A Cloud Privacy-Preserving Model using Gaussian mutation-based Bacterial Foraging Optimization and Genetic Crossover Operation

Dr. K. Anand¹ *, Dr. A. Vijayaraj², Dr. M. Vijay Anand³

¹Associate Professor, Department of Information Technology, Chennai Institute of Technology, Kundrathur, Chennai-69, India.
²Associate Professor, Department of Information Technology, Vignan’s Foundation for Science, Technology & Research, Vadlamudi, Guntur, Andhra Pradesh, India.
³Professor, Computer Science and Engineering, Saveetha Engineering College, India

*Corresponding Author: anandk.phd@gmail.com

Abstract:

The necessity of security in the cloud system increases day by day in which the data controllers harvest the rising personal and sensitive data volume. The cloud has some unprotected private data as well as data that has been outsourced for public access, which is crucial for cloud security statements. An advanced legal data protection constraint is required due to the resultant of repeated data violations. While dealing with sensitive data, most of the existing techniques failed to handle optimal privacy and different studies were performed to take on cloud privacy preservation. Hence, the novel model of privacy preservation in the cloud and artificial intelligence (AI) techniques were used to tackle these challenges. These AI methods are insight-driven, strategic, and more efficient organizations in cloud computing. However, the cost savings, agility, higher flexibility businesses are offered with cloud computing by data hosting. Data cleansing and restoration are the two major steps involved in the proposed privacy replica. In this study, we proposed Chaotic chemotaxis and Gaussian mutation-based Bacterial Foraging Optimization with genetic crossover operation (CGBFO-GC) algorithm for optimal key generation. Deriving the multi-objective function parameters namely data preservation ratio, hiding ratio, and modification degree that accomplishes optimal key generation using CGBFO-GC algorithm. Ultimately, the proposed CGBFO-GC algorithm provides more efficient performance results in terms of cloud security than an existing method such as SAS-DPSO, CDNNCS, J-SSO, and GC.

Keywords: Cloud computing, Security, CGBFO-GC algorithm, data cleansing, and data restoration.

1. Introduction:

In the cloud computing environment, on-demand platforms over the internet provide all resources and applications. Due to the privacy and security concerns, the advantages of cloud
computing are taken by several potential cloud customers. A highly significant element is the security feature of some cloud computing communications [1]. The behavior of elasticity and scalability on the fly cloud with the virtual environment is called the cloud, which is widely classified into public, private, and community cloud. Generally, the infrastructure as services, platform as service and software as service are provided by the cloud. When data is saved in a data center, it raises concerns over maintaining customer data security. The cloud computing industry provides significant support to the global environment in a variety of fields, including business, medicine, and defense [2]. Data confidentiality is one of the main significant features of a cloud environment. The cloud data list several kinds of security risks namely user and physical access control, audit scheduling, identity and access management, encryption, and key management. Data storage and processing security are the privacy maintenance aspects of a cloud environment. Both cloud consumer and cloud server consist of identical cloud security in which trust is the vital requirement [3]. From diverse backdrops, several local techniques are embraced using cloud computing. Numerous encryption algorithms carry data encryption in recent days. In the cloud sector, the privacy-preserving develops numerous approaches. The fuzzy Grouping attribute proposes the privacy-aware access control approach for enhanced data privacy in the cloud sector. The efficiency of this verification scheme is still inefficient [4].

The user security and privacy in the cloud environment are protected using the ABSO protocol but it has extremely higher computational complexities. In order to provide security, the perturbation-oriented model modifies the noise data [5]. The careful calibration, on the other hand, necessitated a change to locate the usability for improved stability and model privacy. While dealing with plaintext, the attributes of privacy are tracking with security threats. According to the cloud, numerous numbers of individual data are demoralized and the CSPs are analyzed. Sensitive data handling is simpler if the CSP is considered to be confidential [6]. The above-mentioned problems are solved using numerous approaches in the last decades [7]. Hence, we needed an effective cloud security model to overcome these issues present in the existing techniques. In this study, we have proposed data sensitization and restoration with CGBFO-GC algorithm for optimal key generation in cloud data security model. The main involvement of this paper is summarized as below:

- Chaotic chemotaxis and Gaussian mutation-based Bacterial Foraging Optimization with genetic crossover operation (CGBFO-GC) method is proposed for best key making.
- The data preservation ratio, hiding ratio, and modification degree are the multi-objective functions during cloud data security.
- The proposed method demonstrates optimal and superior results in terms of key sensitivity evaluation, CPA and KPA attack analysis, and Convergence analysis.

The rest of the paper is arranged as: Section 2 explains the existing methods based on the cloud security model. Section 3 delineates the system architecture. The proposed methodology with respect to optimal key generation using the proposed CGBFO-GC algorithm is delineated in
section 4 followed by the investigationalconsequences are discussed in section 5. At last, section 6 concludes the paper.

