IOT AWARE TVERSKY CORRELATED PIECEWISE REGRESSIVE MERKLE–HELLMAN KNAPSACK CRYPTOSYSTEM FOR SECURE INDUSTRIAL DATA STORAGE IN CLOUD

Flindon W.G[1], Divya S.V[2*]

[1]Department of Computer Application, Noorul Islam Centre for Higher Education, Thuckalay, Kumaracoil, Tamil Nadu 629180. mrflindon@gmail.com

[2*]Department of Computer Science and Engineering, V.S.B College of Engineering Technical Campus, Coimbatore, Tamil Nadu 642109. divyasadasivam@gmail.com

ABSTRACT

Data storage in Cloud Computing gives an coherent and rapid manner of conceding access to authorized users. The cloud data storage system provides a service for accumulating larger data in storage servers. Data saved in the cloud are stored as well as accessed from the internet over a longer time. This in turn directs to data integrity and confidential data loss. However, existing new identity-based sign cryption scheme proven to be failed in supporting revocation facility. To address this issue, novelty of Internet of Things-aware Tversky CorrElated Piecewise RegressivE Knapsack Cryptosystem is performed into secured data storage on a cloud server with improved data confidentiality rate of data storage. At first, Internet of Things devices are obtained for collecting an enormous quantity of industrial informations. Internet of Things-aware Tversky CorrElated Piecewise RegressivE Knapsack Cryptosystem method is split into three process specifically, creating keys, encrypting data and decrypting it to develop cloud storage security levels. Employing Cloud simulator on the Java working platform, the suggested strategy is put into practice. Overall quantitative results confirm that Internet of Things-aware Tversky CorrElated Piecewise RegressivE Knapsack Cryptosystem has an enhanced data confidentiality efficiency of 13%, 8% accuracy and Data Integrity level of 12% in addition to minimize the 24% of Computational Overhead and 15% minimum storage overhead when compared to conventional methods.

KEYWORDS: Cloud computing, Secured data storage, Tversky correlated piecewise regression, Merkle–Hellman Knapsack Cryptosystem

INTRODUCTION

Cloud storage system requires for offering effective as well as flexible storage services through quality promise for users (Malallah *et al.* 2023). Due to the development of more internet applications with big data and a large number of users, cloud storage service becomes a serious challenge. Industrial Internet of Things (IIoT) application plays a significant part, where secure, scalable storage service is being developed for smart industry (Ahmed *et al.* 2023). Traditional IIoT architectures are vulnerable to several cyber-attacks. Therefore an efficient technique is required to solve the cyber-attacks for guaranteeing the security of cloud storage. The reasons behind the commencement of the cloud storage system have led to an upsurge in industry scenarios. Also, the IIoT work should be a modern society by improving prospective for relocation with no exit the daily life necessities. The IIoT uses the concept of IoT and it is employed to over the several real time industrial applications (Qi *et al.* 2023).

A SmartCrypt, a data storing and sharing system was developed by Halder and Newe for encrypting the time-series data by using symmetric homomorphic encryption (Halder & Newe, 2022). However, it failed to use innovative encryption technologies, to support high-volume and high-velocity data storage. For secure cloud storage, A productive Identity-based Privacy-Preserving Provable Data Possession (ID-P^3DP) method was introduced based on RSA cryptographic technique by Ni *et al.* (2022). However, higher confidentiality and integrity were not achieved.

Blockchain-Enabled Shark Smell Optimization with Hopfield Chaotic Neural Network (SSO-HCNN), based on reliable encryption for the IoT environment was developed by Khayyat *et al.* (2022) Designed approach failed to employ lightweight cryptographic techniques through authentication schemes for promising security in IoT field. A method has been presented for sharing information that is reliable as well as safe based on blockchain by Chi *et al.* (2020). However, the overhead was not minimized.

A blockchain-based architecture was developed for secure and trustworthy industrial operations by Latif, *et al.* (2022). But the secure data storage was a major challenging issue. The Cipher text-Policy Attribute-Based Encryption (CP-ABE) technology has been created, enabling safe cloud data preservation by Ning *et al.* (2021). But it failed to maintain a similar level of security. A data integrity auditing method was introduced for secured data storage by Shen, *et al.* (2021). However, it failed to enhance data confidentiality. To preserve cloud data confidentiality, A Cloud Secure Storage Mechanism called CSSM was designed by Song *et al.* (2021). But it was not effective to protect cryptographic materials from a storage perspective.

For cloud storage to minimize computational overhead, A Secure and Efficient Privacy-Preserving Provable Data Possession (SEPDP) technique was designed (Nayak & Tripathy, 2021; Tian *et al.* 2024). But the minimal storage overhead was not achieved. For secure data storage on cloud servers, the shuffle standard onetime padding encryption (S^2OPE) method was developed (Devi & Devi, 2021; Flavia & Chelliah, 2024). The designed technique enhances the integrity level but the storage overhead was not considered.

The following outlines the primary contribution of the suggested TEPEK method:

- At first, a novel TEPEK method is determined by few process specifically, creation of keys, authentication, encryption and decryption/decoding. Security of industrial data storage on cloud server was improved by proposed TEPEK method.
- Proposed TEPEK method uses the Merkle–Hellman Knapsack Cryptosystem for session key generation. After getting key pair, user utilises the public key to encrypt the data and saves it on a server in the cloud. Improved safety measures and increased integrity and privacy of data are implemented by the server when storing user data.

