Techniques for secure distributed systems

Huda H Ali 1,2, Prof. Assist. Dr. Shaimaa H Shaker3
1 Al- imam Al-Kadhum College (IKC).
2 Iraqi Commission for Computer & Informatics-Information Institute for Postgraduate Studies
3University of Technology.
Corresponding author’s e-mail address huda.hamdan33@gmail.com

Abstract. In recent networked world, more of the times computers work not isolation. They work with each other in order to satisfy communication purpose, processing, transfer of data, saving, etc., the systems can be described as distributed when it contain several systems work together in harmonic way with each other that are geographically lies over far distance. The security of distributed systems is very important issues and it may be describe as the most difficult part in the design for a system. In this paper the basic principles of distributed system is presented, then the security vulnerabilities are shown and the recent technique that used for solving the security problems are discussed and also analysis for their performance is done where the advantages and disadvantages with robustness for each technique are shown. So as conclusion for the results that are gotten by researchers, the selection of security technique depends on what really need to be protected, type of distributed systems and available resources. Although there are several efficient techniques, but the most powerful one is using of biometric and encryption for authentication and security purposes.

Keywords Distributed system, security, authentication, encryption and biometric

1. Introduction
There are many definitions of distributed systems. but the nearest one for as that describe the Distributed system as “a collection of independent computers that operates as an integrated system from the perspective of users”. The definition has many significant phases. The first phase is that distributed systems, containing autonomous computers. The other phase is that users, think that the system is exclusively use by them. So that the independents parts must be participate. State of participation is the base of distributed systems [1].

Some people think that a computer network and distributed system are the same thing, in the fact, really it’s have different mean but related in other side. The computer network can define as a set of computers device that connected with each other. Persons who using a computer network sees that he practices different resources on different devices, its mean that the existence of multiple computers does not hidden, on the other hand, a distributed system make the user feel that he is deal with a single extra controlling homogenous computer with additional resources [2].

There are many type of distributed systems. Which can be categorized as, distributed computing systems, distributed storage systems, and distributed database systems. Distributed computing systems are usually used for high-performance requests, it’s represented as Cluster Computing Systems and Grid Computing Systems, where the computers in a cluster are mostly the same, with the same operating system for all, and also they are connected over the same network as shown in figure (1). In the other side, grid computing systems take a extraordinary degree of heterogeneity, the most interesting side in a grid computing system is that resources from diverse societies are brought together to permit the partnership of a set of people or organizations. Such a partnership is understood
in the form of a virtual organization [14]. In fig. (1) and (2) cluster and grid computing examples is shown.

**Figure 1.** Example of cluster computing [2].

**Figure 2.** Example of grid computing [2].

The key impartial of distributing storage through several devices is to keep the data in situation of disk failure through jobless storage in multiple devices and to make data accessible faster to the user in vastly distributed system, as shown in fig. (3).

**Figure 3.** Distributed storage system [2].
Finally, Distributed database system is a collection of autonomous database systems distributed through several computers that collaboratively store data in a means that a user can admission data from anyplace as if it has been stored locally regardless of where the data is really kept, see fig. (4) [2].

![Figure 4. Distributed database system [2].](image)

Distributed systems must have to accomplish the features like: Transparency, Openness, Reliability, Performance and Scalability [3]. Where transparency means hide the fact that its progressions and resources are actually distributed across several computers.

An open distributed system is defined as a system that present services depending to regular rules that designate the syntax and semantics of those services, and that what is means by Openness.

There is necessity for distributed systems to preserve availability even at low levels of hardware/software/network reliability. And at least in three different dimensions Scalability of a system can be founded. At the first dimension, a system can be scalable with respect to its size, its mean that more users and resources can simply add to the system. At the second dimension, a physically scalable system is one in which the users and resources may lie far apart. At the third dimension, a system can be administratively scalable, it’s mean that it can be easily to manage even if it have many independent managerial organizations [4]. To achieve the objectives that described above, must to deals with a system security as it is one of the essential concerns in distributed system [3]. Distributed system, design, carrying out, task and management must be taken in consideration.

Security known as “protection of information and information system from unauthorized access, modification and misuse of information”. The aim of distributed system security is the keeping the data from individuals also from software that may having wicked aims. There are four main security requirements for distributed system security, authentication, authorization, Encryption, and multi-level access control [5].