2. Related works:

Ahmad et al. [8] proposed advancements of artificial intelligence techniques to build up the privacy conservation replica in the cloud environment. Cloud computing provides cost savings, agility, and businesses high flexibility by hosting the data. Data restoration and sanitization are the two major stages of the planned privacy conservation system. According to optimal key generation, the hybrid meta-heuristic algorithm of Jaya-based Shark Smell Optimization (J-SSO) performs the sanitization process. Deriving multi-objective functions including data preservation ratio, modification degree and hiding ratio accomplish optimal key generation. This method provided less space with a higher computational cost. A Reputation-aware Trust and Privacy Preservation (RTPP) scheme was introduced by Ahmad et al. [9] for mobile cloud computing. The leverage cloud services and reputation-aware selection of CCs deals with trust management. To security solutions and key management, a hybrid policy tree mechanism is proposed to select a dynamic attribute. The AS-CABE against security attacks is analyzed to present security analysis. The RTPP with AS-CABE demonstrated higher performances in the case of resilience, trust, computation, storage, reputation score, and encryption and decryption time but the execution time is higher.

The novel non-abelian rings over homomorphic encryption framework was proposed by Li et al. [10]. The matrix-ring proposes homomorphic encryption. The intermediate result of any ciphertext operations is a fast ciphertext homomorphic comparison without decryption. The maximum ciphertext expansion rate required more efficiency according to the encryption scheme. In a trusted environment, improved cloud safety using a crypto-deep neural network was proposed Abirami et al. [11] for privacy preservation. The crypto-deep neural network enhances distributed security (CDNNCS outsourcing scheme. This model consists of a cloud agent, data center, web server, and cloud server. Crypto-deep neural network cloud security (CDNNCS) has targets for managing impersonation attacks. The linear algebraic equation scheme is secured by improving the level of trust between cloud users. The CDNNCS accomplished superior performances in terms of throughput, Jitter, and Delay with minimum packet loss but required a more optimal traffic and security model.

A security vulnerability measurement with cost-optimization was introduced by Park et al. [12] for efficient security enhancement. The security controls considering the limited security budget are enhanced by considering vulnerability-based mitigation and risk assessment. According to security vulnerability measurement, the security cost allocation strategies are evaluated. This model required low-latency resources, data analysis, and various environmental results. For resource allocation with security concerns, Meng et al. [13] suggested security-aware scheduling based on distributed Particle Swarm Optimization (SAS-DPSO). The dynamic workflow model based on a dynamic scheduling mechanism is introduced by means of the mobile industrial applications mobility and dynamics of edge resources. The experimental
investigations accomplished effective balancing between security performances and scheduling performances but it has higher latency resources with poor scalable.

3. **System model and Architecture:**

In the current researches, cloud computing security is a significant factor to be addressed. If the security measures for data transmission and operation are not given correctly, the data will be at risk. The possibility to be elevated risk in data dispensation, and the stored data is evaluated by a number of users as cloud storage provides the ability. To handle security limitations in the cloud, the security challenge and solutions are identified by developing effective security measures. There is a possibility of data misuse when huge organizations are sharing the resources [14]. Hence, the data and data archives are necessarily protected to avoid the risk. For the cloud data, we present a novel cybersecurity model to avoid different data security schemes limitations in the literature.

To analyze the real-time challenges, the proposed cybersecurity model uses cloud data. The proposed cybersecurity model considers the datasets obtained from UCI repositories such as Wholesale customer data, Heart disease, and Air quality data. Data cleansing and data restoration are two major stages of the projected privacy conservation scheme. In a cloud, beating the responsive information or data process is called data cleansing. As a result, exposure to an unauthorised point is avoided. However, an optimal key generation performs the proposed data cleansing process [15]. The proposed Chaotic chemotaxis and Gaussian mutation-based Bacterial Foraging Optimization with genetic crossover operation (CGBFO-GC) improve this process. The consideration of multiobjective function regularizes an optimal key generation that considers the parameters including modification degree, hiding ratio, and preservation ratio of data. The

![Proposed privacy preservation architectural diagram](image-url)
data cleansing and restoration by means of best key making in the cloud are executed by using this multi-objective purpose with CGBFO- GC algorithm.

4. Proposed methodology:

In this section, we discuss data cleansing, data restoration by best key making using the CGBFO-GC algorithm (fig 1).

4.1 Dataset details:

In this study, there are three datasets as wholesale customer data, heart disease data, and air quality data were chosen from the UCI repository related to cloud data. Each dataset is explained as follows:

4.1.1 Wholesale customer data:

Eight attributes with 440 instances present in the wholesale customer dataset (https://archive.ics.uci.edu/ml/datasets/Wholesale+customers) [16]. The Lisbon, Oporto, region, deter gents-paper, delicatessen, milk, frozen, grocery, fresh, etc are examples of the few attributes.