- A Tversky correlated piecewise regression is applied to verify the authenticity of the cloud user by using TEPEK method. This method helps to identify the authorized and unauthorized cloud users with improved accuracy. The authenticated user stores or accesses data from cloud server.
- Finally, experimental analysis are performed in several parameters to underline the improvement of TEPEK performs excellent in contrast to traditional techniques.

The rest of this article is structured as follows, Section 2 provide analysis of related works. Section 3 is introduced a TEPEK by brief explanation. In Section 4, the experimental parameters and data set specification are described. Section 5 explains performance of outcomes as well as analysis by different metrics. Finally, concludes with summarization of the work and concludes in section 6.

RELATED WORK

A new cryptographic technique was designed by Sohal & Sharma (2022) for encrypting data before uploading it onto cloud. It reduces encryption time however confidentiality rate was not minimized. For IoT-aware cloud storage, Secure and Efficient Multi-Authority Access Control technique was developed by Xiong *et al.* (2020). Though the method minimizes the computational overhead and lower storage costs, authentication was not performed. A novel Security framework that integrates blockchain and Coalition Formation theory was developed to facilitate the generalized security for the IIoT by Sharma *et al.* (2021). But it failed to identify the different security attacks for the IIoT network. For encoding uploaded documents on cloud storage, An encryption algorithm was designed by Jahan *et al.* (2021). Proposed algorithm takes lesser time for encryption and decryption of files however failed to minimize storage overhead.

To store data on centralized cloud, centralized cloud-based cross-domain data-sharing approach was introduced (Singh, 2021). However, approach was not suitable for large-scale data. For number of different data encryption to enhance safe cloud data storage as well as computation, distributed data protection system was developed by Rafique *et al.* (2021). But it failed to analyze the reliability, resilience, and scalability conditions. A QoS-oriented replica approach was designed for secure cloud storage system by Yu *et al.* (2021). However, the costs associated with time and space for a safe cloud storage solution were not reduced. For enhancing cloud data storage, Hybrid blockchain architecture was developed by Hasan *et al.* (2021). Other than it failed to use efficient models to handle the big data for cloud data storage.

For attaining maximum level of privacy and storage on cloud, a new Dynamic Searchable Symmetric Encryption (DSSE) method was introduced by Hoang *et al.* (2021). But it failed to perform large datasets analysis. To minimize computation cost, Delegatable Proofs of Storage (DPOS) approach has been put forward by Yang *et al.* (2021). However higher data confidentiality was challenging issue. Re-definable access control (RDAC) was introduced to preserve data privacy via encryption by Zhang *et al.* (2019). However, computation overhead was not considered. Modified artificial neural network (MANN) was developed for higher storage security (Ruth *et al.* 2019). The security issues are addressed by enhanced secure data backup scheme (Hu *et al.* 2019). Secure and privacy-preserving protocol named SPCOR was examined to obtain end-to-end security by Alex *et al.* (2020). But, the storage overhead was not measured. Privacy preserving weighted similarity search was employed for achieving secure data communication (Guo *et al.* 2019). However, it failed to consider multiple users.

METHODOLOGY

Performance, improved accessibility and reduced expenses make CC an appealing field. The internet is used by cloud users to access a multitude of applications. With many services, data storage service is a main aspect which CS gives for storing large amount of storage capacity. A cloud server handles different IoT applications. A new IoT system, specifically for production automation and smart factories, is suggested by an appliance in the industrial sphere to safeguard confidential information, mainly product quality flaws. CC's storage security procedure entails preserving stored industrial sensitive information from unauthorized user access. But, authentication accuracy in data storage was not enhanced by conventional methods. However still, numerous industries not prepared for executing CC technology because of not have appropriate security control policy which leads to many challenges in cloud computing. To overcome the above issue, a new TEPEK method is presented in this work to raise the degree of security when storing data on cloud computing.

Multiple entities that contribute to safe storage and accessibility of information are included in the architecture of the cloud. The design includes cu_1, cu_2, cu_3….cu_n cloud users and the quantity of data that the industrial d_1, d_2, d_3….d_n wants to retain and access from the cloud server (C_r). Cloud service provider (Cp_r) provides storage services to CU.

Figure 1 depicts process involved in TEPEK to guarantee security of data storage as well as access in CC environment. IoT devices are initially used to gather information $d_1, d_2, d_3, \ldots. d_n$ from industrial settings. After collecting information, the user desires for storing in data on CS for additional processing. The cloud user initializes their details with the server by registering them through data storage. Following the registration procedure, the server utilizes the Merkle-Hellman Knapsack Cryptosystem for creating a set of keys for each registered member.

Merkle–Hellman knapsack is public key cryptosystem is comprised two types of session keys specifically public and private in order to store data safely on cloud. Public key is determined for encryption and private key for decryption. Main advantage of knapsack cryptosystems to improve the speed of encryption with decryption than additional cryptography.

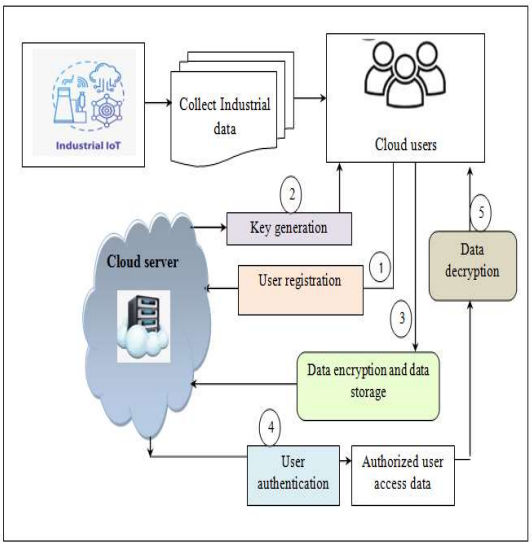

Figure. 1. Architecture diagram of the proposed TEPEK

After both registered users' keys are generated, the user encodes data using the recipient's public key. Encryption of information is a process of encoding plain text or original data into cipher text (i.e., encrypted data). Encrypted data are stored in CS. Whenever user needs for accessing data from server, they need to validate authenticity by using Tversky correlated piecewise regression. Piecewise regression is machine learning technique employed for finding relationship between two variables by using Tversky correlation coefficient. It is mathematical model employed for finding correlation among user entered as well as already stored password in CS. Based on the correlation coefficient results, authorized user and unauthorized users are properly recognized. Finally, authorized user

