorithms by making use of the TOPSIS technique for storage selection of servers have been done in Sect. 3. The necessary conclusions and advantages have been given in Sect. 4.

2 Preliminaries

In this section, we are presenting the basic notions and definitions of various other fundamental sets which are available in the literature. These preliminaries would help to understand the proposed notions of picture fuzzy hypersoft set and increase the readability for the researchers.

Definition 1 Intuitionistic Fuzzy Set(IFS) [15]. "An intuitionistic fuzzy Set R in V is given by $R = \{v, \rho_R(v), \omega_R(v) | v \epsilon V\}$;where $\rho_R : V \to [0, 1]$ is the degree of membership of v in R and $\omega_R : V \to [0, 1]$ is the degree of non-membership of v in R and ρ_R, ω_R satisfies the constraint $\rho_R(v) + \omega_R(v) \leq 1 \ (\forall \ v \in V)$; and $\pi_R(v) = (1 - (\rho_R(v) + \omega_R(v)))$ is called the degree of indeterminacy v in V. We denote the set of all intuitionistic fuzzy sets over V by $IFS(V)$".

Definition 2 Picture Fuzzy Set(PFS) [16]. "A picture fuzzy Set R in V is given by

$$R = \{v, \rho_R(v), \tau_R(v), \omega_R(v) | v \epsilon V\};$$

where $\rho_R : V \to [0, 1]$ is the degree of positive membership of v in R, $\tau_R : V \to [0, 1]$ is the degree of neutral membership of v in R and $\omega_R : V \to [0, 1]$ is the degree of negative membership of v in R and ρ_R, τ_R, ω_R satisfies the constraint

$$\rho_R(v) + \tau_R(v) + \omega_R(v) \leq 1 \ (\forall\, v \in V);$$

and, $\beth_R(v) = (1 - (\rho_R(v) + \tau_R(v) + \omega_R(v)))$ is called the degree of refusal membership of v in V. We denote the set of all the picture fuzzy sets over V by $PFS(V)$".

As per the findings by Prasad et al. [17], it is understandable that "*A proficient multi-agent-based static and dynamic data integrity protection by periodically confirming the hash estimation of the files stored in the massive data storage. Their proposed model depended on the multi-agent system (MAS). The agent here had a capacity for self-ruling, ingenuity, social ability, and so on. The proposed architecture incorporates three entities (i.e. client, service provider, and data owner) and has different agents to screen and keep up the data integrity.*"

Also, Sood et al. [18] stated that "*The concept of data security sections, which is followed in our paper in a different manner. Confidentiality, availability, and integrity parameters for cryptography in addition to the message authentication code (MAC) for checking the data integrity are utilized as a part of this procedure. The strategy provides classification, uprightness, authorization, verification, and non-repudiation and anticipates data spillage. The security degree that they provide in ascending order is MAC, classification of data, and execution of index and encryption system.*"

3 Security Classification Using Picture Fuzzy Information

This stage follows the encryption stage on the part of the data owner. The data owner would process the data in accordance with his selected security requirements after successfully logging into the CSP. The required P, Q, and R security parameters—P for "proactive threat detection and management", Q for "quality data backup", and R for "maximum uptime and lowest downtime" are listed and sent collectively to the CSP in order to be stored. Along with those three parameters are the encrypted message M', encrypted index IM' containing the user's most frequently searched phrases, and the secret key K_1 discussed in the previous section.

Here, a fuzzy-based method for storing data with various access kinds on various cloud storage servers depending on the three aforementioned crucial security characteristics is described (i.e., P, Q, and R). With a shorter execution time and without the additional load of dataset training, the proposed fuzzy-based approach has been used to categorize the access kinds where ambiguity or fuzziness will be handled utilizing membership functions. The CSP gives the user an option of various fuzzy variables for each security parameter. These variables will be selected by the user. These choices will be converted into a security factor (S_f) using the suggested method illustrated in the sections that follow.

The user might not be familiar with the process for assigning values to the aforementioned security parameters because they are qualitative in nature. The user will be able to list their needs on a more detailed level with the aid of the fuzzy linguistic variables as shown in Table 1.