4.1.2 Heart disease data:

There are 303 instances with 75 attributes included in the heart disease dataset (https://archive.ics.uci.edu/ml/datasets/Heart+Disease) [17]. This dataset considers id, age, chol, htn, smoke, and cigs.

4.1.3 Air quality data:

From five metal oxide chemical sensor arrays, there are 9358 instances present in the air quality dataset (https://archive.ics.uci.edu/ml/datasets/Air+Quality) [18]. The sensor is located in the most polluted region of the Italian city. From March 2004 to February 2005 data recorded is chosen that contains different attributes such as relative humidity, data, temperature, absolute humidity, reference analyzer, and time.

4.2 Dataset cleansing and restoration:

Hiding the sensitive information or data in a cloud process is data cleansing. The data from escape on to an illegal point is prevented. Data cleansing is the reverse operation of data restoration. The cleansing process is evaluated here. Fig 2 illustrates the data cleansing and restoration process. The key matrix generation and cloud data determine the binary conversion during cleansing. The optimal key is generated using the CGBFO-GC algorithm. The cleansing data achieves received binary data under the XOR operation. Equation (1) produces the cleansing data from the key matrix generation and original cloud data.
Fig 2: Data cleansing and restoration process

\[ C_{\text{data}} = \hat{C}_{\text{data}} \oplus Key_2 \] (1)

Where \( \hat{C}_{\text{data}} \) is the cleansing data and \( C_{\text{data}} \) is the original data. Here, \( Key_2 \) is an optimally made key. The cleansing is performed using \( C_{\text{data}} \), after cleansing, \( C_{\text{data}} \) is performed as known in the intentionutility of the proposed representation. The cleansing process hides sensitive rules and transferred them to the cloud. The safety in the cloud division devoid of some cyber attack performance is improved and the data will be secluded to further use. The proposed CGBFO-GC algorithm with the same key recovers the original data during restoration. Equation (2) explains the restoration process.

\[ \hat{C}_{\text{data}} = C_{\text{data}} \oplus Key_2 \] (2)

Where, the restored data is \( \hat{C}_{\text{data}} \).

4.3 CGBFO-GC algorithm for optimal key generation:

The major role in cleaning and restoration is key extraction in the proposed cloud data cybersecurity model. In this section, we used the CGBFO-GC algorithm for optimal key generation. The Kronecker method converts the key into a new form. Equation (3) converts the
key into $Key_1$. Where, $\sqrt{Nu^N \times \text{Max}_p}$ considers the size for both $Key_1$ and $Key_2$. Equation (3) explains the key matrix for $Key=5,6,1$.

$$Key_1 = \begin{bmatrix}
5 & 5 & 5 \\
6 & 6 & 6 \\
1 & 1 & 1
\end{bmatrix}$$

Equation (3)

Where, $Nu$ is the number of transactions with the highest score $Nu$. Here, $\text{Max}_p$ is the maximum transaction length. The executing row-wise duplication produces the reconstructed key matrix $Key_1$. The Kronecker method generates the key matrix $Key_2$ as shown in equation (4).

$$Key_2 = key_1 \oplus key_1$$

(4)

The Kronecker product is represented using a symbol $\oplus$. The proposed CGBFO-GC method is used for best key making in the cloud security model. The Chaotic chemotaxis and Gaussian mutation-based Bacterial Foraging Optimization with genetic crossover (CGBFO-GC) algorithm are used in this work. The following steps explain the CGBFO-GC algorithm stages for optimal key generation.

### 4.3.1 Operation of chaotic chemotaxis step length:

The Bacterial Foraging Optimization (BFO) algorithm to change the original set chemotaxis step length [18]. The Gaussian mutation is followed by a genetic crossover operation (GCO) in the chemotaxis step. The optima positions in the current bacterial individuals are operated by allowing chemotaxis step. Equation (5) introduces the chaos theory as shown below:

$$CH_{l+1} = \gamma CH_{l} * (1 - CH_{l}) \quad l = 1, \ldots, R - 1$$

(5)

Where the control parameter is $\gamma$, The bacterial population is $R$. The intervals 0 and 1 locate based on the logistic map and random map. The logistic map generates the chaotic series $CH$. The logistic map contains a higher probability of creating standards near zero and one than the random map between the intervals 0 and 1. The chaotic search is very effectual in local optimization owing to its randomicity and ergodicity [19].