performs decryption for attaining original data. In this manner, secured data storage carried out in CS. These three processes of TEPEK are described as follows.

Security model

In this proposed work, Secure Industrial Data Storage in Cloud using IoT Aware Tversky Correlated Piecewise Regressive Merkle–Hellman Knapsack Cryptosystem (TEPEK) method introduced into secured data storage on a cloud server with enhanced data confidentiality rate. TEPEK method is involves into three process such as, key generation, encryption, and decryption for developing security levels in cloud storage. Initially, the cloud member enters their data to Cloud Server (CS) is store data. Following registration, CS leverages Merkle-Hellman Knapsack Cryptosystem strategy to generate a session key pair - that is, a private and public key, using each registered member in the key development phase in order to store information in CS. After obtaining the key pair, the Cloud User (CU) utilizes the public key to encrypt information which is then sent to CS to be stored at an even greater level of security. A cloud user sends a request message to CS in order to obtain access to the information gathered from CS. Following reception of the request, CS provided the transaction password in order to use Tversky correlated piecewise regression to verify CU authentication. When a cloud user inputs a password that CS sends back and matches, the user is considered authorized and is allowed to retrieve or access information obtained from CS. After that, data attain decrypted with a private key of CU. It helps to improve cloud storage security and integrity by industrial IoT. The proposed technique explained in below section.

User registration and key generation

User registration is essential method of TEPEK. In order to utilize cloud services, users have to initially register their information with CS. Flow procedure of user registration has been illustrated in Fig. 2. The user must complete the registration form that the CS provider gives. The user's birthdate, gender, phone number, email address, real name, middle name, last name and other details are included in the information provided. Subsequent to entering personal information, CS sent successfully registered messages to user. Once user completes registration, CS generates pair of keys for secure data storage on CS.

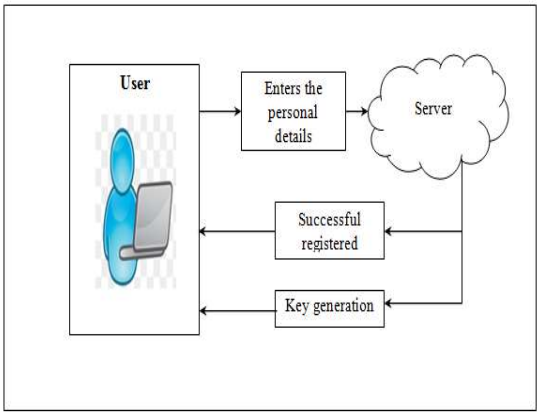

Figure. 2. Flow process of the user registration

CS uses Merkle–Hellman Knapsack Cryptosystem for generating pair of session keys for every registered user for encryption as well as decryption. Here generated key pair valid for one login session or communication. After the session time expired, the key was disabled automatically. This in turn reduces unauthorized access.

Let us consider the random increasing sequence of positive integers $q_1, q_2, q_3, \ldots . q_n$

$$Q = (q_1, q_2, q_3, \ldots . q_n) \quad (1)$$

After that, the random integer 'B' is selected which is greater than the sum of positive integers.

$$\mathrm{B} > q_1, q_2, q_3, \ldots . q_n \quad (2)$$

Followed by, a random integer 'K' i.e. $GCD\,(K, \mathrm{B}) = 1$ i.e. K, B are co-prime numbers. As a result, public key of user is generated as below,

$$Y = [c_1, c_2 \ldots c_n] \text{ Where, } c_i = K\, q_i \mod B \quad (3)$$

Private key of user is generated as follows,

$$X = (Q, \mathrm{B}, \mathrm{K}) \quad (4)$$

In above Eq. (3) and Eq. (4), 'X' is private key and 'Y' is public key were generated. The session key pairs are produced for every individual in the following manner. The generated keys are sent from CS to user.

Tversky correlative piecewise regressive merkle–hellman knapsack cryptosystem-based secure data storage

After that, the user wants to store the collected data on CS for additional processing by using key generation process. Prior to storing information, the cloud customer encodes industrial data using the public key of the recipient and delivers it to CS to ensure greater security data storage. The process of transforming raw data into unintelligible form (i.e., cipher text) is known as encryption. The encryption diagram for transforming raw industrial data into cipher text is displayed in Fig. 3. When user requests to store data, they perform encryption with aid of receiver session public key.

Let us consider the industrial data $d_1, d_2, d_3, \ldots . d_n$. The binary string $b_1, b_2, \ldots b_n$ that is [0, 1] was created from the input data, which is a positive integer. Then cipher text of information is generated by receiver public key is formulated as,

$$CT = \sum_{i=1}^{n} b_i * q_i \quad (5)$$

From above Eq. (5), CT is specified as cipher text of input data obtained as sum of product of the input binary string 'b_i' and 'q_i' is a public key of the receiver. In this approach, encryption is determined with a better security level and cipher text stored on CS.

