International Journal of Computer and Information System (IJCIS)
Peer Reviewed – International Journal
Vol : Vol. 03, Issue 03, August 2022
e-ISSN : 2745-9659
https://ijcis.net/index.php/ijcis/index

A Scheduling Genetic Algorithm For Real-Time Data Freshness And Cloud Data Security Over Keywords Searching

1stDivya Surendran, 2ndSasikala K
[1]Research Scholar, Department of Computer Science
[2]Research Supervisor, Department of Computer Science and Engineering
[1]Vinayaka Missions Research Foundation, Salem, India
[2]Vinayaka Missions Kirupananda Variyar Engineering College, Salem, India
[1]surendrdivya@gmail.com, [2]sasijk78dr@gmail.com

Abstract— *Cloud storage services allow customers to ingress data stored from any device at any time. The growth of the Internet helps the number of users who need to access online databases without a deep understanding of the schema or query. The languages have risen dramatically, allowing users to search secured data and retrieve desired data from cloud storage using keywords. On the other hand, there are fundamental difficulties such as security, which must be provided to secure user'spersonal information. A hybrid scheduling genetic algorithm (SGA) is proposed in this research. SGA technique enhances the security level and provides data freshness. For evaluation and comparison, parameters such as execution time throughputs are used. According to experimental results, the proposed technique ensures the security of user data from unauthorized parties. Furthermore, SGA is strong and more effective when compared to a set of parameters to the existing algorithm like Data Encryption Standard (DES), Blowfish, and AdvancedEncryption Standard (AES).*

Keywords : *Cloud Storage, Keyword Search, Security Issue, Scheduling Genetic Algorithm, Data Freshness.*

I. INTRODUCTION

Recently, cloud computing has replaced cloud storage that serves as anonline infrastructure for consumers. Self-service on-demand, broad network access, resource pooling, rapid elasticity, measurable service are some of the qualities and benefits of cloud computing that entice organizations, businesses, and people to employ cloud services. The preservation of outsourced data's privacy is a severedifficulty to the growth of cloud storage services.Individuals can submit their personal information, such as email addresses, personal well-being records, and financial information, to the cloud to share with others or use on their own in any location. Cloud computing provides a great deal of ease to users. It does, however, bring with it a slew of security concerns. Among the most significant security disadvantages of cloud computing is that the data owner may not entirely respect the cloud server and no longer control their data. Many approaches have been presented to secure cloud data privacy and security [1-3].

With the development of keyword-based searching, the work has gotten a lot simpler, and the required file can now be found quickly by using phrases related to it. The traditional keyword search strategies were limited to exact keyword searches. Various researchers have recently suggested the use of fuzzy keyword searching. The term is available for downloading an encrypted file even if it is misused while maintaining the keyword's privacy. A fuzzy keyword search system based on wild cards has been developed [4] in a semi-trusted server.

A searchable public-key encryption system that is securefrom inside keyword guessing attacks. [5] Achieves (single) ciphertext (CI) indistinguishability and (single) trapdoor indistinguishability security requirements (TI). Any data owner may feasibly encrypt data and its keywords employing the receiver's public key in Public-key Encryption along Keyword Search (PEKS). The receiver can produce the trapdoor and direct it to the server via a safe connection to search for encrypted data. [6] The secure channel assumption is removed using a PEKS program with a delegated tester (dPEKS). In dPEKS, a ciphertext is enciphered under receiver's and the specified server's public key. The specified server alone can exploreendlesslyenciphered data utilizing the trapdoor and its secret key, even though the trapdoor is publicly exposed. Nevertheless, these two techniques are both subject to off-line Keyword Guessing Attacks (KGA) [7]. i.e., provided with a trapdoor, anyone may quickly determine whether or not a keyword was accustomed to creating the trapdoor. This attack will be particularly effective if the entropy of the keyword space is low. In fact, in the public-key setup of PEKS/dPEKS, it is intrinsically challenging to ensure keyword guessing attack protection from an inside attacker. Because anyone may construct a ciphertext containing a guess at a keyword in the PEKS/dPEKS schemes, the malware server can consistently receive the test consequence and estimate the keyword.