The obtained chaotic chemotaxis step length is sorted using equation (6). The falling into local optimal predicament is prevented via chemotaxis step length.

$$D = \text{sort} (D, \text{descend})$$

(6)

### 4.3.2 Gaussian mutation process:

The mutated position $G_{gb}$ is generated. According to the current bacterial population, the Gaussian mutation is applied towards the best position $G_b$. 
The standard normal distribution is $Gaus(0,1)$. Replace $G_b$ with the fitness of $G_b Gao$ if the mutated position of the fitness functions $G_b Gao$ is better than $G_b$, then update $G_b$ with $G_b Gao$.

Based on $G_b$, the Gaussian distributed random disturbance term $G_b * Gaus(0,1)$ increased using equation (8). The convergence speed is improved that converges to global optima.

4.3.3 Genetic crossover operation (GC):

The key part of the basis bat algorithm (BA) is local searching. For complex tasks, the random walk is enough. The Genetic Algorithm (GA) with BA combination crossover operation is to solve this problem [19]. The novel solution among the most excellent solution and global greatest solution of the current iteration is generated using the crossover process [20]. Equation (8) defines the mathematical representation of the crossover operation.

$$Y_{new} = PG_b * (1 - E) + Y_b * E$$

(Fig 3: Proposed CGBFO-GC algorithm for optimal key generation)
The current iteration with the best solution is $PG$, and the worldwidedgreatest solution is $Y$. Where, the steady is $E$. When the global best solution is unequal towards the most excellent solution of the current iteration then the operation of genetic crossover is preceded [21]. Fig 3 explains the overall procedure of the CGBFO-GC algorithm.

4.4 Cloud cyber security-based objective model:

In this work, the three major objective functions including Modification degree, Hiding ratio, and Data preservation ratio are delineated as follows:

4.4.1 Modification degree:

The modification degree describes the modification degree that occurs between the cleansed dataset $C_{\text{data}}$ and the original dataset $C_{\text{data}}$ in which the Euclidean distance determination is measured. Equation (9) formulate the modification degree [22].

\[ G_1 = C_{\text{data}} - \hat{C}_{\text{data}} \]  

(9)

4.4.2 Hiding ratio

The sensitive rate defines hiding ratio in which is correctly concealed in $C_{\text{data}}$. The difference among the original data of the relevant index considers the term $E$. Equation (10) offers the difference between $E_1$ and $E_2$.

\[ E_{\text{difference}} = xyz(E_1 - E_2) \]  

(10)

The non-zero indexes of $E_{\text{difference}}$ length are $LN_1$. Equation (10) explains the mathematical form of the hiding ratio.

\[ H_{\text{ratio}} = \frac{LN_1}{H_{\text{data}}} \]  

(11)

Where, $H_{\text{data}}$ is the number of data indexes have to hide. The best performance maximizes the hiding ratio.

4.4.3 Ratio of data preservation

The non-sensitive rules rate defines the information preservation ratio that concealing in $H_{\text{data}}$ [23]. Equation (12) represents the ratio of data preservation.

\[ D_{\text{preservation}} = \frac{LN_2}{P_{\text{data}}} \]  

(12)

Where, $LN_2$ is the number of zero indexes and the term $P_{\text{data}}$ preserving the total number of data indexes. The proposed security model maximizes the preservation ratio.
4.4.4 Encode the solution

The data cleansing and restoration is performed by using the proposed CGBFO-GC algorithm for key optimization. The key length is changed according to the number of transactions or data size. Where, \( N_i = 1, 2, \ldots, N_o \). From this, the attribute or field length is \( N_o \). From \( I \) to \( 2^6 - I \), that gives the bounding limit. To generate the best solution, the CGBFO-GC algorithm optimizes the key vectors [24].

4.3.5 Last objective function

The multi-objective function achieves optimal key generation. Equation (13) represents the parameters such as data preservation, hiding ratio, and modification degree [25].

\[
Obj = H_1 + (1-GS) + (1-HS)
\]  
(13)

Where the constant terms are \( GS \) and \( HS \). The data preservation, hiding ratio, and modification degree are \( H_1 \).

Table 1: Parameter settings based on proposed CGBFO-GC algorithm

| Parameters               | Ranges |
|--------------------------|--------|
| Population size          | 10     |
| Maximum number of iteration | 50   |
| Swimming length          | 5      |
| Crossover ratio          | 0.9    |
| Mutation ratio           | 0.1    |

5 Result and Discussion:

The performance of the proposed work is discussed using various performance analyses with a state-of-art comparison. MATLAB 2018 a software that implements the model of proposed cloud data cybersecurity. The virtualization executes extensive data analysis and computational codes are developed easily as the benefits of MATLAB. The complete solution is constructed to leverage different technologies given MATLAB to handle the cloud data. The size of the population is 10 and the maximum numeral of iteration is 50 to experiment [26, 27]. Table 1 explains the parameter settings of the proposed method. The proposed method presentation is evaluated using key sensitivity analysis and state-of-art comparison. In this experiment, we have chosen SAS-DPSO, CDNNCS, J-SSO [8], and GC [21] with the proposed CGBFO-GC algorithm as the state-of-art method. The relevant parameters including hiding ratio, modification degree and ratio of data preservation are considered to obtain the best key making using the proposed CGBFO-GC method. The cleansed data with the original information the modification degree demonstrate the loss of information [28]. According to data preservation, non-hiding other data and hiding sensitive data is to demonstrate the hiding rate effectively. The
optimal key generation with data cleansing and restoration ability is analyzed. All kinds of cloud data security is confirmed using this algorithm.