Table 1 Linguistic variables for computing the security parameters

Qualitative term	PFNs
"Absolutely bad (AB)"	(0.83, 0.04, 0.11)
"Very very bad (VVB)"	(0.75, 0.05, 0.15)
"Very bad (VB)"	(0.62, 0.1, 0.2)
"Bad (B)"	(0.55, 0.11, 0.25)
"Medium bad (MB)"	(0.50, 0.15, 0.30)
"Medium (M)"	(0.45, 0.20, 0.35)
"Medium high (MH)"	(0.40, 0.22, 0.37)
"High (L)"	(0.35, 0.25, 0.40)
"Very high (VH)"	(0.25, 0.30, 0.43)
"Very very high (VVH)"	(0.15, 0.35, 0.48)

The range for security factor S_f is between $0 \leq S_f \leq 0.399$ for public access type and for private access type it is ranging between $0 \leq S_f \leq 0.799$. Also, for the limited access owner, the range is between $0.8 \leq S_f \leq 1$.

Selection of server storage and data storage

After data encryption on the owner's end, the data will be sent to the cloud for storage. Instead of being stored at a single server, the suggested approach will divide the data over a number of separate, geographically dispersed storage servers. During this stage, the CSP's registered storage servers are chosen from a pool of available servers for data storage. Some cloud service providers split apart the data that they get from the data owner. Next, each data component is saved on a distinct storage server with a different storage type, level of security, etc. at a different geographic location. To ensure an effective, secure, and quick data storage process, the task is divided across several access level sites (based on S_f) on various storage servers.

Computation of the weights of the criterion

The area that concerns the most is data security. As a result, the storage server's level of security is the most crucial one, because they reduce execution time and network bandwidth, processing speed, and time delay. These factors are regarded as the most important factors when storing data concurrently on several storage servers, which aids in determining the additional communication cost in a cloud environment.

Analytic Hierarchy Process (AHP) is used in this case to generate the weights of the criteria through pairwise comparisons for each of the selected criteria (AHP). AHP converts empirically based comparisons into numerical numbers for further analysis and comparison. The most popular scale is the relative importance scale between two criteria, as proposed by Saaty [19]. The scale, which has values ranging from 1 to 9, is shown in Table 2 and it is used to compare one criterion's importance to another criterion. The consistency ratio (CR) gauges how consistently respondents answer questions on the AHP forms. Using Saaty's significant scale [19] provided in

Table 2 Scale of significance for criterions

Qualitative term	PFNs
"Extremely important (EI)"	(0.83, 0.04, 0.11)
"Very important (VI)"	(0.60, 0.05, 0.21)
"Important (I)"	(0.53, 0.12, 0.25)
"Less important (LI)"	(0.45, 0.15, 0.30)
"Very less important (VLI)"	(0.30, 0.25, 0.35)

Table 3 Computations of criterion weights

Criteria	Weights
"No. of CPU"	0.0403
"Avg. processing speed"	0.1276
"Security level"	0.3894
"Avg. transmission speed"	0.2611
"Avg. time delay"	0.0998
"Avg. memory utilization"	0.0268

Table 2 to build the pairwise comparison matrix. The final weight for each criterion determined using the AHP technique is shown in Table 3. These weights will then be processed further in the TOPSIS method for the selection of storage servers.

Multiple Storage Servers Selection

For handling the uncertainty found in the process of selection of data storage centers, the role of fuzzy set theories and soft computing decision-making techniques become quite important and can be extensively utilized for better results. In the selection process, a number of quantitative and qualitative criteria must be taken into account. In order to solve the storage server selection problem, a combination of these two approaches is used here. The suggested model uses a picture fuzzy TOPSIS technique, in which the ratings of various storage servers under various criteria are appraised in linguistic words represented by picture fuzzy numbers, to choose the best storage servers under real-time conditions. The implementation of fuzziness for three qualitative criteria termed as "degree of security", "memory use", and "time delay", for which the linguistic scaling/fuzzification are presented in Table 4. Some significant decision factors for storage servers are listed in Table 5.

