ISSN 2347 - 3983

Volume 8. No. 9, September 2020

International Journal of Emerging Trends in Engineering Research

Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter256892020.pdf

https://doi.org/10.30534/ijeter/2020/256892020

Cloud data Search and verification using Order Preserving Encryption

R Shankar [1],S Janardhanarao[2], Syed Inthiyaz[3], Syed Shameem[4]

[1,2] Department of CSE, [3,4] Department of ECE,Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Guntur.

ABSTRACT

Cloud Storage system stores the huge quantity of data in their servers. In the present system, the data owner must upload the files in the cloud server and view all the data. If data owner can modify the any data, then Third party Arbitrary (TPAR) identify the modified or missing data. If there is no change in the data stored in the cloud, then TPA will produce the verification report to the data owner. Otherwise, if TPA found that any change in the data then TPAR is called for providing the report to the owner. In case if any outer people got the permission credentials of the data owner then providing the security for the data owners files is a difficult task because the possibility for the modification of the data is very high. So, reducing this modification chance is a very important issue in these days. So, for protecting the data stored in the cloud servers here, we introduced an order preserving Encryption (OPE) scheme. In this paper, the data owner cannot view all the files uploaded by him directly. By using the keyword only, the data owner can view the files related to the searched keyword. By this, if anyone enters into data owners account, he/she unable to view the files. They can view the files only after searching the particular keywords. With this proposed system, we maintain security for the owner's data.

Key words: Cloud server, searchable encryption, order preserving encryption, cloud computing.

1. INTRODUCTION

As we know that the data of the clients expanding step by step so it is vital to store this data and access this large volume of data remotely for efficient processing of this data. So, for storing this large amount of data users should connect with the cloud server with the help of the internet [1]. these servers are helpful in processing this data. Without paying too much of energy and payment these servers provides on demand services. during the usage of these servers the cloud users feel very convenient [2]. This cloud provides us enormous benefits such as low-cost, increased storage capacity, flexibity among them the most attractable benefit is its low cost [3]. These cloud models mainly have five essential characteristics they are on-demand self-service [4], broad network access, resource pooling, rapid elasticity and measured service.

In the present system the owner can store their data in the cloud. Even though the data is deposited in the cloud large number of people still worries about safety of their data. If the cloud server or third-party people get access on all the user's data they may try to analyze the documents to get the private information. In the existing system we mainly used two independent cloud servers. Among these 2 servers one is used for auditing and another one is used for data storage. High storage capacity is not required for the audit server. This audit server is used for the auditing the files which are remotely stored in the cloud storage server. Here owner can view all their data directly. So providing the efficient security for the stored data is an important issue. Performing the encryption on the spitted data we can provide the security in some manner, but direct view of the uploaded files provides more chance for the data modification and deletion. This encryption operation is done by TPA [13].

Cloud server stores this encrypted data. If the login credentials of the owner is hacked by the hackers and if they modify any data of the owner this modified data can be identified by the TPA when the data owner put the request to the TPA for verifying the data. If data is not modified, then TPA generates the report. But if the data is modified then TPAR [12] generates the report on the modified data and this report is send to data owner. This generated report consists of result about either owner changed the data or cloud server changed the data. Then owner will check the verification status of the verification request file. An efficient verification scheme is used for the verification. The verification of the requested file can be complete with the help of the signature of the encrypted data. The main disadvantage is the possibility for the alteration of the owner uploaded data is high because all the uploaded files are visible whenever the owner logged in.

So to overcome this disadvantage and to reduce the possibility for the uploaded data we are proposing this paper. In this paper the owner uploaded files can be visible to him only after searching with a keyword. If the searched word is available in the uploaded file then it will show the particular file. Here we are using the binary search algorithm for searching the keyword in the file and order preserving encryption (OPE) is used for encryption of the uploaded file. Applying the order preserving (OPE) is one of the practical ways of supporting the fast-ranked search [16].

2. RELATED WORK

A study on security problems in service distribution models of cloud computing,[7] in this paper, we tend to address the various safety issues that create a threat to the cloud. In this work may be a survey a lot of specific to the various security problems that has originated due to the nature of the service delivery models of a cloud system.