This research proposes a hybrid scheduling genetic algorithm (SGA). The deferrable technique enhances the security level, and the genetic algorithm provides freshness to the data. Furthermore, the SGA has strong and more effective performances than existing algorithms like DES, Blowfish, AES.Theimplementation timing of encryption, decryption and throughput computations is contemplated when evaluating the proposed SGA model.The remaining portion of the paper is structured as follows: Section 2 explains the literature survey,while Section 3 delves into the proposed technique.Section.4 illustrates the security analysis, and results of the exploratory phase are discussed in Section 5. The exposition is wrapped up in the sixth area.

II. LITERATURE SURVEY

Miao et al. [8] presented the primitive of verifiable database (VDB) that helps forward secure keyword searches, and the symmetric searchable encryption (SSE) cannot be adopted to VDB. According to the results, the VDB method has attained the requisite security features with excellent efficiency.Liu et al. [9] bestowed an SPKS (safe and privacy-preserving keyword search) decryption technique that decreases computational and communication overhead for consumers while maintaining user data and querying privacy. The result showed that the SPKS was more felicitous to a cloudenvirons.Najafi et al. [10] obtainedtherFSMSE increased the complexity of time complexity, memory, and communication by allowing searching multiple terms, which was more secure and search time complexity. The result demonstrates that the rFSMSE required less time to search many keywords than the existing randomized symmetric searchable encryption (SSE).Pandiaraja et al. [11] presented the apriori technique is based on secure inner product computation, which decreases the time it takes to access data in the cloud. The result demonstrates that the cloud server securely searches the data. It has fewer computational requirements in configuration, trapdoor generation, and queries.Qin et al. [12] introduced a public key authenticated encryption with keyword search (PAEKS), thatcaptures inner keyword assaults and captures multi-keyword attacks from the outside. The implementation result shows comparable efficiency with previous PEKS/PAEKS schemes.Tariq et al. [13] presented asecure keyword-based search approach that allows the client to safely keep his informations using wild-card techniques. The fuzzy keyword searching scheme seeks and recovers encrypted data. The result shows that the introduced method enhances the security system and protects confidential client information against unwanted disclosure.Ogata et al. [14] presented internal security by maintaining encryption, previously employed in secure kNN systems and secured ranking keyword search algorithms.The result shows that one was ineffective, and the first encryption function was breakable.A security mechanism for conjunctive keyword search systems, including trapdoor security, was provided by Byun et al. [15]. The result shows that the introduced method analyzed their weakness and countermeasure.Pakniat et al. [16] presented thesecurity in the enhanced security mechanism of Certificate Less Authenticated Encryption with Keyword Search (CLAEKS). This method allowed to upload the ciphertextsto the cloud with more privacy, searchability, sharability, and authentication all at the same time. The result shows that the CLAEKS scheme was secured in the magnified security mechanism.The difficulties of computation and communication denote that safety was accomplished at a reasonable fare.Zhang et al. [17] presenteda secure and efficient SEPSE that provided the key resumption to restore aprevailing keyand a new one on a key server to the art the key comparison. The result shows that the security analysis and performance of SEPSE provided more robust security.

III. PROPOSED METHODOLOGY

Cloud computing entrusts to a collection of hardware and software used to provide various computing services. The cloud maintains services for delivering software, infrastructure, and platforms over the Internet as users require. Cloud computing is essential in the IT sector because it allows users to access services from anywhere in the world. As the need for and popularity grows, so do many threats and vulnerabilities. Data security is a primary concern in cloud computing, and it must be taken into consideration because data is kept in multiple locations. Multi-keyword search and verifiability are two significant features of searchable encryption. This research introduced a new and robust security framework: scheduling genetic algorithm (SGA). The proposed algorithm is used for data security in the cloud, and also we introduce a Deferrable scheduling algorithm that maintains the data freshness problem.