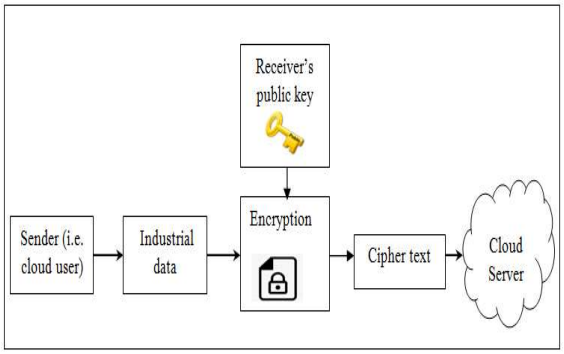

Figure. 3. Block diagram of the encryption

When CU needs for accessing stored data from CS, first user transmits the request message to CS. Subsequent to receiving request, the server confirms cloud user's authenticity by sending transaction password. When CU entered password attain matched by server-sent password, CU is authenticated and allowed for storing or accessing data from CS. Password-matching process carried with aid of Tversky correlative piecewise regression. Piecewise regression is machine learning technique applied for examining relationship with two variables namely passwords by using Tversky correlative index. The regression analysis partitions the user into the authorized user or unauthorized user. Therefore, the

authentication schemes were developed to ensure the security in IoT environment. The Tversky correlation between two passwords is mathematically formulated as,

$$S = \frac{(PW_S \cap PW_E)}{\tau(PW_S \Delta PW_E) + \sigma(PW_S \cap PW_E)} \quad (6)$$

From Eq. (6), S is specified as Tversky correlation coefficient, PW_S is a stored password, PW_E is an entered password. $PW_S \Delta PW_E$ exhibits a difference between both of the passwords, $PW_S \cap PW_E$ suggests a mutual reliance between the two passwords. From Eq. (1), τ and σ denote the Tversky index parameters ($\tau, \sigma \geq 0$). Coefficient (s) gives resultant value between [0, 1].

$$S = \begin{cases} 1, & authorized\ user \\ 0, & unauthorized\ user \end{cases} \quad (7)$$

In this manner, user authentication is determined for improving data confidentiality level. After authentication, the server permits data access to authorized user. Authorized user performs decryption and obtains original data. Initially, modular inverse of the 'K' is calculated as,

$$K^{'} = K^{-1} \ (mod\ B) \quad (15)$$

The computation of $K^{'}$ is independent of the input data and it is computed when private key 'B' is generated. By applying multiplicative congruential generator, the cipher text is generated as given below,

$$CT^{'} = CT\ K^{'} \ (mod\ B) \quad (8)$$

Where, $CT^{'}$ denotes cipher text. First, initialize the empty set of index list $R = \{r_1, r_2, \ldots r_3\}$. Then find the largest element in 'Q' which is less than or equal to $CT^{'}$ then it is denoted as q_j .

$$CT^{'} = CT^{'} - q_j \tag{9}$$

Then add the j to the empty set of the index list. 'This process repeated until the $CT^{'} = 0$ *or* $CT^{'} \geq 0$. Finally, the original data is obtained as follows,

$$d = \sum_{i=1}^{k} 2^{n-R} \tag{10}$$

From above Eq. (10), d is a original data, n is a length of the input bit string, R is a set of index list. Like this, data decryption is achieve, and obtain the original data from CS. The algorithm for secure data storage and data access is given below,

In above algorithm 1, represent the Tversky correlative piecewise regressive Merkle–Hellman Knapsack Cryptosystem-based secure data storage in cloud. Initially, the numbers of industrial data are gathered by using IoT devices. Afterward, user desires for storing collected data on server for further processing. At first, user is performed by registration and enters their details to CS. Following registration, CS generates a private and public key combo for every registered user. The CU encodes its data before storing it on a cloud server in the CS. Data encryption reduces the space complexity in CS. Secure data storage on cloud enhances data confidentiality. When cloud user accesses data from server, server confirms the authenticity of cloud user for avoiding unauthorized access. After that server grant data access to authorized user. An authorized member decodes the contents and retrieves the original information. This helps to guarantee data access confidentiality.

EXPERIMENTAL SETUP

With the help of the CloudSim simulator, a pilot study of the recommended TEPEK and the four presently in use has been implemented via Java. The WUSTL-IIOT-2021 Data set for IIoT Cyber security Investigation is used in experiments to store data securely on CS by preventing unwanted access, or attacks. It may be accessed at https://www.cse.wustl.edu/~jain/iiot2/index.html. For four existing methods safe namely SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song & Li, 2021) and RDAC (Zhang *et al.* 2019). In order to facilitate the exchange of Industrial Internet of Things information between organizations, CS offers an internet-based platform featuring storage security. When user needs to access stored data, CS provides access for authorized users (i.e., normal samples) in cipher text form. Otherwise, the cloud server removes access for unauthorized users i.e attack samples. The user performs decryption and receives the original data. The primary goals of gathering information are to detect actual cyber attacks and to track industrial systems in the real world as closely as possible.

RESULTS AND DISCUSSION

An investigation of calculations comparing the recommended TEPEK and four current systems, including SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song & Li, 2021) and RDAC (Zhang *et al.* 2019) are discussed by different matrics namely,

- Space complexity
- Data integrity rate
- Computation overhead
- Data confidentiality rate
- Authentication accuracy

Impact of data confidentiality rate (DC)

DC is the quantity of authorized users' access (i.e., normal samples) to cloud user information and it is protected from unauthorized users. It is measured as,

$$DC = \left\{\frac{D_{prot}}{D}\right\} * 100 \quad (11)$$

The data confidentiality level, the amount of data samples and the protection of industrial data from unauthorized users are all determined by Eq. (11). Measured in percentage (%). Results of data confidentiality rate calculated utilizing proposed TEPEK method outperforms existing SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song & Li, 2021) and RDAC (Zhang *et al.* 2019) is explained in Table 1.