Girishma, Satyanaryana [18] showed an access control decentralized key distribution methodology providing user with anonymous authentication, revocation avoiding attacks by replay. Hierarchical attribute based Encryption (HBASE) scalable, flexible and fine-grained information scheme is properly managed in cloud computing [19]. In paper [20] the system proposed provides the cloud's integrated security to secure the data in cloud storage. In paper [21] discussed the overview and evolution of secure multi-party computation till recent advances and also stressed that SMC is very handy in collaboration in solving complex computations and keeping data from private users secure. As the increasing knowledge thieving attacks have turning into a severe threat to cloud service suppliers, [22] the projected approach helps in minimizing knowledge thieving over the illegitimate access by observance the user behavior and inundate the malicious corporate executive with decoy information.

A schema using user-based attribute encryption to encrypt cloud data, which is a public key crypto technique in which key will be based on user attributes. The attributes we have used are user biometrics which will upload the data [23]. External parties such as Third-Party Auditors (TPA) conduct an audit to verify on behalf of the user, this remote data. In paper [24] surveys various mechanisms of cryptography proposed by various researchers to check the veracity of remote data. In paper [25] shows about another necessity of ABE with outsourced unscrambling obviousness and furthermore projected a strong ABE plot with certain outsourced deciphering and showed that it is secure and obvious. A comparison of existing designs from the angle of inclusion of security infrastructure at intervals cloud system is bestowed in conjunction with a comprehensive architecture that's enclosed with each facet of security taking under consideration the foremost of the vulnerabilities. [26]. In paper [27] shows a secure privacy migration using a honey encryption cryptographic algorithm for data that is outsourced to the cloud and use migration protocol while migrating data from existing server storage to cloud server storage system that ensures data integrity and data confidentiality.[28]

Order-preserving encryption reentered: enhanced security analysis and alternative solutions, [8] this paper proposes an easy and economical transformation which will be applied to any OPE theme. In this paper we introduce standard order-preserving coding (MOPE), in which the theme is pretended with a random shift cipher. MOPE enhance the safety of OPE that means it doesn't leak any details regarding plaintext location. Thus, our analysis will resists the attacks of the data.

Security investigation for order conserving encryption schemes, [9] the progress of third-party hosting, IT outsourcing, service clouds, etc. will produce the security problems related to data. so it's a necessity to encrypt data before hosted by a third-parity, however meantime, the data should be in a position to process the queries on encrypted data .many researchers are targeted on search query process on encrypted information, in addition as the order preserving encryption (OPE) schemes. Security investigation plays an significant role in cloud computing. Currently, OPE schemes are restricted to provide security to the data. In this paper we proposed cryptographic-based OPE theme, $SE_{m,n}$. During this paper we analyze the plain text attacks that which block of data is changed.

Enabling protected and efficient graded keyword search over outsourced cloud data [10] Cloud computing economically enables the data outsourcing. To provide data security, it is necessary to encrypt the data before outsourcing to public cloud. So no one can understand the original data because the data will be in encrypted form. The traditional searchable encryption techniques allow the user to search over encrypted information through keywords, they provision only Boolean search but it does not provide the results efficiently when we perform the searching on large amount data files.

Achieving secure, fine-grained and scalabledata access control in cloud computing,[11] To protect the user sensitive data from untrusted servers, the prevailing papers apply only crypto graphical methods during which they supply cryptography to licensed users. Here for each request the owner can send decoding keys to users thus inevitably introduce a major computation overhead on the data possessor for key circulation and data organization when fine grained data access administration is desired, and therefore do not scale well. This paper projected some services for shielding the data confidentiality whereas sharing the data on cloud servers. This allowing the data owner to assign the superior part of the calculation works engaged with fine grained information get to administration to untrusted in cloud servers while not uncovering the hidden information contents. We tend to win this aim by exploiting and unambiguously combining techniques of attribute-based secret writing.[29-33]

3. METHODOLOGY

In the existing system the data owners have access to upload their information into cloud for providing the security for the uploaded data the data is encrypted [14] and then uploaded into cloud server. Here the owner can send request to the TPA for the verification. The verification of the requested file can be done by using TPA if data is modified it calls TPAR for generating the report about the who modified the data i.e., either owner or cloud after that the report will be generated and owner checks the verification status. To verify the data modified by the cloud server owner send the verification request to TPAR. Then TPAR send the generated report to the owner and owner check the status. Here we have more chance to modify the data.