Fig 4: Modification degree performance analysis, (a) wholesale customer data, (b) heart disease data, and (c) air quality datasets

5.1 Modification degree analysis:

Fig 4 depicts the performance of degree modification analysis using three datasets namely wholesale customer data, heart disease data, and air quality data. The graph is plotted between several iteration and distance. A Euclidian distance between cleansed data and original data is the modification degree as mentioned earlier [29, 30]. While correlated along with the original data, this demonstrates the loss of information that occurred in cleansed data. During the data cleansing process, it ensures that there is no loss of information and the modification degree of the cleansed data is minimum. For all the iterations from 0 to 50, minimal distance than the conventional algorithm produced by the proposed CGBFO-GC algorithm. The proposed
CGBFO-GC algorithm accomplished optimal modification degree results than existing SAS-DPSO, CDNNCS, J-SSO, and GC methods.

5.2 Preservation ratio analysis:

Fig 5 illustrates the reservation ratio performance analysis with respect to wholesale customer data, heart disease data, and air quality datasets are as shown in Fig 5(a) to Fig 5(c). For all the datasets, the proposed CGBFO-GC algorithm demonstrates enhanced ratio of preservation. For various iterations, the proposed CGBFO-GC algorithm demonstrates better results than conventional algorithms. For dataset 1, the CGBFO-GC algorithm accomplishes 0.75%, 0.85% and 2.5%. The conventional methods regarding the relevant preservation data in the cloud outperform all the datasets of the proposed CGBFO-GC algorithm.

![Preservation ratio performance analysis](image-url)

**Fig 5:** Preservation ratio performance analysis, (a) wholesale customer data, (b) heart disease data, and (c) air quality datasets

5.3 Performance analysis based on convergence:
Fig 6 delineates the performance of convergence of cloud security. Fig 6(a) to Fig 6(c) delineates the convergence performance of the proposed method in terms of wholesale customer data, heart disease data, and air quality datasets during cloud security. The convergence value of the proposed CGBFO-GC method accomplishes 1.2%, 0.2%, and 4% values for wholesale customer data, heart disease data, and air quality datasets. However, the proposed CGBFO-GC method demonstrates optimal performances than existing methods such as SAS-DPSO, CDNNCS, J-SSO, and GC methods.

**Fig 6:** Convergence performance analysis, (a) wholesale customer data, (b) heart disease data and (c) air quality datasets

### 5.4 CPA and KPA attacks analysis:

For arbitrary plaintexts, the attack model describes the CPA that presumes ciphertexts are obtained through the attacker. Both encrypted version and plaintext are accessed by an attacker in
which the attack model for cryptanalysis is KPA. Table 2 delineates the different attack effects using various datasets. If the CPA attack is completed then the correlation among original data and restored information is calculated in this section. Minimum connection among restored and original information is obtained when the KPA attack is performed.

KPA attack effect analysis: For whole sale customer data, the SAS-DPSO, CDNNCS, J-SSO, GC and proposed CGBFO-GC methods demonstrated 0.9999%, 1%, 1%, 1% and 0.9979%. Further, SAS-DPSO, CDNNCS, J-SSO, GC and proposed CGBFO-GC methods obtained 0.9969%, 0.9983%, 0.9972%, 0.99979% and 0.9949% for heart disease data. Similarly, the SAS-DPSO, CDNNCS, J-SSO, GC and proposed CGBFO-GC methods provided 0.9998%, 0.9992%, 0.9998%, 0.9999% and 0.9997% for air quality datasets.

CPA attack effect analysis: The CPA attack effect is analyzed using different state-of-art methods including DPSO, CDNNCS, J-SSO, GC, and proposed CGBFO-GC methods. However, the proposed CGBFO-GC method demonstrates 0.998%, 0.9978% and 0.99799% correlation values than previous studies. To secure the cloud data, the proposed CGBFO-GC method is more effective against attacks while contrasted to the other methods.