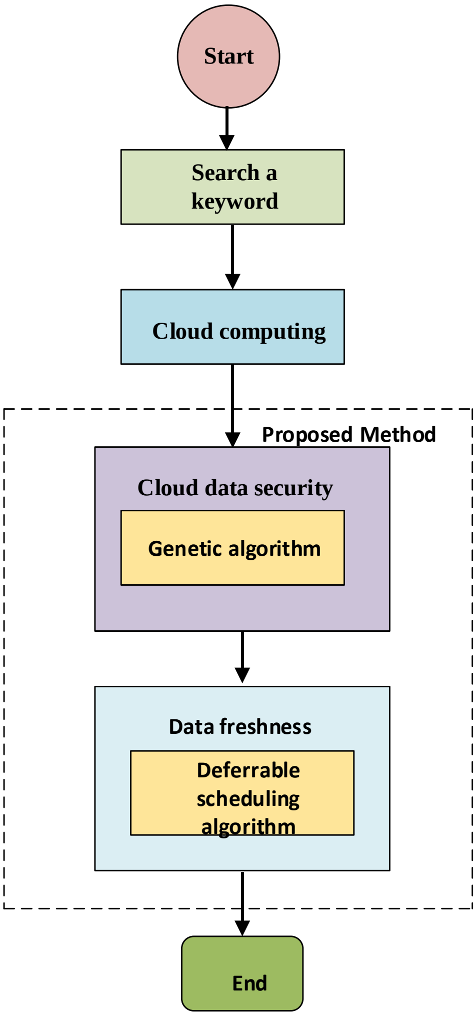

Figure 1: Proposed Methodology

The above diagram explains the overall process of the proposed keyword search security method. In this process, input is a query word that is the keyword for what the user wishes to search. First, the user is asking the request query to the server. The searching keyword is encrypted data, and the system enters the cloud storage. By resolving security concerns as quickly as feasible, the upgraded genetic algorithm (GA), which is utilized for cloud data safety, decreases the massive computing complexities.Thedeferrable scheduling mechanism maintains the real-time data freshness.

3.1 Proposed Scheduling Genetic Algorithm(Sga)

In computer applications, GA is utilized for restricted and unrestricted optimization and security problems. In order to solve NP-hard problems, GA resolves optimization problems in the shortest amount of time possible [18]. This reduces the massive complexity of computing. Population creation, crossover, and mutations are GA's three main roles. As the design of the approach differs from the traditional security algorithms in crossover and mutation, increasing the security level, this leads to a more intricate and complex mapping between the input and output.

3.2 Key Generation

An alphabetic and special characters letter series produced by a stochastic mechanism makes up the initial chromosomal population. 200 persons make up the initial population, and each chromosome has 16 characters encoded in 128 bits. Everyone is sent individually to the fitness function through a loop. Because the fitness function is maximum, the individual with the highest fitness value will be chosen for further analysis. [19] After that, the procedure chooses two people and performs byte-wise one-point crossover, with the crossover point determined by a random number. After performing crossover, acquire the offsprings of the selected individuals. The outcome of previous phases serves as the input for the mutation process. Following mutation, the final key used for encryption is obtained.The steps in the critical generation process are as follows:

INITIAL POPULATION GENERATION:

An experimental population of 200 chromosomes of 16 characters each, with alphabetic and specialized symbols recorded as 8 bits per character, making each chromosome 128 bits long, is being produced using the randomized function.

FITNESS CALCULATION:

The Shannon Entropy (H(X)) is used to calculate each individual's fitness value. It compares the level of randomness in the end population set of data to initial population using eq.1.

$$H(X) = \sum_{i=1}^{n} P(x_1)\, log_2 P\,(x_1) \quad (1)$$

Here P is the likelihood of every chromosomal feature being evaluated.The greater the entropy, the more challenging it is to penetrate.

CROSSOVER:

A random value is used to accomplish byte-by-byte single-point crossover on specified chromosomes; means that as parents, two 128-bit chromosomes are selected, and a haphazardly created value between 1 and 8 is used for crossover operation in each byte to produce an offspring.

MUTATION

The freshly produced child chromosome undergoes byte-wise modification based on a random number generated between 1 and 8.

If the number of iterations is less than or equal to 100, the preceding steps are repeated until the stopping requirements are met.In each iteration, the value of the individual with the highest level of fitness is recorded. The chromosome with the greatest fitness value is selected as encryption key if stopping condition is met.

Next, we show how to use our deferrable scheduling technique to keep data fresh. $D_i$isestablished by the responding timing of the first job, that is the worst-case response timing of all T_i jobs. ML is pessimistic on the deadlines and periods assignment because it employs a periodic task concept with a fixed period and deadline for every task, and the deadline is comparable to the worst-case response time.It should be observed that as long as $P_i + D_i \leq V_i$The validity requirement can always be satisfied.