To minimize this chance during this paper we introduced a searching algorithm to view the files. For searching the here we used one keyword in this searching process. If this keyword is available in the particular, then only we can view the file details. For searching with keyword, we are using the Binary search algorithm in which we can compare the each word of the file with the entered keyword. This searching can be done from top to bottom of the file. For providing the security of the owner uploaded file we are using the OPE algorithm for encryption of the data. Then similarly in the present system the owner will send the verification request to TPA and TPAR for verifying the uploaded files. Then TPA and TPAR send the generated reports to the owner. Then owner can check the verification status.

3.1 ARCHITECTURE

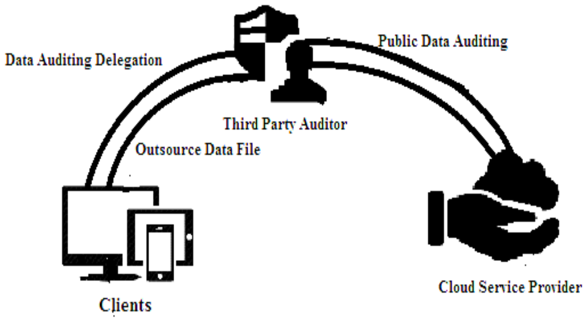


Figure 1: Cloud data storage architecture

The construction of the cloud storage system is in above figure. In this architecture we have mainly 3 modules

i. Clients

ii. Third party auditor

iii. Cloud service provider

Clients: here clients are the data owner which uses the cloud storage servers to store their data and depends on the cloud to manage their data files.

Third party auditor:here third-party auditor is also called as cloud audit server. And this TPA is used for generating the tags before storing into the cloud service provider.

Cloud service provider:this cloud service provider manages all the cloud storage servers and this csp has efficient storage space and calculation resources to maintain the client's data.

3.2 OPE (order preserving encryption)

OPE is a symmetrical so it has an another name called as order preserving symmetric encryption(OPSE).The property order preserving explains that if the plain text have the relation like $z_1 < z_2$ then the cipher texts associated with the above data E(z_1) and E(z_2) ought to additionally satisfy the $E(z_1) < E(z_2)$ .the security in our OPE theme are often provided by using an ideal object[5].we can note that any order-preserving function g from domain G={1,2,3,.....,A} to range H={1,2,....,B} are often decidedly outlined by a integration of A out of B ordered things. we can randomly pick the perfect object from all the order preserving functions and it's known as random order preserving function (ROPF). AN OPE theme is claimed to be secure if the opponent cannot differentiate the OPE from ROPF. We will additionally construct AN economical OPE theme that satisfies this secure criterion. This theme is often made by using the relation between ROPF and HGD. The HGD is used to select an order preserving function in pseudorandom manner. During this OPE theme the vary H is split into completely different buckets with different sizes. The scale of the bucket is decided by using the binary search based on a ransom HGD sampler.

The procedure [6] for this binary search algorithmic rule is described below. During this algorithmic rule Tape Gen () could be a random coin generator.

In this we also calculated the relevance score of the searching keyword also. The relevance score is defined as the number of times that the searched keyword is present in the file.

$$\text{Relevance score}=\frac{\textit{occurance of seraching keyword}}{\textit{total no of words}}$$

Keyword	w			
Field ID	F1	F2	F3	F4
Relevance Score	5.2	6.8	3.2	5.6

3.3 BINARY SEARCH:

```
Input: {L, G, H, j}
$A \leftarrow$ length (G); $B \leftarrow$ length (H)
$s \leftarrow \min(G) - 1$; $t \leftarrow \min(H) - 1$
$e \leftarrow t +$ ceil (B/2)
coin $\longleftarrow$ Tape Gen (K, (G, H, e||0))
$z \longleftarrow s +$ HGD (coin, A, B, $e - t$)
$z = s + f$
if $j \le z$ then
G $\leftarrow \{s + 1, \cdots, z\}$
H $\leftarrow \{t + 1, \cdots, e\}$
else
G $\leftarrow \{z + 1, \cdots, s + A\}$
H $\leftarrow \{e + 1, \cdots, t + B\}$
end if
Output: {G, H}
```

3.3 OPE ALGORITHM

Input: searching keyword
Output: no of matched files.
1. Enter the keyword.
2. Convert the plain text to cipher text using encryption and then generate the hash values for the given keyword.
3. For each i € total no of files

{ Perform binary search and comparison can be done by using the hash values.}

4. Get the no of matched files.

After accomplishment of the binary search the plain text j is mapped into a pail in the range H, and then the OPE algorithm allocate a secure value in the bucket similar to the encrypted rate of j. In this binary search algorithm the plaintext j_i is always mapped to fixed cipher text C_i belong to the pailnominated by the binary search algorithm.