Table 2: Attack effect analysis using a different dataset

| Name of the attacks |Datasets               | Name of the methods | SAS-DPSO |CDNNCS   | J-SSO    | GC       | CGBFO-GC (proposed) |
|---------------------|-----------------------|---------------------|----------|----------|----------|----------|---------------------|
| KPA                 | Whole sale customer data | 0.9999              | 1        | 1        | 1        | 0.99799             |
|                     | Heart disease data    | 0.99698             | 0.9983   | 0.99972  | 0.99979  | 0.99498             |
|                     | Air quality datasets  | 0.9998              | 0.9992   | 0.9998   | 0.9999   | 0.9997             |
| CPA                 | Whole sale customer data | 1                   | 1        | 1        | 1        | 0.998               |
|                     | Heart disease data    | 0.9993              | 0.99986  | 0.9994   | 0.9993   | 0.99786             |
|                     | Air quality datasets  | 0.9999              | 0.9999   | 0.9999   | 0.9999   | 0.99799             |

5.5 Literature bio-inspired models with its comparative analysis:

Fig 7 delineates the proposed CGBFO-GC based cloud data security model in which the wholesale customer data, heart disease data, and air quality datasets performances are depicted in Fig 7(a) to 7(b). In this experiment, the graph is plotted between a number of iterations and cost functions. The proposed method is 97% better than SAS-DPSO, CDNNCS, J-SSO, and GC models when the iteration is 40 for wholesale customer data. While comparing with the related works, the proposed method demonstrates efficient cloud data security.
Fig 7: Illustration of proposed CGBFO-GC based cloud data security, (a) wholesale customer data, (b) heart disease data and (c) air quality datasets

5.6 Evaluation of key sensitivity:

The novel cloud data security with the evaluation of key sensitivity based on the various datasets are delineated in Table 3. To evaluate the key sensitivity, we execute 15%, 30%, 45% and 60% variations. Minimum variation in the key is obtained and the analysis offered the association between original and restored data. The below table demonstrates key sensitivity results with respect to Wholesale customer data, Heart disease data, and Air quality datasets. However, the proposed CGBFO-GC technique accomplishes superior performances in terms of key sensitivity evaluation.
Table 3: Proposed cloud data security with the evaluation of key sensitivity based on a different dataset

| Name of the methods | Wholesale customer data (%) | Heart disease data (%) | Air quality datasets(%) |
|---------------------|-----------------------------|------------------------|------------------------|
|                     | 15  | 30  | 45  | 60  | 15  | 30  | 45  | 60  | 15  | 30  | 45  | 60  |
| SAS-DPSO            | 0.999 | 1   | 0.996 | 0.990 | 0.997 | 0.995 | 0.994 | 0.63  | 0.997 | 0.999 | 0.984 | 0.993 |
| CDNNCS              | 0.998 | 1   | 0.999 | 0.999 | 0.999 | 0.899 | 0.924 | -0.2  | 0.999 | 0.997 | 0.4002 | 0.1957 |
| J-SSO               | 1    | 0.99 | 0.99 | 1    | 0.906 | 0.985 | 0.619 | 0.834 | 0.7314 | 0.3182 | 0.997 | 0.989 |
| GC                  | 0.999 | 1   | 0.999 | 0.994 | 0.993 | 0.944 | -    | 0.438 | 0.860 | 0.993 | 0.998 | 0.997 | 0.995 |
| CGBFO-GC (proposed) | 0.997 | 0.999 | 0.996 | 0.993 | -    | 0.9905 | -    | 0.530 | 0.395 | 0.954 | 0.993 | -0.245 | -0.434 |

5.7 State-of-art comparison:

The state-of-art comparison of different datasets with secure data rate is delineated in Fig 8. The comparative analysis is performed using MSCryptoNet, J-SSO, and the proposed CGBFO-GC technique. For wholesale data, we have obtained 0.587%, 0.837%, and 0.901% for MSCryptoNet, J-SSO, and CGBFO-GC methods. Further, the MSCryptoNet, J-SSO and CGBFO-GC method demonstrate 0.752%, 0.902% and 0.1293% for heart disease data. Similarly, we have obtained 0.528%, 0.678%, and 0.7340% for MSCryptoNet, J-SSO, and CGBFO-GC methods. However, the proposed CGBFO-GC method demonstrates efficient performance in terms of cloud data security than existing methods.

![Fig 8: State-of-art comparison of the different dataset with a secure data rate](image)

5.8 Computational time analysis:

The state-of-art comparison with respect to computational time analysis is delineated in Fig 9. This graph is plotted between the different methods and computation time in which the
computational time is measured using seconds (sec). The SAS-DPSO, CDNNCS, J-SSO, GC, and proposed CGBFO-GC demonstrate 134.72 sec, 142.38 sec, 129.78 sec, 159.11 sec, and 103.21 sec computational time. When compared to all the existing methods, the proposed CGBFO-GC algorithm provides lower computational time with better speed.