Table 1 depicts the outcomes of data privacy level with plenty of data from industries. The data confidentiality rate's performance is measured by applying five methods namely the TEPEK method over the existing SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song & Li, 2021) and RDAC (Zhang *et al.* 2019). When compared to other approaches the acquired findings reveals that the TEPEK has produced an improved data confidentiality level. For example, the TEPEK obtains a 98.74% data confidentiality rating when 10,000 industrial data sets have been included, while the other four current models performance levels SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song & Li, 2021) and RDAC (Zhang *et al.* 2019) was found to be 92.45%, 91.12%, 90.79% and 86.43%. Similarly, with varying amounts of information, various degrees of performance are seen for every approach. Compared the overall observed findings of TEPEK to those obtained using traditional procedures. Ultimately, the performance enhancement of TEPEK over

the other available approaches was demonstrated using the average of ten comparative outputs. When compared to the current SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song & Li, 2021) and RDAC (Zhang *et al.* 2019), the data confidentiality rate analysis utilizing TEPEK is enhanced by 9%, 13%, 17% and 22%. This is due to applying Tversky correlative piecewise regressive Merkle–Hellman Knapsack Cryptosystem in the cloud.

Table 1.Comparison of data confidentiality rate over the number of data

No. of data	**Data confidentiality rate (%)**				
	TEPEK	**Smart Crypt**	**ID-P^{3}DP**	**CSSM Method**	**RDAC**
10000	98.74	92.45	91.12	90.79	86.43
20000	98	91	89.28	87.56	84
30000	97.83	90.4	88.28	86.16	82.65
40000	96.41	90.31	87.03	83.75	80.66
50000	96.2	88.22	86.25	84.28	80.33
60000	96.18	87.43	84.09	80.25	77
70000	96.07	87.36	83.46	79.56	75.65
80000	96	86.93	83.23	79.53	75.42
90000	95.68	86.28	82.36	78.44	75.22
100000	94.25	85.64	82.45	79.27	75

The proposed cryptosystem encrypt the data and stored in the CS. Encryption of data on cloud enhances data confidentiality. Therefore, it aids to avoid unauthorized access and authorized user accesses data from CS by decryption process. As a result, improved data confidentiality is achieved by protecting the data.

Impact of data integrity rate (DI)

The term DI pertains to the entire quantity of industrial data that are stored in the cloud and are not altered by unauthorized individuals. The mathematically calculated way is given below:

$$DI = \left(\frac{D_{NA}}{D}\right) * 100 \quad (12)$$

According to Eq. (12), 'DI' data integrity is evaluated depending on industrial data that has not been altered into'D_{NA}' data storage to industrial data involved in simulation process 'D'. The percentage (%) is used to express data integrity.

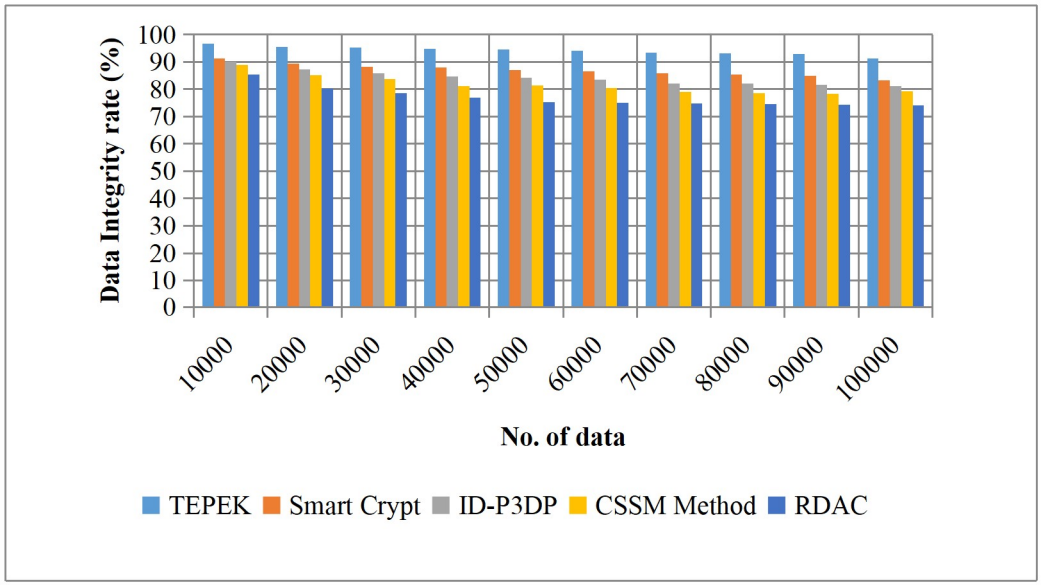

Figure. 4. Graphical illustration of data integrity rate

The data integrity rate's result utilizing proposed TEPEK method outperforms existing SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022) and CSSM (Song & Li, 2021) is explained in Figure 4. Figure 4 depict performance results of data integrity rates by various methods, the TEPEK method,

SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song & Li, 2021) and RDAC (Zhang *et al.* 2019) with number of data. TEPEK gives enhanced results when compared to conventional methods. Assume that 10,000 data points are provided as input. Data integrity level of TEPEK had been determined to be 96.58% while the data integrity percentage using SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song & Li, 2021) and RDAC (Zhang *et al.* 2019) was reported to be 91.25%, 90.12%, 88.99% and 85.46% correspondingly.