Here we used the signature algorithm for identifying the modified content. In this algorithm we compare the signature at the server side and at the owner side. If both the signatures are equal, then we can assume that no file data is modified. If the signature is modified at any side that is at owner side or at server side, then we can assume that the information is modified and this report is send to owner. And this report consists of the result about who changed the data i.e, cloud server or the owner.

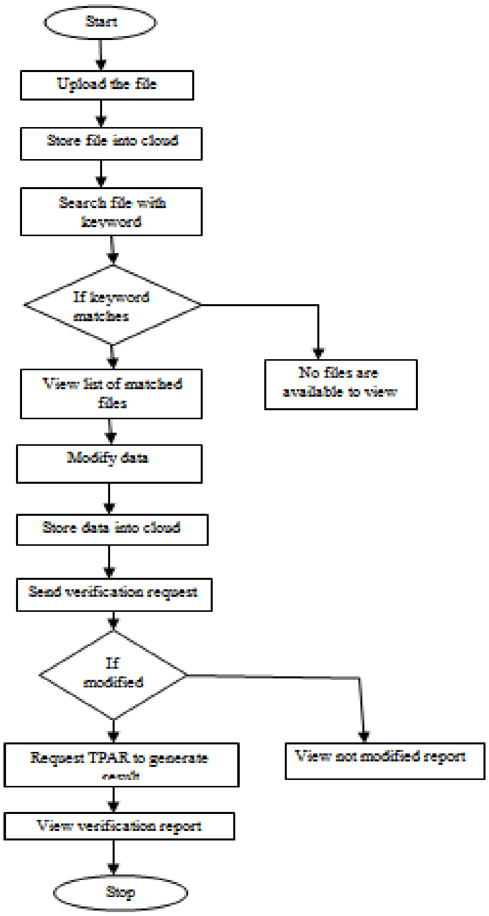


Figure 2: Flow chart of proposed method

4. RESULTS AND DISCUSSION

In this below screen we can view that the dealer enters keyword to view data:

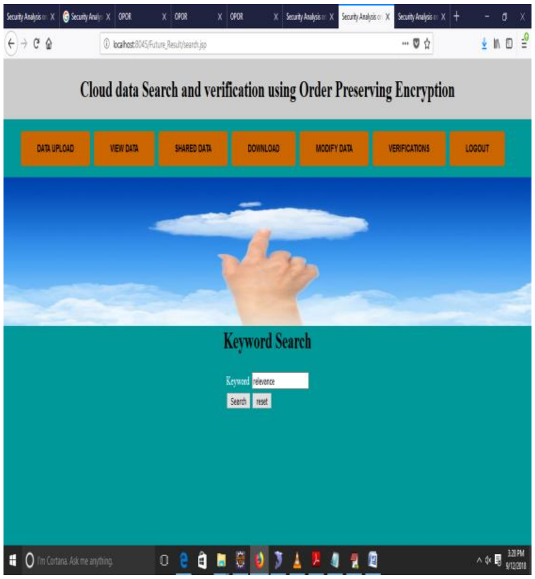
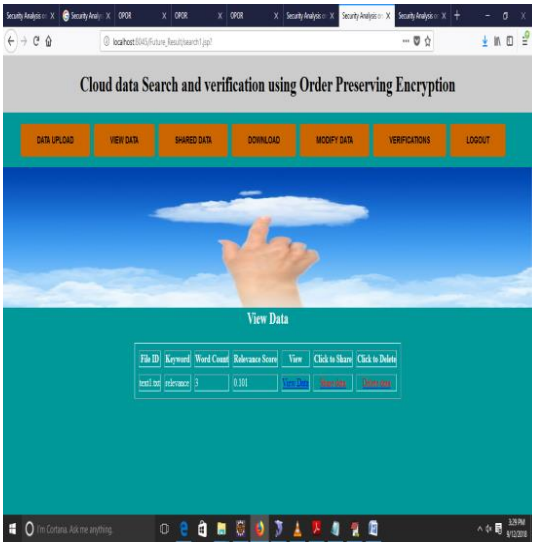


Figure 3: entry of keyword by dealer



View Data

File ID	Keyword	Word Count	Relevance Score	View	Click to Share	Click to Delete
text1.txt	relevance	3	0.101	View Data	Share data	Delete data

Figure 4: search result

View search results of the searched keyword here we can view the file id and relevance score of the searched file.