In order to achieve these requirements, there are a lot of researches, papers, articles that try to cover this side, some of these literatures that related with security side are presented in this paper.

2. Secured technique for distributed systems

For the safety of distributed systems several techniques are suggested and used, it may be focused on protection of stored information or the protection of data during transport for example using encryption or ensure the security by authentication of users in different ways like using password or identification by code or biometric, to ensure the safe access.in the next sections some of these techniques are presented and discussed their result by performance analysis.

2.1. Encryption of Stored Data

A.Dafalla, k.subair in 2019, in this work many experiments are implemented for encrypted different text file size to choose the suitable algorithm and also the time duration is calculated for each algorithm in case of encryption and decryption.as a result its notes that all algorithm are worked for
any file size but the difference appeared on duration time of encryption and decryption for different file size.

Five algorithms were using in one program using in the form that if encryption is needed, depending on text file size the suitable encryption algorithm will be selected to encrypt the specific text file. Where the file size before or after encryption will not change or is stable [6].

By taken a set of text file that have size between 10 K. Byte to 5M. Byte ,five encryption algorithms are implemented as shown in table (1). Comparing the result of algorithms results, depending on duration of time that it’s need for encryption and for decryption with different size of text file.

Table 1. Results of encryption different text file size and time Duration [6].

| File Size | AES       | DES      | TDES     | RC2       | Rijndael  |
|-----------|-----------|----------|----------|-----------|-----------|
| 10 KB     | 0.0424665 | 0.120598 | 0.046154 | 0.153397  | 0.120982  |
| 20 KB     | 0.0760034 | 0.125278 | 0.150168 | 0.252929  | 0.119187  |
| 50 KB     | 0.0850285 | 0.107128 | 0.127364 | 0.183079  | 0.126646  |
| 1 MB      | 0.1590523 | 0.154506 | 0.336127 | 0.174721  | 0.197938  |
| 2 MB      | 0.2439725 | 0.182302 | 0.324648 | 0.165653  | 0.204957  |
| 5 MB      | 0.6157387 | 0.440977 | 0.521788 | 0.355612  | 0.338767  |

The result was described in the table (1) that showed all algorithms which have been used in the work and their performance depending on time duration that can take to encrypt various sizes of text file for each of them. As a result for this work it can be said that the using of cryptographic algorithm is to ensure data protection from unapproved user to achieve the part of secrecy, reliability, and availability, many algorithm of cryptographic are advanced by several persons, some are suitable for the big size of data, and other are suitable for small size of data. But usually one algorithm can be used for both encryption and decryption for the same file.

Pedro G. M. R. Alves and Diego F. Aranha in 2018, They propose a framework for encrypting the database, which ensures confidentiality of data in an unreliable environment and search and update functions. The encryption was used to make the selection with the complexity of the time in $\theta (\log n)$ and the symmetric encryption to activate the encrypted text account. The Netflix price is explored as a case study in which an anonymous record with real user behavior information that was later identified by link attacks that contained public databases [7].

In this work, they offered a proof of concept implementation with Python via the document-based database MongoDB. To demonstrate the functionality, the key predicates required and tested by the successful Netflix Grand Prize solution were selected and tested, and the performance impact of performance is calculated in an encrypted version of the record of data.

They present that their proposal offers robustness against a compromised server. Compared to Crypt DB, their proposal enhances security because it delegates the data owner's sole responsibility for encrypting and decrypting data, as shown in Fig. (5). Thus, the confidentiality remains valid even in a scenario of compromise of database or application. The result of the work provides greater security and flexibility than the current state of knowledge. Their confidentiality even in case of compromise of the database and application servers. But in other words, it's not easy, so improve the framework is nessesary. Finally, the development of a version optimized for space and speed is very important.