Likewise, different performance outcomes observed for every method. Observed outcomes of TEPEK compared to conventional methods. Overall comparison outcomes evidently prove the performance of data integrity rates of the proposed TEPEK is significantly improved by 8%, 12%, 15% and 23% when compared to SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song & Li, 2021) and RDAC (Zhang *et al.* 2019) respectively. This is due to the application of the Merkle–Hellman Knapsack Cryptosystem based on data storage. The cryptographic technique encrypts the input industrial data as well as it stored securely on CS. Then the cloud server permits data access only for the authorized user. This assists for evading data alteration through unauthorized user resulting in increasing data integrity.

Impact of authentication accuracy (AA)

AA is defined as ratio of number of authorized users properly recognized for data access from server to total number of users. Authentication accuracy is estimated as,

$$AAcc = \left(\frac{Number\ of\ users\ correctly\ identified}{n}\right) * 100 \quad (13)$$

From Eq. (13), '$AAcc$' is authentication accuracy, 'n' is a number of users. It is measured in percentage (%).The outcome of the authentication accuracy using

proposed TEPEK method outperforms existing SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song & Li, 2021) and RDAC (Zhang *et al.* 2019) is explained in Table 2.

Table 2: Comparison of authentication accuracy over the number of users

Number of users	**Authentication accuracy (%)**				
	TEPEK	**SmartCrypt**	**ID-P^{3}DP**	**CSSM Method**	**RDAC**
50	98	94	92	90	88
100	97	93	91	89	86
150	96.66	92.66	89.33	86	84.33
200	96.5	92	89	86	84
250	96.4	91.6	88.8	86	83.75
300	96.33	91.33	88.33	85.33	83.66
350	96	91.14	88	84.46	83
400	95.75	90.75	87.75	84.75	82.4
450	95.55	90.22	87.55	84.88	82.22
500	95	89.6	87	84.4	82

Table 2 indicate performance analysis of authentication accuracy with number of users. As shown in outcomes, various accuracy output are observed for every method. Tabulated outcome represent the proposed TEPEK gives better performance compared to conventional methods. For each method, ten various accuracy results are attained. Average of ten various outcomes indicates the accuracy is improved by 5%, 8%, 12% and 15% when compared to existing SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song &

Li, 2021) and RDAC (Zhang *et al.* 2019) respectively. This is due to the application of Tversky correlative piecewise regression. As a user needs for accessing data from server, first user transmits request message to CS. The password-matching process is performed with the help of the Tversky correlative piecewise regression. When the CU entered password obtain matched through server-sent password, user is authorized as well as permitted for accessing data from CS. This process helps to improve the accuracy of user authentication.

Impact of computation overhead (CO)

CO is measured as amount of time taken by algorithm for secure data storage on cloud server. CO is estimated as follows,

$$CO = [n] * Time\ (SSD) \quad (14)$$

From Eq. (14), 'CO' is the computation overhead, 'n' denotes the number of data, and time taken for storing single data (SSD). It is measured in milliseconds (ms). The outcome of the computation overhead using proposed TEPEK method outperforms existing SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song & Li, 2021) and RDAC (Zhang *et al.* 2019) is explained in Figure 5.

Figure 5 indicates the performance results of computational overhead in terms of time consumption of secured data storage with distinct number of data. TEPEK minimizes the overall performance of computational overhead. The experiment is demeanored for 10000 data, and time consumption of data storage by overhead technique was 4 $2ms$. whereas, the time consumption of data storage using SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song &

Li, 2021) and RDAC (Zhang *et al.* 2019) was found to be 50*ms*, 65*ms* , 75 ms, and 85 ms respectively.

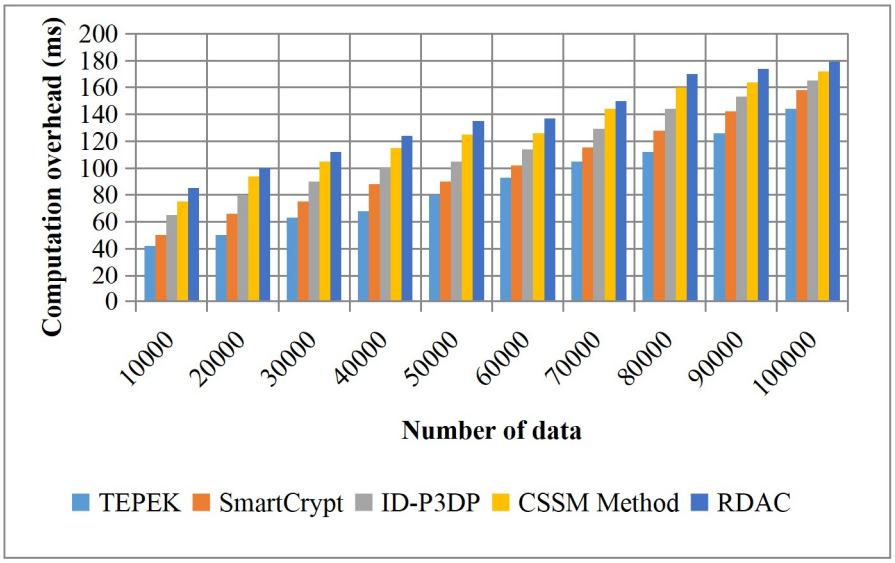

Figure. 5. Graphical illustration of computation overhead

As a result, the TEPEK is compared to observed outcomes of conventional methods. Finally, average value of comparison outcomes denotes the overall computational overhead of TEPEK was reduced by 14%, 25%, 33% and 37% when compared to extisting the existing SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM (Song & Li, 2021) and RDAC (Zhang *et al.* 2019) methods respectively. This is because of CS performing data storage under three processes such as user registration, key generation, as well as data encryption. Cloud server generates a pair of session keys, for each registered user. Then the data encryption process is said to be performed by using Merkle–Hellman Knapsack Cryptosystem. The cloud user performs the Cryptosystem system for storing the user data with minimum time.