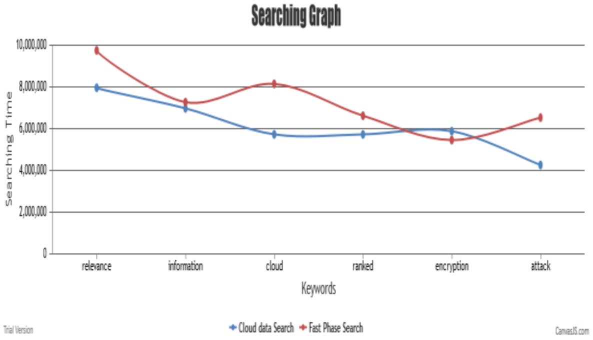


Figure 5: searching time graph

From the above graph we can represent that our proposed system will perform the fast searching than compared with the existing system[17]. In the above graph we can take the searching keyword on the X-axis and the time required to search this keyword will be taken on the Y-axis. Here the time is represented in Nano seconds (ns).so our proposed system is more effective in terms of searching.

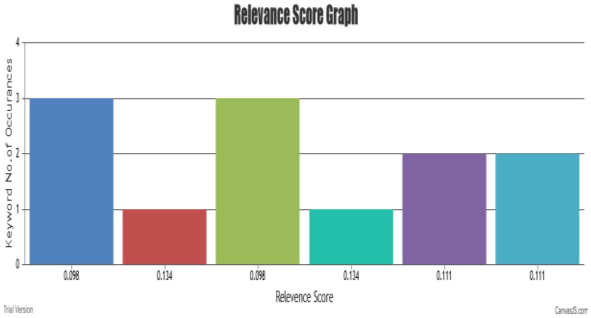


Figure 6: relevance score graph

The above graph we shown the relationship between the relevance score value and for the count of the searching keyword in the file. On X-axis we take the calculated relevance score value and on Y-axis we take the total count of the searching word.

5. CONCLUSION

In this paper, we proposed a new process in which a trusted cloud server is used for storing the data owner's files. The data is encrypted before storing in the cloud server. High security is provided for the owner's data in the cloud server. In this paper whenever the owner identity credentials are known to anybody there is a chance to modify or delete the data which is stored in the cloud server. To reduce the chance we introduced a new searching technology in which we can the files only after searching with the particular keyword. Then we will get all the files related to the search keyword. Files are available after searching only so the security for our files is improved more. For searching we implemented binary search algorithm and order-preserving encryption is used for encryption.The result shows that the accuracy and performance of our proposed system are high.