Figure 5. Sequence diagram for encrypted query [7]
2.2. Biometric for Authentication

S. Jahan, et al in 2018 introducing a general view of biometrics uses in eHealth and also introduced a good and competent method for user certification by use fingerprint biometrics to ensure powerful security and privacy in eHealth information systems. The issues of security and privacy in eHealth field are principally associated with user authentication, integrity of data, confidentiality of data, and person’s privacy protection. Biometrics technology has significant chances to handle with the security issues by providing trustworthy and secure user authentication. However, because of the nature of health data environment, an important challenge deceits toward the improvement of an powerful security model that can assurance data privacy and consistency, confirming that only authorized individuals can access their corresponding health data. Additional aspects like the complexity of system and time of processing that is used for biometrics should be taken into account [8].

The main objective of this work was to use the biometric with healthcare system to ensure security and the privacy of people data that stored in distrusted way. It was not simple and its real time application. The work in espial way focused on reducing the cost of computation in biometric matching. For user authentication local minutia features are used and fast-stereo as an algorithm for matching to match the tested fingerprints features with the features that extracted from the fingerprints stored template to confirm an individuals. To estimate the performance of the suggested fingerprint biometric verification scheme, FVC2002 fingerprint data-set is used. The data set has four databases. Each database contain 110 sample of fingers with 8 impressions per finger so it have 880 fingerprints in all. The fingers are split into set A that contain 100 fingers (evaluation set) and set B that contain 10 fingers as (training set). The result is compared with other method of matching by using statistical measurement, table (2) shows the result of comparison.

Table 2. Result of comparisons for different matching method
On FVC 2002 database [8].

| Data-set | MSM (%) | FISIA (%) | Gabor (%) | Fusion (%) | Proposed (%) |
|---------|---------|----------|----------|-----------|--------------|
| DB1     | 3.5     | 1.8      | 2.3      | 5.2       | 0.42         |
| DB2     | 4.8     | 2.5      | 3.1      | 4.8       | 0.61         |
| DB3     | 6.4     | 3.2      | 5.4      | 6.4       | 1.27         |
| DB4     | 7.3     | 4.1      | 5.9      | 7.1       | 1.31         |

As described in table (2), after the implementation appear that with all using datasets, the suggested algorithm needs a smaller amount of calculation cost in fingerprint matching. Experimental results appear that the proposed method give a cooperation between the calculation effectiveness and the verification process accuracy. They be certain of that, their suggested method could be work in real-time applications.

Also K. A. Shakil etal., in 2017 proposed BAM Health Cloud which is a cloud-based system for administration of healthcare information, it ensures security of data using biometric authentication. It has been advanced after carrying out a comprehensive case study on healthcare sector in a developing country. Resilient Backpropagation neural network is used for the training of the signature samples for authentication purpose, also Hadoop Map Reduce framework is used at the same time [9].

The main goal of this work was to use biometric for a healthcare system to make it more secure, so the signature of persons who use the system was stored then the verification and authentication is done using backpropagation neural network as recognition tool then matching the result with stored data to make sure that the user is genuine or not. The huge size of data take in account, so the suggested system has two phases, one of it concerned with the controlling massive data, which is daily basis produced and the second phase conserved with the security side. ALGOHealth Security Check was offered in this work, which use Map Reduce programming model for checking security. Implement this model guarantees the scalability, flexibility and strength of the system. After execution a set of experiments it can be concluded that it accomplishes a speedup of 9x, 0.12 Equal error rate, 0.98 sensitivity and 0.95 specificity, by way of matching to other approaches existing in literature. The model is compatible for
different size application as it offers straight scaling of resources and therefore is proper for the
processing of big data.

P. Nivetha and V. Sarala, in 2014 emphasized on the data access security in cloud. Batch homomorphic encryption method is used with biometric recognition. At the same time, privacy increasingly come to be significant in the cloud atmosphere. A technique for keeping the data from attackers who already obtain, over authorized or other ways, a user’s put in storage data and private decryption keys is implemented. It bases sensitive information, such as account numbers, passwords and notes to permanently “self-destruct”, with no achievement on the user’s side. The other part of the system is emphasized on security side. Unidentified users should not contact the authorized data that was the reason for providing security. The server offers access privileges to access the data in the cloud [10].

In the proposed system user’s fingerprints samples are encrypted by using the proposed encryption schema. So and after the decryption of biometric information individual is certified to access the cloud. This is the way which is used to decide whether the user is allowed or not, after the decision of user authorization take place, the user can use the cloud services like, “show, upload or download the stored data in cloud”. System architecture with all details is illustrated in in fig. 6.