Each storage server registered with the CSP has a direct relationship with its relevant properties. Some of them are static, while others were captured at the moment. In Table 6 , the values of qualitative variables like "degree of security", "average memory use", and "average time delay" are taken into account. While the qualitative qualities are expressed in fuzzy linguistic terms, the quantitative attribute values are simply determined from the various storage servers' current behavior. By using

Table 4 Security level

Qualitative term	PFNs
"Very low (VL)"	(0.23, 0.04, 0.11)
"Low (L)"	(0.51, 0.05, 0.21)
"Moderate low (ML)"	(0.53, 0.12, 0.17)
"Fair (F)"	(0.38, 0.15, 0.30)
"Moderate high (MH)"	(0.29, 0.25, 0.23)
"High (H)"	(0.10, 0.25, 0.35)
"Very high (VH)"	(0.05, 0.25, 0.22)

Table 5 Utilization of memory

Qualitative term	PFNs
"Low (L)"	(0.33, 0.02, 0.12)
"Medium (M)"	(0.54, 0.03, 0.21)
"High (H)"	(0.33, 0.04, 0.17)

Table 6 Time delay

Qualitative term	PFNs
"Close (C)"	(0.73, 0.04, 0.11)
"Adequate (A)"	(0.52, 0.05, 0.18)
"Fair (F)"	(0.37, 0.12, 0.23)

the TOPSIS algorithm, it is feasible to order the storage servers according to their priority in the selection process.

Picture Fuzzy TOPSIS Algorithm for the selection of storage server
Step 1: In the first step, there is the conversion of linguistic variables into picture fuzzy numbers.
Step 2: In this step, there is the conversion of the picture fuzzy numbers to crisp numbers.
Step 3: In this step, the different attribute values which are not in the range of 0 to 1 need to be normalized so that the computations in decision-making can be done. Normalization can be done by $x_{ij} = \frac{y_{ij} - \min y_{ij}}{\max y_{ij} - \min y_{ij}}$, where, y_{ij} is the value of the jth criteria.
Step 4: Now, calculation of the normalized weighted decision matrix:
$v_{ij} = x_{ij} \times w_{ij}$, where v_{ij} is the weighted normalized data and w_{ij} denotes the weight of the jth criteria.
Step 5: Next, calculation of positive and negative ideal solution:

$$V_j^+ = \{v_1^+, v_2^+, ..., v_m^+\}$$
$$V_j^- = \{v_1^-, v_2^-, ..., v_m^-\}.$$

Step 6: In this step, the computation of the distance of every attribute value v_{ij} from a positive ideal solution (V_j^+).

$$D_i^+ = \sqrt{\sum_{j=1}^{n} \frac{1}{3}(v_{ij} - v_j^+)^2}; i = 1, 2, \ldots, n; j = 1, 2, \ldots, m.$$

Step 7: Similarly, the computation of the distance of every attribute value v_{ij} from a negative ideal solution (V_j^-).

$$D_i^- = \sqrt{\sum_{j=1}^{n} \frac{1}{3}(v_{ij} - v_j^-)^2}; i = 1, 2, \ldots, n; j = 1, 2, \ldots, m.$$

Step 8: In this step, computation of the coefficient of relative closeness can be done by making use of the following formula:

$$CRC_i = \frac{S_i^-}{S_i^- + S_i^+}; \text{where}, 0 \leq CRC_i \leq 1; i = 1, 2, \ldots, n; j = 1, 2, \ldots, m.$$

Step 9: In the final step, ranking of the alternatives can be done on the basis of the values of CRC_i in decreasing order (Table 7).

Table 7 Computation of the ranking orders of the storage servers

Storage server	D_i^+	D_i^-	CRC_i	Order
"(SS_1)"	0.073	0.017	0.315	7
"(SS_2)"	0.042	0.054	0.623	3
"(SS_3)"	0.058	0.047	0.346	6
"(SS_4)"	0.018	0.074	0.812	1
"(SS_5)"	0.039	0.052	0.587	4
"(SS_6)"	0.030	0.058	0.628	2
"(SS_7)"	0.053	0.061	0.558	5

4 Conclusions and Scope for Future Work

In order to secure the privacy of distributed cloud storage systems, this study proposes a cloud storage framework/technique that uses a 128-bit encryption key generated by synchronizing a deoxyribonucleic acid (DNA) cryptographic approach with the Hill Cipher algorithm. With the aid of a picture fuzzy information-based classification methodology, the data in this document have been categorized in accordance with several security parameters. Additionally, this architecture would let you pick the best storage server from a selection. Further, a picture fuzzy information-based technique for order of preference by similarity to ideal solution (TOPSIS) decision-making algorithm has been implemented to determine the best storage server where the data can be saved, reducing the execution time in the process. Also, the extension of this work can be done by executing various other methods like AHP, WASPAS, VIKOR in different techniques.