Given release time $r_{i,j}$ of job $J_{i,j}$ and the deadline $d_{i,j+1}$ of job $J_{i,j+1}(j \geq 0)$,

$$d_{i,j+1} = r_{i,j} + V_i \quad (2)$$

Makes sure that the constraint on validity will be met.

As a result, the following equation is derived promptly from equation (2)

$$(r_{i,j+1} - r_{i,j}) + (d_{i,j+1} - r_{i,j+1}) = V_i \quad (3)$$

If $r_{i,j+1}$ is moved onward to $r'_{i,j+1}$Along the timeline. The deferral of job $J_{i,j+1}$ release time diminishes the relative deadline of the job if it's an utter deadline. Hence its relative deadline, $D_{i,j+1}$ becomes $d_{i,j+1}$-$r_{i,j+1}$, that is less than $d_{i,j+1}$-$r_{i,j+1}$. The deadline of $J_{i,j+1}$subsequent job $J_{i,j+2}$could be additionally prorogued to ($r_{i,j+1}$+V_i) to gratify the validity constraint. Therefore, the processor application for completion of three jobs, $J_{i,j}$, $J_{i,j+1}$and $J_{i,j+2}$ then becomes $\frac{3C_i}{2V_i-(d_{i,j+1}-r'_{i,j+1})}$ (4)

It's lower than the utilization rate $\frac{3C_i}{2V_i-(d_{i,j+1}-r'_{i,j+1})}$(5)needed for the same amount of work to be completed.

IV. SECURITY ANALYSIS

Regardless of these preferences, there are various security risks associated with cloud services; as a result, many organizations and customers are hesitant to employ cloud services. [20] Demonstrates one of the many security flaws that exist.

DATA BREACHES

It occurs when the client's highly confidential and classified data stored in the cloud gets stolen, viewed, and presented to unauthorized elements [21]. It contains trade secrets, bank account information, and other information.

INSIDER ATTACKS (THREAT)

This authorized employee of an organization takes advantage of his privileges to gain access to confidential client information such as account numbers, budgetary structures, and so on. Most organizations do not devote much attention to this attack because they primarily focus on external threats [22].

DENIAL OF SERVICE ATTACK

It targets the cloud network's availability service. The attacker tries to overwhelm the cloud foundation, system, and services such that accredited users cannot access them [23].

DATA LOSS

Data loss can occur due to malicious attacks or catastrophic events such as natural disasters. It can also happen due to a lack of a recovery plan, mismanagement, and incorrect cloud data management [24].

MALWARE INFUSION

It occurs when cloud services are created utilizing code or scripts that function as a "valid instance" and run as SaaS on a cloud server. [25]. It appears to carry out a well-coordinated operation; yet, it aids hackers in monitoring, listening in, and stealing personal information.

V. RESULT

We carried outvarious experiments to calculate the proposed hybrid method SGA's performance under various algorithms. Specifically, the data security evaluation is addressed in the experimental setup, as depicted in figure 2. Effective data security architecture should overcome all of the potential challenges associated with cloud computing, allowing the benefits of cloud computing to reach their full potential. The proposed model is compared to other data security frameworks in Figure 2.