![Figure 6. System architecture [10]](image-url)

2.3 Secure Encryption Scheme with Key Exchange

D. Siva Santosh Kumar and Mrs. Kalyani in 2017 proposed model that serves the authentication and key-exchange mechanisms under a secure channel where it will be incased if one or multiple (n-1) servers are compromised even then the intruder cannot determine the password as they are hashed and the hash values are split into multiple servers making it further more difficult for the intruder to crack the password [11].

The mechanism justifies the concept of secure authentication and key exchange and as per the analysis the current model can withstand passive and active attacks and there is always scope for continuously improving and developing more advanced and strong defending mechanisms in authentication and key-exchange protocols. The main objective of this model is to ensure the system safety even if one or more serves are compromised. As it is a boring task to determine the password among the multiple servers by cracking the hash values. Even if the attacker tries with the replay attack the server can easily identify as he may not forge the nonce value which is randomly generated for every active session by the server.

2.4 Biometric Encryption in Cloud Computing

Mehreen Ansar et al in 2018, introduced paper that deal with biometrics that represent one of the most powerful authentication method, so the work is focused on who to protect the template of that biometrics which stored in cloud database using Biometric cryptosystem, the idea simply is encrypting the templates and deal with the encryption result which discloses zero percent of original information about the
template which is encrypted [12].
Several techniques to ensure high security for biometric template and also reliability of the system are implemented like fuzzy fault, fuzzy commitment schemes.
The templates are stored in different database as shown in fig. (7).

DNA is used as biometric, genetic algorithm is use to search and schedule data which are requested from the database of biometric.
Using lattice based cryptography to reduced short comings of quantum cryptography that make the templates more secure. The system became needing more resources to make computation, storing, scans and other operation on database because of increasing number of users which need to identify their templates.
Surely there are several problems that faced the work in all of its phases and there was solutions suggested to solve these problems, these problem and solution are shown in table (3).

![Figure 7. Basic Biometric process [12].](image)

| Area                          | Problem                               | Solution                                                        |
|-------------------------------|---------------------------------------|-----------------------------------------------------------------|
| Biometric Templates           | Storage                               | Cloud Data Bases                                                |
| Cloud Data Bases              | Insecure                              | Cryptography                                                    |
| Cryptography                  | Key distribution                      | Quantum Cryptography                                            |
| Quantum Cryptography          | Size, traffic, and delay of data Qu bits | Hybrid of QC algorithms and Pair Hand protocol or Lattice based Cryptography |
| Hybrid of QC algorithms and Pair Hand protocol | Heavy computation cost | Cognitive Cryptography                                           |
| Cognitive Cryptography        | Complex computing power               | Hybrid of cognitive and quantum cryptography                    |
| Hybrid of cognitive and quantum cryptography | High storage required for storing distributed templates | Hybrid of QC~CC=DNA Storages                                    |
| DNA Storages                  | Time consuming                          | Still in working how to overcome the limitations of DNA Storages |

2.5. Security and Concurrency Control
Manoj k. sah and et.al in 2014, this paper handed with disrupted system concurrency control. The techniques that can be used to get concurrency control in distributed system environments are
discussed. The concurrency control algorithms are described as basic timestamp ordering, distributed 2PL, optimistic algorithm and wound wait. Actually it is significant for database to implement the ACID properties.

With focusing on security, fragmentation and replication issues in systems with disrupted data base, there was show to the efficient properties of distributed system upon central systems [13]. The work aimed to ensure the security of the system, to achieve this aim the confidentiality, availability, and also integrity must be guaranteed.

From the study it was clear that distribution of data includes the problem of deadlock. Which is result from poor in concurrency control of the systems.so, there is need to find out the means to data distribution and accessing which leads to minimization of deadlock and thus resulting in good deployment of resources.