Impact of storage overhead (SO)

SO is defined as amount of memory space utilized by server for storing encrypted user data. It is expressed as,

$$SO = n * Mem\,(storing\ one\ data) \quad (15)$$

From Eq. (15), SO is specified as storage overhead, 'n' is a number of data, 'Mem' denotes a memory space. Storage overhead is measured in megabytes (MB).

Table 3: Comparison of storage overhead over the number of data

Number of data	Storage overhead (MB)				
	TEPEK	SmartCrypt	ID-P^{3}DP	CSSM Method	RDAC
10000	57	60	72	82	88
20000	64	72	80	92	99
30000	81	96	102	112	120
40000	100	124	132	140	148
50000	115	120	130	140	150
60000	126	132	138	145	158
70000	140	154	159.6	163.6	170
80000	156	168	176	184	194
90000	162	171	180	189	200
100000	170	180	188	196	208

Table 3 depict comparative analysis of storage overhead using five methods TEPEK, SmartCrypt (Halder & Newe, 2022), ID-P^{3}DP (Ni *et al.* 2022), CSSM

(Song & Li, 2021) and RDAC (Zhang *et al.* 2019), TEPEK gives lesser storage overhead than the existing methods. Besides, while increasing the number of data, space consumption for storing the data gets also increased using all four methods. But comparatively, the figure shows that the storage cost of TEPEK is minimized as compared to other existing works. This is owing to the cloud server using the Merkle–Hellman Knapsack Cryptosystem that stores the encrypted data into a server. The storage overhead of ten results concludes that the performance of TEPEK reduces storage overhead by 9%, 15% , 21% and 25% than the conventional methods.

CONCLUSION

The TEPEK is proposed to industrial IoT via Cloud Computing environment. By means of the Cloud Sim, the proposed method was performed. The major advantage of the proposed method over the methods available in the existing literature is that it helps to reduce the computation overhead and storage overhead. Initially, CU registers detail to CS to store the data. After that, generates the session key pair using Merkle–Hellman Knapsack Cryptosystem for every registered user for storing data in CS. Then CU encrypts data and transmits it into CS. As user wants for accessing data from server, the server performs the Tversky correlated piecewise regression for user authentication. The server permits authorized user for accessing data from server. The performance of the method is analyzed and compared with other conventional methods. The overall quantitative results confirm that TEPEK provide by 13% of improved Data confidentiality rate, 12% of Data Integrity rate and 8% of accuracy, in addition to minimize the 24% of Computational Overhead and 15% minimum storage overhead when compared to conventional methods. So that the proposed structure is more secure. In future, the proposed method is further extended to various

encryption methods and verification techniques to achieve higher excellence in performance.

DATA AVAILABILITY

Dataset name: WUSTL-IIOT-2021 Dataset for IIoT Cyber security Research
Download link: https://www.cse.wustl.edu/~jain/iiot2/ftp/wustl_iiot_2021.zip

COMPETING INTERESTS

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 manuscript.

FUNDING STATEMENT

No funding received for the research

AUTHOR'S CONTRIBUTIONS

Flindon W.G. contributed to the design, execution of the research, analysis of the results and to writing of the paper. Dr. Divya S.V.contributed to the design of research and review of article. All authors reviewed the manuscript.

ACKNOWLEDGMENTS

The authors are thankful to Department of Computer Science, Noorul Islam Centre for Higher Education for supporting the research.

REFERENCES

Ahmed, S. F., Alam, M. S. B., Hoque, M., Lameesa, A., Afrin, S., Farah, T., ... & Muyeen, S. M. (2023). Industrial Internet of Things enabled technologies, challenges, and future directions. *Computers and Electrical Engineering*, *110*, 1-16.

Alex, S., Pattathil, D. P., & Kakkanattu Jagalchandran, D. (2020). SPCOR: a secure and privacy-preserving protocol for mobile-healthcare emergency to reap computing opportunities at remote and nearby. *IET Information Security*, *14*(6), 670-682.

Chi, J., Li, Y., Huang, J., Liu, J., Jin, Y., Chen, C., & Qiu, T. (2020). A secure and efficient data sharing scheme based on blockchain in industrial Internet of Things. *Journal of Network and Computer Applications*, *167*, 1-10. https://doi.org/10.1016/j.jnca.2020.102710

Devi, K. G., & Devi, R. R. (2021). S^2OPE security: Shuffle standard onetime padding encryption for improving secured data storage in decentralized cloud environment. *Materials Today: Proceedings, Elsevier,* (2021) 1-9.

Flavia, B. J., & Chelliah, B. J. (2024). BO-LCNN: butterfly optimization based lightweight convolutional neural network for remote data integrity auditing and data sanitizing model. *Telecommunication Systems*, *85*(4), 623-647.

Guo, C., Tian, P., & Chang, C. C. (2019). Privacy preserving weighted similarity search scheme for encrypted data. *IET Information Security*, *13*(1), 61-69. https://doi.org/10.1049/iet-ifs.2018.5187

Halder, S., & Newe, T. (2022). Enabling secure time-series data sharing via homomorphic encryption in cloud-assisted IIoT. *Future Generation Computer Systems*, *133*, 351-363.