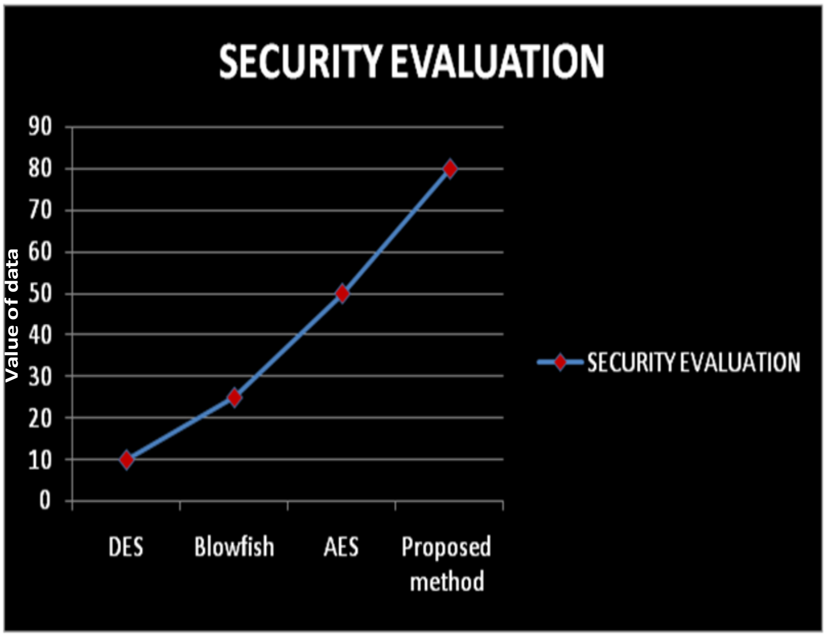

Figure 2: Security Evaluation

The comparison of the postulated model reveals that it addresses the majority of potential security concerns by providing functions and strategies that are adequate for dealing with cloud storage security concerns.

The execution duration for all tests is measured in seconds, and the throughput effectiveness for both encryption and decryption is measured in bytes per second. SGA enhancementaboveothers iscalculated inefficiency percentage for consistency and clarity. Furthermore, during experimental assessment from datasets, the throughput efficient performance of encryption and decryption procedures is noted and addressed.Following a thorough review of the literature, it has been determined that several state-of-the-art methodologies exist, i.e., DES, AES, and Blowfish, which perform well in a keyword cloud condition. As a result, these algorithms have been chosen for comparison. Table 1 shows the encryption of time analysis the efficiency of the SGA algorithm.

DES	AES	Blowfish	Proposed method
5.321	4.269	3.658	1.487
5.457	4.466	3.789	1.541
5.558	4.578	3.884	1.625
6.75	5.69	4.487	2.238
8.147	9.147	7.856	2.472

Table 1: Encryption of time analysis

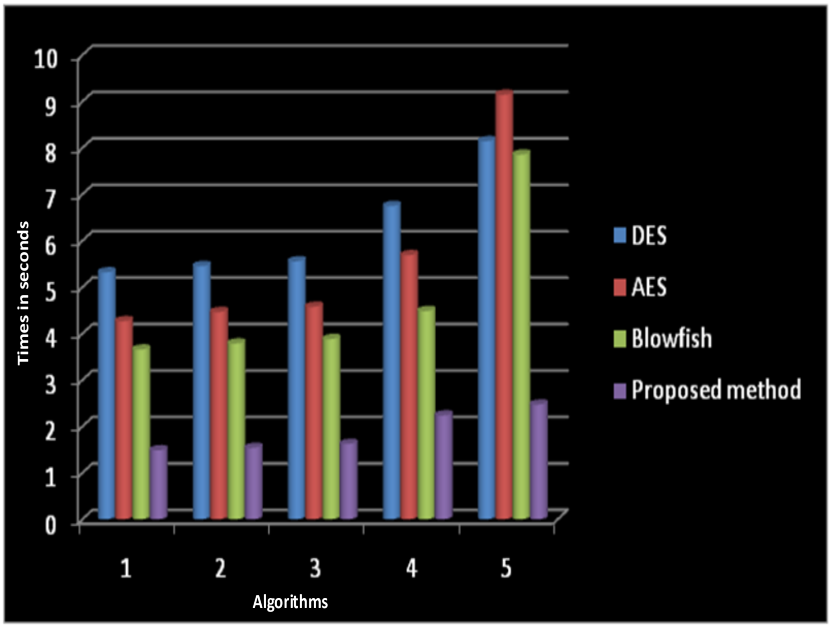

Figure 3: encryption of time comparison

Time (seconds) along the x-axis and data sizes are the parameters taken into account (MB) beside the y-axis, and the encryption time is contrasted (pic 2 and table 1). The result analysis demonstrates that the introduced method SGA is higher than DES, AES, and Blowfish.