3. Result and Discussion

Depending on the result of studies that are presented in this paper, noted that there are several techniques are used for protection the distributed system, each of it have effect on the system also it have some problem that must be taken in account when use it.

| Technique            | Effects on system                  | problems                                      |
|----------------------|-----------------------------------|-----------------------------------------------|
| Biometric            | Very Secure access, authentication, identification and verification | Privacy, choosing the suitable type of biometric, need storage, need recognition algorithm |
| Encryption           | Good security, high privacy       | Need to choose suitable algorithm, suitable key size, key exchange |
| Biometric encryption | High protection                   | More processing and storage                   |
| Password protection  | Security, confidentiality         | Need more than one server                     |
| Secure channel       | Protect data during transposition, data integrity | Need suitable protocol, Transmission time |
| Concurrency control  | Integrity, availability           | deadlock                                       |

Table (4) show the techniques, their effect, and also the problems that faced it and need to deal with it. And from the previous studies and table we can conclude that using biometric and encryption is the most interesting technique which can be used for ensuring secure and more reliable system. The others techniques like using password and secured channel with concurrency control are used also and gave good result with some restrictions or problems. Finally it can be said that the biometric is the best and favorite one because of its unique characteristic that make it the most powerful technique in this domain.

4. Conclusion

Distrusted systems have enormous number of security problems, the researchers worked hardly especially on the last years, with the increasing of using network, to make solution for theses problem, from these studies it can be conclude that the more interesting technique which is used for authentication and to ensure high security is the biometric, different type of biometric was using. Encryption techniques were used in different phases of distributed system like encryption of stored and transform data, password, also the biometric template need to be encrypted to increase and ensure security. Also there was a need for secure channel during moving data between system components using some techniques, in addition to that the concurrency, availability and integrity of the system must take in account.

References
[1] Reza Nayebi Shahabi 2015 Security Techniques in Distributed Systems Computer Science and Information Technology 3(2): 49-53

[2] M.Firdhous 2011 implementation of security in distributed system International Journal of Computer Information Systems, vol.2, no.2.
[3] Zhidong Shen and Xiaoping Wu China 2010 The Protection for Private Keys in Distributed Computing System Enabled by Trusted Computing Platform in International Conference On Computer Design And Applications (ICCDA 2010), Qinhuangdao, Hebei pp. 576-580.

[4] Andrew S.Tanenbaum Maarten Van Steen 2007 Distributed Systems Printed in the United States of America in (Pearson Education, Inc.).

[5] Md.TabrezQuasim 2013 Security Issues in Distributed Database System Model. compusoft, an international journal of advanced computer technology, 2 (12),Volume-II, Issue-XII.

[6] AllajabuDafalla Khamis, Saad Subair 2019 Security Framework for Distributed Database System Journal of Data Analysis and Information Processing volume 7, 1-13.

[7] Pedro G. M. R. Alves, Diego F. Aranha,m 2018 A framework for searching encrypted Databases Journal of Internet Services and applications.

[8] Sharmin Jahan, Mozammel Chowdhury & Rafiqul Islam 2018 Robust user authentication model for securing electronic healthcare system using fingerprint biometrics International Journal of computers and Applications.

[9] Kashish A. Shakil, Farhana J. Zareen, mansafalam, And Suraiya Jabin 2017 BAM Health Cloud: A Biometric Authentication and Data Management System for Healthcare Data in cloud Journal of King Saud University – Computer and Information Sciences.

[10] P.Nivetha , V.Sarala, 2014 Secure Data Access Using Biometric Recognition In Cloud Computing International Journal Of Advanced Research In Computer Engineering &Technology (IARCT) Volume 3 Issue 3.

[11] D.Siva Santosh Kumar and Mrs. Kaljyani 2017 Secure Encryption scheme with key exchange for Two server Architecture, arXiv (preprint arXiv:1712.04817).

[12] Mehreen Ansar, Muhammad Sheraz Arshad Malik, Mubeen Fatima, Sadaf Aslam, Anum Rasheed, Iqra Nazir, 2018 Biometric Encryption in Cloud Computing International Journal of Computer Science and Network Security volume 18 No.8.

[13] Manoj Kumar Sah, Vinod Kumar , Ashish Tiwari,2014 security and Concurrency Control in Distributed Database System International Journal of scientific research and management (IJSRM),Volume,2,Issue,12,Pages,1839-1845.

[14] Andrew S.Tanenbaum Maarten Van Steen, 2007 Distributed Systems principles and paradigms (second edition, Pearson Prentice Hall, Pearson Education, Inc.).