![Fig 9: State-of-art comparison of computational time analysis](image)

### 5.9 Big ‘O’ notation for proposed CGBFO-GC time complexity analysis:

The time complexity is analyzed using the most common metric called Big ‘O’. The execution time requirement is described using the worst-case scenario called Big ‘O’. In this study, \( O(\text{Max}_{ir} \times SP \times 2) \) is the time complexity of the proposed CGBFO-GC algorithm. Where the maximum iteration is \( \text{Max}_{ir} \) and population size is \( SP \). The proposed CGBFO-GC algorithm with its space complexity is \( O(\text{Max}_{ir} \times SP^2) \).

### 6 Conclusion

This paper presented Chaotic chemotaxis and Gaussian mutation-based Bacterial Foraging Optimization with genetic crossover operation (CGBFO-GC) algorithm. For the cloud sector, this paper implements the developed privacy preservation model. The information sanitation and restoration with optimal generation are the major steps involved in the proposed model. The multi-objective parametric function such as hiding ratio, data preservation ratio and modification degree derives the CGBFO-GC algorithm optimizes the optimal key. While dealing with varied key and different attacks, the proposed method contains a fast convergence rate with the ability in solving multi-objective privacy preservation issues. The proposed CGBFO-GC has a lower computation time than existing methods such as SAS-DPSO, CDNNCS, J-SSO, and GC. When compared to the conventional algorithms, the proposed method demonstrates optimal and
superior results in terms of key sensitivity evaluation, CPA and KPA attacks analysis, and convergence analysis.

**ETHICAL STATEMENT:**

*Compliance with Ethical Standards*

*Conflict of interest*

The authors declare that they have no conflict of interest.

*Human and Animal Rights*

This article does not contain any studies with human or animal subjects performed by any of the authors.

*Informed Consent*

Informed consent does not apply as this was a retrospective review with no identifying patient information.

**Funding:** Not applicable

**Conflicts of interest Statement:** Not applicable

**Ethics approval:**

**Consent to participate:** Not applicable

**Consent for publication:** Not applicable

**Availability of data and material:**

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

**Code availability:** Not applicable

**Authors’ contributions**

KA agreed on the content of the study. KA and AV collected all the data for analysis. MVA and KA agreed on the methodology. KA, AV and MVA completed the analysis based on agreed steps. Results and conclusions are discussed and written together. Both author read and approved the final manuscript.

**References:**
1. Begum, R. Sabin, and R. Sugumar. "Novel entropy-based approach for cost-effective privacy preservation of intermediate datasets in cloud." Cluster Computing 22, no. 4 (2019): 9581-9588.

2. Kanwal, Tehsin, Adeel Anjum, and Abid Khan. "Privacy preservation in e-health cloud: taxonomy, privacy requirements, feasibility analysis, and opportunities." Cluster Computing (2020): 1-25.

3. Kaaniche, Nesrine, and Maryline Laurent. "Data security and privacy preservation in cloud storage environments based on cryptographic mechanisms." Computer Communications 111 (2017): 120-141.

4. Ma, Zhuoran, Jianfeng Ma, Yinbin Miao, Ximeng Liu, Kim-Kwang Raymond Choo, Ruikang Yang, and Xiangyu Wang. "Lightweight privacy-preserving medical diagnosis in edge computing." IEEE Transactions on Services Computing (2020).

5. Fang, Chen, Yuanbo Guo, Na Wang, and Ankang Ju. "Highly efficient federated learning with strong privacy preservation in cloud computing." Computers & Security 96 (2020): 101889.

6. Berlato, S., Carbone, R., Lee, A. J., & Ranise, S. (2020, October). Exploring Architectures for Cryptographic Access Control Enforcement in the Cloud for Fun and Optimization. In Proceedings of the 15th ACM Asia Conference on Computer and Communications Security (pp. 208-221).

7. Shankar, K. "Improving the Security and Authentication of the Cloud with IoT using Hybrid Optimization Based Quantum Hash Function." Journal of Intelligent Systems and Internet of Things 1, no. 2 (2021): 61-1.

8. Ahamad, D., Hameed, S. A., & Akhtar, M. (2020). A multi-objective privacy preservation model for cloud security using hybrid Jaya-based shark smell optimization. Journal of King Saud University-Computer and Information Sciences.

9. Ahmad, Waqas, Shengling Wang, Ata Ullah, and Zahid Mahmood. "Reputation-aware trust and privacy-preservation for mobile cloud computing." IEEE Access 6 (2018): 46363-46381.

10. Li, J., Kuang, X., Lin, S., Ma, X. and Tang, Y., 2020. Privacy preservation for machine learning training and classification based on homomorphic encryption schemes. Information Sciences, 526, pp.166-179.