DES	AES	Blowfish	Proposed method
5.875	4.874	4.235	1.247
6.045	5.023	4.245	1.774
6.14	5.18	4.188	2.45
6.126	5.104	4.688	2.441
8.004	7.003	5.206	2.841

Table 2: Decryption of time analysis

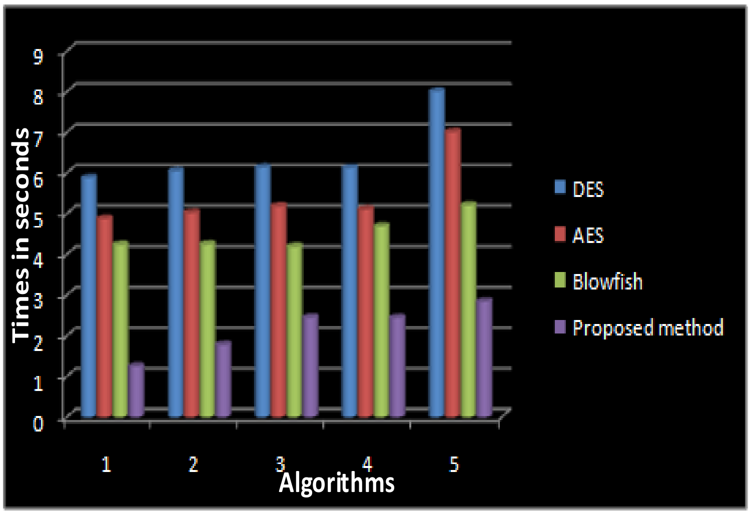

Figure 4: Decryption of time comparison

Table 2 and figure 3 show the decryption of time analysis a maximum decryption time of 500 runtime implementations for the dataset. The result shows that the introduced methodologydiminishes the amount of time spent on all datasets compared to others.

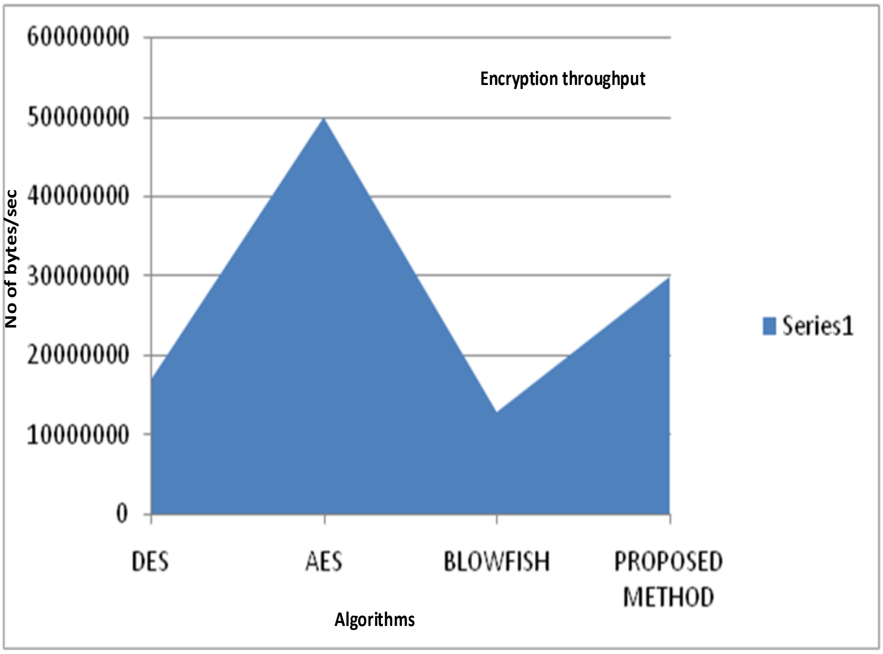

Figure 5: Average Encryption Throughput Comparison

Fig 4 displays the encryption throughput efficacy of the postulated model SGA and other methodologies. SGA has greater throughput efficiency than the others, according to the investigation.