Hasan, M., Ogan, K., & Starly, B. (2021). Hybrid blockchain architecture for cloud manufacturing-as-a-service (cmaas) platforms with improved data storage and transaction efficiency. *Procedia Manufacturing*, *53*, 594-605. https://doi.org/10.1016/j.promfg.2021.06.060

Hoang, T., Yavuz, A. A., & Guajardo, J. (2019). A secure searchable encryption framework for privacy-critical cloud storage services. *IEEE Transactions on Services Computing*, *14*(6), 1675-1689.

Hu, H., Lin, C., Chang, C. C., & Chen, L. (2019). Enhanced secure data backup scheme using multi-factor authentication. *IET Information Security*, *13*(6), 649-658. https://doi.org/10.1049/iet-ifs.2018.5380

Jahan, R., Suman, P., & Singh, D. K. (2021). An algorithm to secure data for cloud storage. *Information Technology In Industry*, *9*(1), 1382-1387.

Khayyat, M. M., Khayyat, M. M., Abdel-Khalek, S., & Mansour, R. F. (2022). Blockchain enabled optimal Hopfield Chaotic Neural network based secure encryption technique for industrial internet of things environment. *Alexandria Engineering Journal*, *61*(12), 11377-11389.

Latif, S., Idrees, Z., Ahmad, J., Zheng, L., & Zou, Z. (2021). A blockchain-based architecture for secure and trustworthy operations in the industrial Internet of Things. *Journal of Industrial Information Integration*, *21*, 1-12.

Malallah, H. S., Qashi, R., Abdulrahman, L. M., Omer, M. A., & Yazdeen, A. A. (2023). Performance analysis of enterprise cloud computing: a review. *Journal of Applied Science and Technology Trends*, *4*(01), 1-12.

Nayak, S. K., & Tripathy, S. (2018). SEPDP: Secure and efficient privacy preserving provable data possession in cloud storage. *IEEE Transactions on Services Computing*, *14*(3), 876-888.

Ni, J., Zhang, K., Yu, Y., & Yang, T. (2020). Identity-based provable data possession from RSA assumption for secure cloud storage. *IEEE Transactions on Dependable and Secure Computing*, *19*(3), 1753-1769.

Ning, J., Cao, Z., Dong, X., Liang, K., Wei, L., & Choo, K. K. R. (2018). CryptCloud $^+ $+: secure and expressive data access control for cloud storage. *IEEE Transactions on Services Computing*, *14*(1), 111-124.

Qi, Q., Xu, Z., & Rani, P. (2023). Big data analytics challenges to implementing the intelligent Industrial Internet of Things (IIoT) systems in sustainable manufacturing operations. *Technological Forecasting and Social Change*, *190*, 122401.

Rafique, A., Van Landuyt, D., Beni, E. H., Lagaisse, B., & Joosen, W. (2021). CryptDICE: Distributed data protection system for secure cloud data storage and computation. *Information Systems*, *96*, 1-23. Ruth, J., Sirmathi, H., & Meenakshi, A. (2019). Secure data storage and intrusion detection in the cloud using MANN and dual encryption through various attacks. *IET Information Security*, *13*(4), 321-329.

Sharma, M., Pant, S., Kumar Sharma, D., Datta Gupta, K., Vashishth, V., & Chhabra, A. (2021). Enabling security for the Industrial Internet of Things using deep learning, blockchain, and coalitions. *Transactions on Emerging Telecommunications Technologies*, *32*(7), 1-19.

Shen, W., Qin, J., Yu, J., Hao, R., Hu, J., & Ma, J. (2019). Data integrity auditing without private key storage for secure cloud storage. *IEEE Transactions on Cloud Computing*, *9*(4), 1408-1421.

Singh, P., Masud, M., Hossain, M. S., & Kaur, A. (2021). Cross-domain secure data sharing using blockchain for industrial IoT. *Journal of Parallel and Distributed Computing*, *156*, 176-184.

Sohal, M., & Sharma, S. (2022). BDNA-A DNA inspired symmetric key cryptographic technique to secure cloud computing. *Journal of King Saud University-Computer and Information Sciences*, *34*(1), 1417-1425.

Song, H., Li, J., & Li, H. (2021). A cloud secure storage mechanism based on data dispersion and encryption. *IEEE Access*, *9*, 63745-63751.

Tian, Y., Zhou, X., Zhou, T., Zhong, W., Li, R., & Yang, X. (2024). A secure certificate-based data integrity auditing protocol with cloud service providers. *Mathematics*, *12*(13), 2-26.

Xiong, S., Ni, Q., Wang, L., & Wang, Q. (2020). SEM-ACSIT: Secure and efficient multiauthority access control for IoT cloud storage. *IEEE Internet of Things Journal*, *7*(4), 2914-2927.

Yang, A., Xu, J., Weng, J., Zhou, J., & Wong, D. S. (2018). Lightweight and privacy-preserving delegatable proofs of storage with data dynamics in cloud storage. *IEEE Transactions on Cloud Computing*, *9*(1), 212-225.

Yu, H., Liu, S., & Fan, Z. (2021). MDupl: A replica strategy of cloud storage system. *Procedia Computer Science*, *188*, 4-17.

Zhang, Z., Chang, C., Guo, Z., & Han, P. (2019). Re-definable access control over outsourced data in cloud storage systems. *IET Information Security*, *13*(3), 258-268.