11. Abirami, P., and S. Vijay Bhanu. "Enhancing cloud security using crypto-deep neural network for privacy preservation in trusted environment." Soft Computing 24, no. 24 (2020): 18927-18936.

12. Park, Jun-Young, and Eui-Nam Huh. "A Cost-Optimization Scheme Using Security Vulnerability Measurement for Efficient Security Enhancement." Journal of Information Processing Systems 16.1 (2020).

13. Meng, Shunmei, Weijia Huang, Xiaochun Yin, Mohammad R. Khosravi, Qianmu Li, Shaohua Wan, and Lianyong Qi. "Security-aware dynamic scheduling for real-time
optimization in cloud-based industrial applications." IEEE Transactions on Industrial Informatics (2020).
14. Sugumaran, M., B. Bala Murugan, and D. Kamalraj. "An architecture for data security in cloud computing." In 2014 World Congress on Computing and Communication Technologies, pp. 252-255. IEEE, 2014.
15. Gupta, I., N. Singh, and A. Kumar. "Layer-based privacy and security architecture for cloud data sharing." Journal of Communications Software and Systems 15, no. 2 (2019): 173-185.
16. Cardoso, M.G.M.S. (2013). Logical discriminant models â€“ Chapter 8 in Quantitative Modeling in Marketing and Management Edited by Luiz Moutinho and Kun-Huang Huarng. World Scientific. p. 223-253. ISBN 978-9814407717
17. Wlodzislaw and Rafal Adamczak and Krzysztof Grabczewski and Grzegorz Zal. A hybrid method for extraction of logical rules from data. Department of Computer Methods, Nicholas Copernicus University.
18. De Vito, E., Massera, M., Piga, L. Martinotto, G. Di Francia, On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario, Sensors and Actuators B: Chemical, Volume 129, Issue 2, 22 February 2008, Pages 750-757, ISSN 0925-4005.
19. Rajasekar, N., Neeraja Krishna Kumar, and Rini Venugopalan. "Bacterial foraging algorithm based solar PV parameter estimation." Solar Energy 97 (2013): 255-265.
20. Xu, Yueting, Huiling Chen, Ali Asghar Heidari, Jie Luo, Qian Zhang, Xuehua Zhao, and Chengye Li. "An efficient chaotic mutative moth-flame-inspired optimizer for global optimization tasks." Expert Systems with Applications 129 (2019): 135-155.
21. Yue, Xiaofeng, and Hongbo Zhang. "Modified hybrid bat algorithm with genetic crossover operation and smart inertia weight for multilevel image segmentation." Applied Soft Computing 90 (2020): 106157.
22. Manogaran, Gunasekaran, Chandu Thota, and M. Vijay Kumar. "MetaCloudDataStorage architecture for big data security in cloud computing." Procedia Computer Science 87 (2016): 128-133.
23. Talib, A.M., Atan, R., Abdullah, R. and Murad, M.A.A., 2010. Security framework of cloud data storage based on multi agent system architecture: Semantic literature review. Computer and Information Science, 3(4), p.175.
24. Ramasamy, S., and R. K. Gnanamurthy. "Cluster Based Multi Layer User Authentication Data Center Storage Architecture for Big Data Security in Cloud Computing." Journal of Internet Technology 21, no. 1 (2020): 159-171.
25. Sharma, S., Mishra, A., & Singhai, D. (2020). Secure cloud storage architecture for digital medical record in cloud environment using blockchain.
26. Thakare, Vaishali Ravindra, and K. John Singh. "Cloud Security Architecture Based on Fully Homomorphic Encryption." In Architecture and Security Issues in Fog Computing Applications, pp. 83-89. IGI Global, 2020.
27. Ghani, Anwar, Afzal Badshah, Saeedullah Jan, Abdulrahman A. Alshdadi, and Ali Daud. "Cloud storage architecture: research challenges and opportunities." arXiv preprint arXiv:2004.06809 (2020).

28. Kumar, Y. Kiran, and R. Mahammad Shafi. "An efficient and secure data storage in cloud computing using modified RSA public key cryptosystem." International Journal of Electrical and Computer Engineering 10, no. 1 (2020): 530.

29. Malik, Anju, Mayank Aggarwal, Bharti Sharma, Akansha Singh, and Krishna Kant Singh. "Optimal Elliptic Curve Cryptography-Based Effective Approach for Secure Data Storage in Clouds." International Journal of Knowledge and Systems Science (IJKSS) 11, no. 4 (2020): 65-81.

30. Murugesan, Arun, Balamurali Saminathan, Fadi Al-Turjman, and R. Lakshmana Kumar. "Analysis on homomorphic technique for data security in fog computing." Transactions on Emerging Telecommunications Technologies (2020): e3990.