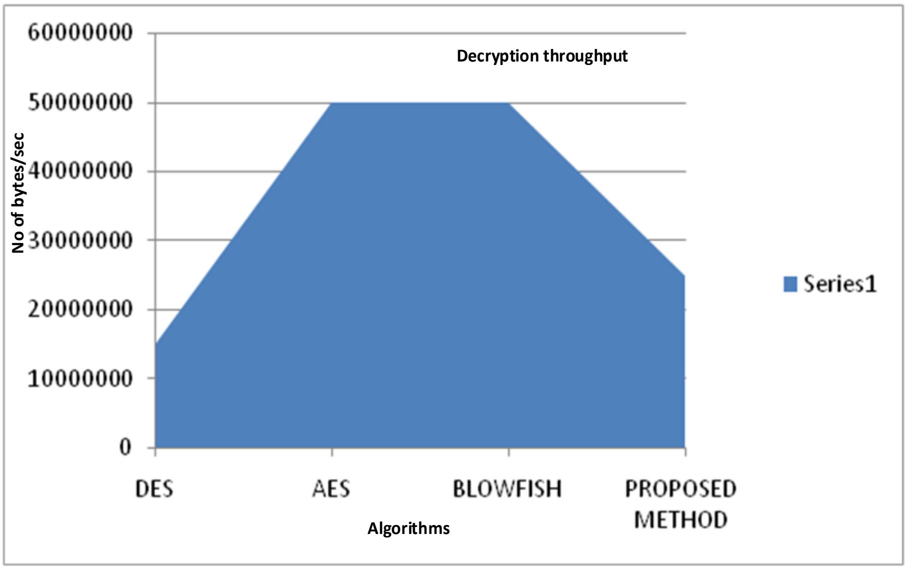

Figure 6: Average Decryption Throughput Comparison

Figure 5 depicts the proposed model SGA's and other techniques decryption throughput efficiency. Its study reveals that the postulated model SGA has greater throughput efficacy than others.

VI. CONCLUSION

Cloud computing is a new discipline of computer science that uses the Internet to deliver quick and efficient services. Data security is always a significant issue for cloud computing. Different strategies and algorithms are employed to assure data security, yet there is a space that must be filled. This research proposes a hybrid scheduling genetic algorithm for real-time data freshness and cloud data security over keywords search. The GA algorithm has been postulated for cloud data security. It is effortless to implement with only two main crossover and mutation procedures. GA is a network security technique that can be employed in cloud data security. In addition, a Deferrable scheduling algorithm is used to maintain freshness to the data. The SGA algorithm is also utilized in the key generation, encryption, and decryption of time processes. In contrast to the old traditional algorithm, i.e., DES, AES, and blowfish output evaluation demonstrated that the introduced method offered a faster execution time and higher throughput when executing encryption and decryption.Finally, our comparison outcomesprove that the SGA algorithm is more secure than other DES, Blowfish, AES.

REFERENCE

1. J. Yu, H. Wang, Strong key-exposure resilient auditing for secure cloud storage, IEEE Trans. Inf. Forensics Secure. 12 (8) (2017) 1931–1940.
2. Y. Yu, Y. Li, J. Tian, J. Liu, Blockchain-based solutions to security and privacy issues in the Internet of things, IEEE Wirel. Commun. 25 (6) (2018) 12–18.
3. J. Li, D. Lin, A.C. Squicciarini, J. Li, C. Jia, Towards privacy-preserving storage and retrieval in multiple clouds, IEEE Trans. Cloud Comput. 5 (3) (2017) 499–509.
4. Shekokar N, Sampat K, Chandawalla C, Shah J (2015) Implementation of fuzzy keyword search over encrypted data in cloud computing. In: Proceedings of international conference on advanced computing technologies and applications (ICACTA-2015) held in Mumbai, India, ISBN: 978-1-5108-0136-3, vol 45. pp 499–505
5. Q. Huang, H. Li, An efficient public-key searchable encryption scheme secure against inside keyword guessing attacks, Inform. Sci. 403 (2017) 1–14.
6. J. Baek, R. Safavi-Naini, W. Susilo, Public key encryption with keyword search revisited, in: O. Gervasi, B. Murgante, A. Laganà, D. Taniar, Y. Mun, M.L. Gavrilova (Eds.), Computational Science and Its Applications - ICCSA 2008, International Conference, Perugia, Italy, June

30, - July 3, 2008, Pro- ceedings, Part I, Lecture Notes in Computer Science, volume 5072, Springer, 2008, pp. 1249–1259.