Declarations and Compliance with Ethical Standards

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.
Funding Details: The authors declare that the research carried out in this article has no source of funding.
Conflict of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Authorship contributions: The authors have equally contributed to the design and implementation of the research, to the analysis of the results, and the writing of the manuscript.

Acknowledgements We are very much thankful to the anonymous reviewers for suggesting the points/mistakes which have been well implemented/corrected for the necessary improvement of the manuscript. We sincerely acknowledge our deep sense of gratitude to the Editorial office and reviewers for giving their valuable time to the manuscript.

References

1. Gartner N (2012) Consumers will store more than a third of their digital content in the cloud by 2016. Press Release
2. Velte AT, Velte TJ, Elsenpeter RC, Elsenpeter RC (2010) Cloud computing: a practical approach. McGraw-Hill, New York, pp 44

3. Kamara S, Lauter K (2010) Cryptographic cloud storage. In: Proceedings of the international conference on financial cryptography and data security. Springer, Berlin, Heidelberg, pp 136–149
4. Cao N, Wang C, Li M, Ren K, Lou W (2014) Privacy-preserving multi-keyword ranked search over encrypted cloud data. IEEE Trans Parallel Distrib Syst 25(1):222–233
5. Chen R, Mu Y, Yang G, Guo F, Wang X (2016) Dual-server public-key encryption with a keyword search for secure cloud storage. IEEE Trans Inf Forensics Secur 11(4):789–798
6. Chase M, Kamara S (2010) Structured encryption and controlled disclosure. In: International conference on the theory and application of cryptology and information security. Springer, Berlin, Heidelberg, pp 577–594
7. Wang C, Wang Q, Ren K, Lou W (2009) Ensuring data storage security in cloud computing. In: Proceeding of the IWQoS, pp 1–9
8. Wang Q, Wang C, Ren K, Lou W, Li J (2011) Enabling public auditability and data dynamics for storage security in cloud computing. IEEE Trans Parallel Distrib Syst 22(5):847–859
9. Buyya R, Murshed M (2002) Gridsim: a toolkit for the modeling and simulation of distributed resource management and scheduling for grid computing. Concurr Comput: Pract Exp 14(13–15):1175–1220
10. Smith JE, Nair R (2005) The architecture of virtual machines. Computer 38(5):32–38
11. Balding C (2008) Is your amazon machine image vulnerable to SSH spoofing attacks?. Cloud Secur
12. Takabi H, Joshi JB, Ahn GJ (2010) Securecloud: towards a comprehensive security framework for cloud computing environments. In: Proceedings of the 34th annual computer software and applications conference workshops (COMPSACW). IEEE, pp 393–398
13. Guleria A, Bajaj RK (2018) Pythagorean fuzzy R-norm information measure for multicriteria decision-making problem. Adv Fuzzy Syst. Article ID 802301
14. Venkatesan S, Vaish A (2011) Multi-agent based dynamic data integrity protection in cloud computing. In: Proceedings of the international conference on advances in communication, network, and computing. Springer, Berlin, Heidelberg, pp 76–82
15. Atanassov KT (1986) Intuitionistic fuzzy sets. Fuzzy Sets Syst 20(1):87–96
16. Coung B (2014) Picture fuzzy sets. J Comput Sci Cybern 30(4):409–420
17. Prasad P, Ojha B, Shahi RR, Lal R, Vaish A, Goel U (2011) 3 dimensional security in cloud computing. In: 2011 3rd international conference on proceedings of the computer research and development (ICCRD). IEEE, pp 198–201
18. Sood SK, Sarje AK, Singh K (2011) A secure dynamic identity based authentication protocol for multi-server architecture. J Netw Comput Appl 34(2):609–618
19. Saaty TL (2005) The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making. In: Multiple criteria decision analysis: state of the art surveys, vol 78. Springer, New York, NY, pp 345–405