REFERENCES

1. Ke Li, Weiming Zhang, Ce Yang, and Nenghai Yu Security Analysis on One-to-Many Order Preserving Encryption Based Cloud data Search 2015 IEEE.
2. P. Mell and T. Grance, "The NIST definition of cloud computing," NIST special publication, 800(145): 7, 2011.
3. Chunye Gong, Jie Liu, Qiang Zhang, Haitao Chen and Zhenghu Gong Department of Computer Sciences National University of Defense Technology Changsha, China The Characteristics of Cloud Computing 2010 IEEE.
4. Colin Ting Si Xue1 , Felicia Tiong Wee Xin2 Benefits And Challenges Of The Adoption Of Cloud Computing In Business ijccsa.2016.
5. A. Boldyreva, N. Chenette and Y. Lee , "Order-preserving symmetric encryption," Advances in Cryptology-EUROCRYPT, 2009. Springer Berlin Heidelberg, pp. 224-241, 2009.
6. C. Wang, N. Cao and K. Ren, "Enabling secure and efficient ranked keyword search over outsourced cloud data," Parallel and Distributed Systems, IEEE Transactions 23(8), pp. 1467-1479, 2012.
7. S. Subashini and V. Kavitha, "A survey on security issues in service delivery models of cloud computing," 2011
8. A. Boldyreva, N. Chenette and A. ONeill, "Order-preserving encryp-tion revisited: improved security analysis and alternative solutions," 2011
9. L. Xiao, I.-L Yen, "Security analysis for order preserving encryption schemes,"2012
10. C. Wang, N. Cao and K. Ren, "Enabling secure and efficient ranked keyword search over outsourced cloud data" 2012
11. S. Yu, C. Wang and K. Ren, "Achieving secure, scalable, and fine-grained data access control in cloud computing", 2010.
12. A. Juels and B. S. Kaliski Jr, "Pors: Proofs of retrievability for large files," in Proc. 14th ACM Conf. Computer and Comm. Security(CCS07), 2007, pp. 584–597.
13. M. A. Shah, R. Swaminathan, and M. Baker, "Privacy-preserving audit and extraction of digital contents." IACR Cryptology ePrint Archive, Report 2008/186, 2008.
14. C.Wang, Q.Wang, K. Ren, andW. Lou, "Privacy-preserving public auditing for data storage security in cloud computing," in Proc. IEEE INFOCOM, 2010, pp. 1–9.
15. Z. Hao, S. Zhong, and N. Yu, "A privacy-preserving remote data integrity checking protocol with data dynamics and public verifiability," IEEE Trans. Knowledge and Data Eng., vol. 23, no. 9, pp. 1432–1437, 2011.
16. R. Agrawal, J. Kiernan and R. Srikant, "Order preserving encryption for numeric data," Proceedings of the 2004 ACM SIGMOD international conference on Management of data. ACM, pp. 563-574, 2004.
17. Hoi Ting Poon, Ali Miri "Fast Phrase Search for Encrypted Cloud Storage", IEEE Transactions on 2017.
18. Ranjeeth Kumar M, N.Srinivasu, Lokanatha C. Reddy, "Fine Grained Multi Access Control Via Group Sharing In Distributed Cloud Data", ournal of Theoretical and Applied Information Technology,2017.
19. Phanikumar, V. & Satyanarayana, K.V.V.. (2017). Advanced data sharing and group scheduling procedure for dynamic resource allocation of cloud computing. Journal of Advanced Research in Dynamical and Control Systems. 9. 396-409.
20. Prasad, G. S., Praneetha, D. L., Srivalli, S., and Sukesh, B. V, "Information security in cloud by using enhanced triple-DES encryption algorithm", 2019.
21. Vijaya Kumar A and Reddy L. S. S, "A critical review on application of secure multi party computation protocols in cloud environment", 2018.

22. Vamsi Krishna K and Srikanth V,“Securing cloud by mitigating insider data theft attacks with decoy technology using hadoop", 2018.
23. Ruth Ramya K, Saikrishna, D. N. V, Sravya Nandini T and Tanmai Gayatri R“A survey on using biometrics for cloud security", 2018.
24. Chaudhari S and Pathuri S. K,“A comprehensive survey on public auditing for secure cloud storage", 2018.
25. Vurukonda N, Trijan Kumar S, Rajasekhar Reddy, J. V, Adithya, A, and Boddu, S. B, “A secure attribute-based encryption scheme in cloud computing", 2018.
26. Trinath Basu M and Sastry J. K. R,“A fully security included cloud computing architecture", 2018.
27. Ravindranadh K, Kiran M. S, Durga Sai Pavan Kumar B and Priyanka D,“Data migration in cloud computing using honey encryption", 2018.
28. Myla, S., Marella, S.T., Goud, A.S., Ahammad, S.H., Kumar, G.N.S., Inthiyaz, S.,” Design decision taking system for student career selection for accurate academic system’,International Journal of Scientific and Technology Research8(9), pp. 2199-2206.
29. Karimunnisa, Syed & Kompalli, Vijaya. (2019). Cloud Computing: Review on Recent Research Progress and Issues. International Journal of Advanced Trends in Computer Science and Engineering. 8. 216-223. 10.30534/ijatcse/2019/18822019.
30. Rathod, SB; Reddy, VK, “NDynamic Framework for Secure VM Migration over Cloud Computing”Journal of Information processing systems, 2017, 10.3745/JIPS.01.0015.
31. Parvez, Muzammil. (2020). Network Security using Notable Cryptographic Algorithm for IoT Data. International Journal of Emerging Trends in Engineering Research. 8. 2169-2172. 10.30534/ijeter/2020/111852020.
32. P, Jahnavi. (2019). Facial expression detection of all emotions and face recognition system. International Journal of Emerging Trends in Engineering Research. 7. 778-783. 10.30534/ijeter/2019/087122019.
33. K, RUTH. (2019). An Efficient and Secured Biometric Authentication for IoT. International Journal of Emerging Trends in Engineering Research. 7. 604-609. 10.30534/ijeter/2019/327112019.