7. J.W. Byun, H.S. Rhee, H. Park, D.H. Lee, Off-line keyword guessing attacks on recent keyword search schemes over encrypted data, in: W. Jonker, M. Petkovic (Eds.), Secure Data Management, Third VLDB Workshop, SDM 2006, Lecture Notes in Computer Science, volume 4165, Springer, 2006, pp. 75–83.
8. Miao, Meixia, Yunling Wang, Jianfeng Wang, and Xinyi Huang. "Verifiable database supporting keyword searches with forward security." *Computer Standards & Interfaces* 77 (2021): 103491.
9. Liu, Qin, Guojun Wang, and Jie Wu. "Secure and privacy preserving keyword searching for cloud storage services." *Journal of network and computer applications* 35, no. 3 (2012): 927-933.
10. Najafi, Aniseh, Hamid Haj Seyyed Javadi, and Majid Bayat. "Efficient and dynamic verifiable multi-keyword searchable symmetric encryption with full security." *Multimedia Tools and Applications* (2021): 1-20.
11. Pandiaraja, P., and P. Vijayakumar. "Efficient multi-keyword search over encrypted data in untrusted cloud environment." In *2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM)*, pp. 251-256. IEEE, 2017.
12. Qin, Baodong, Yu Chen, Qiong Huang, Ximeng Liu, and Dong Zheng. "Public-key authenticated encryption with keyword search revisited: Security model and constructions." *Information Sciences* 516 (2020): 515-528.
13. Tariq, Husna, and Parul Agarwal. "Secure keyword search using dual encryption in cloud computing." *International Journal of Information Technology* 12, no. 4 (2020): 1063-1072.
14. Ogata, Wakaha, and Takaaki Otemori. "Security analysis of secure kNN and ranked keyword search over encrypted data." *International Journal of Information Security* 19, no. 4 (2020): 419-425.
15. Byun, Jin Wook, and Dong Hoon Lee. "On a security model of conjunctive keyword search over encrypted relational database." *Journal of Systems and Software* 84, no. 8 (2011): 1364-1372.
16. Pakniat, Nasrollah, Danial Shiraly, and Ziba Eslami. "Certificateless authenticated encryption with keyword search: Enhanced security model and a concrete construction for industrial IoT." *Journal of Information Security and Applications* 53 (2020): 102525.
17. Zhang, Yuan, Chunxiang Xu, Jianbing Ni, Hongwei Li, and Xuemin Sherman Shen. "Blockchain-assisted public-key encryption with keyword search against keyword guessing attacks for cloud storage." *IEEE Transactions on Cloud Computing* (2019).
18. Naresh, R., Sayeekumar, M., Karthick, G., Supraja, P.: Attributebased hierarchical file encryption for efficient retrieval of files by dv index tree from cloud using crossover genetic algorithm. Soft Comput. 23(8), 2561–2574 (2019)
19. Tahir, Muhammad, Muhammad Sardaraz, Zahid Mehmood, and Shakoor Muhammad. "CryptoGA: a cryptosystem based on genetic algorithm for cloud data security." *Cluster Computing* 24, no. 2 (2021): 739-752.
20. Ahmed M, Hossain MA (2014) Cloud computing and security issues in the cloud. Int J Netw Sec Appl (IJNSA) 6(1):25–36
21. Barona R, Anita EAM (2017) A survey on data breach challenges in cloud computing security: issues and threats. In: International conference on circuit, power and computing technologies (ICCPCT-2017) held in Kollam, India, ISBN: 978-1-5090-4967-7. pp 20–21
22. Yusop ZM, Abawajy JH (2013) Analysis of insiders attack mitigation strategies. In: International conference on innovation, management and technology research held in Malaysia, vol 129. pp 611–618
23. Carlin A, Hammoudeh M, Aldabbas O (2015) Defense for distributed denial of service attacks in cloud computing. In: International conference on advanced wireless, information, and communication technologies (AWICT-2015), vol 73. pp 490–497
24. Wong R (2017) Research on data security technology based on cloud storage. In: 13th global congress on manufacturing and management, (GCMM-2016). pp 1340–1355
25. Watson MR, Shirazi NH, Marnerides AK, Mauthe A, Hutchison D (2016) Malware detection in cloud computing infrastructures. IEEE Trans Dependable Secure Comput 13(2):192–205