Scientific Journal of Impact Factor (SJIF): 4.72

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 4, Issue 10, October -2017

Security Enhancemnt Based On Hybrid Encryption Approach For Cloud Computing

Mandeep Kaur[1], Gurbahar Singh[2]

[1]Computer Science Department, BBSBEC, Fatehgarh Sahib, Maharaja Ranjit Singh Punjab Technical University
[2]Computer Science Department, BBSBEC, Fatehgarh Sahib, Maharaja Ranjit Singh Punjab Technical University

Abstract— *With the advancements in the technology, the amount of data over the internet is getting increased and it becomes tedious task to handle such a huge amount of data. It is noticed that the clients need to hire the data mining tasks from cloud service providers this is known as outsourcing. Mostly the selection of cloud service provider for various services relies upon the choice of clients because the maintenance of confidentiality and privacy is major concern nowadays. This study specifically conducted to develop to a novel approach (DSEkNN i.e. Data Split Encrypted kNN) for managing the security by using outsourced kNN (k-Nearest Neighbor) by distributing dataset randomly to different CSPs on the basis of their limits for handling data. For the purpose of security two-level or multi level encryption is applied by using RSA (Rivest-Shammir Adleman) and DNA. In DSEkNN firstly the data is encrypted with RSA and the encrypted data is further crypted with DNA hence it provides two level security to the datasets. For proving the efficiency of the DSEkNN is compared with OCkNN and PPkNN. After observing the results it is obtained that the proposed work is capable to decrease the burden of data owner's by reducing the computation time and communication cost. The simulation is done on two datasets i.e. red wine and white wine.*

Keywords—*Cloud Computing, Cloud Service Providers, Outsourced Services, RSA, DNA, kNN, Data Security.*

I. INTRODUCTION

Cloud computing is also referred as Distributed computing. It provides new dimensions to various service providers as well end users. To incur these new dimensions, new techniques for designs are implemented that convey infrastructure as service (IaaS), Stage as a Service (PaaS) and also Software as a Service (SaaS) [1]. To get the virtualization of servers, routers, memory unit and various other components, IaaS service is used. After this the PaaS service is used to generate, implement and test the internet based application on the cloud service. Hence, this service is responsible to maintain the operation of software for their whole lifespan. Third service is SaaS, it is responsible for providing the software application when the end user demands for it [2]. Following feature are exhibited by the Cloud process:

- **Utility-based pricing:** As the service is offered on the basis of pay per use therefore the implementation cost has been reduced. The cost of services varies as different service providers offers services with different cost. Example: if the service providers offers one service to the end user on hourly basis and on the other hand he can provide the same service on the basis of number of clients.
- **Broad network access:** To access the Cloud computing services, web connection is mandatory. The devices that are connected with web service can easily harness cloud services. Example of devices as follow: cellular phones, laptops, and tablet etc. The cloud service providers have stored their data at various locations.
- **Shared resource pooling:** To provide different resources to end users, the virtualization concept and multi tenancy technique has been implemented.
- **Dynamic resource provisioning:** the end user can demand for any resource and their demands are fulfilled by without any manual participation.
- **Measured service:** the measurement of the services and resources available to the end users is done on the basis of demand of customer and also the consumption is measured on the basis of "pay-per-use".
- **Self-organizing:** As per the requirement of Cloud service provider they have to coordinate their resources. On the basis of demand of consumer, the resources are coordinated.

Data security is major concern in cloud computing. Because the full fledge data is handled by the cloud server and each and every user has an access to this database which is handled over the server. Hence in this way the confidentiality of the data can be lost. The various authors provide different security mechanism to provide the security to the data over the cloud server but lacks at some points. For example in traditional work [1] the security was provided to the data over the cloud by using the concept of multiple keys i.e. more than one encryption techniques was applied to the dataset. But it can be concluded that the usage of multiple key for securing the data can enhances the complexity and computation time for the cloud service providers which indirectly results to the increased communication cost. Therefore this study provides a novel approach for securing the data by using multilevel key process and KNN as classifier for pattern matching. The objective of this work is to reduce the communication cost, computation time and data owner computations.

II. PROBLEM FORMULATION

As the Volume and variety of data captured by organizations or companies are growing more rapidly than ever, resource constrained clients tend to outsource both data and data mining tasks to cloud service providers (CSP) to improve efficiency and save costs. Nowadays, there's a growing trend for collaborative data mining, that is, clients who have a common goal to make use of shared information are likely to cooperate in computation over their databases vertically or horizontally partitioned and distributed among multiple clouds. Unfortunately, due to the continuous occurrences of privacy breach in cloud computing, concerns about security have significantly impeded the wide adoption of outsourced data mining. To protect confidentiality of data from unauthorized access, sensitive data are usually encrypted. But the major issue that was analyzed in the previous approached based on client based encryption and KNN based cloud privacy preserving is that the approach is best suitable for the small data but as the data set is becoming large it is hard to encrypt it using a same key for security point of view secondly if the data is saved on single cloud it can be easily used by un authorized party if key get decrypted. So there is need to update the data encryption strategy on clients end and uploading it to clouds.

III. PROPOSED WORK

As per the literature study and after analyzing the previous paper methodology the main issues those were discussed in the problem formulation can be resolved by using the approach of splitting data for the large dataset and secondly the encryption approach will be enhance by using the hybridization of algorithms named as RSA-DNA approach. The hybrid model will be used for the encryption process on different sets of data after splitting it on threshold basis.
One other update that is done is the data uploading on the cloud , in proposed work not the whole data will be uploaded to a single cloud server it will be uploaded to multiple servers so cannot be accessed easily and that will be done by choosing the priority of the data and KNN bases cloud server selection.

1. Initialize Data

The top most steps is to initialize the dataset. In this work we initialize dataset in two different categories of wine quality i.e. red wine and white wine**.**

2. Data Pre-Processing

Data preprocessing is done for formatting the data.

3. Dataset Allotment

After formatting the dataset, next step is to divide the dataset randomly to the clouds. In this case total 5 clouds are considered for data allotment. The allotment is done randomly but on the basis of the upper limit of data handling of the clouds. In any case if any of clouds gets allotted with the higher amount of data as compare to its handling capacity then the extra data will be fetched from that cloud and gets assigned to those clouds where there data allotment is low.

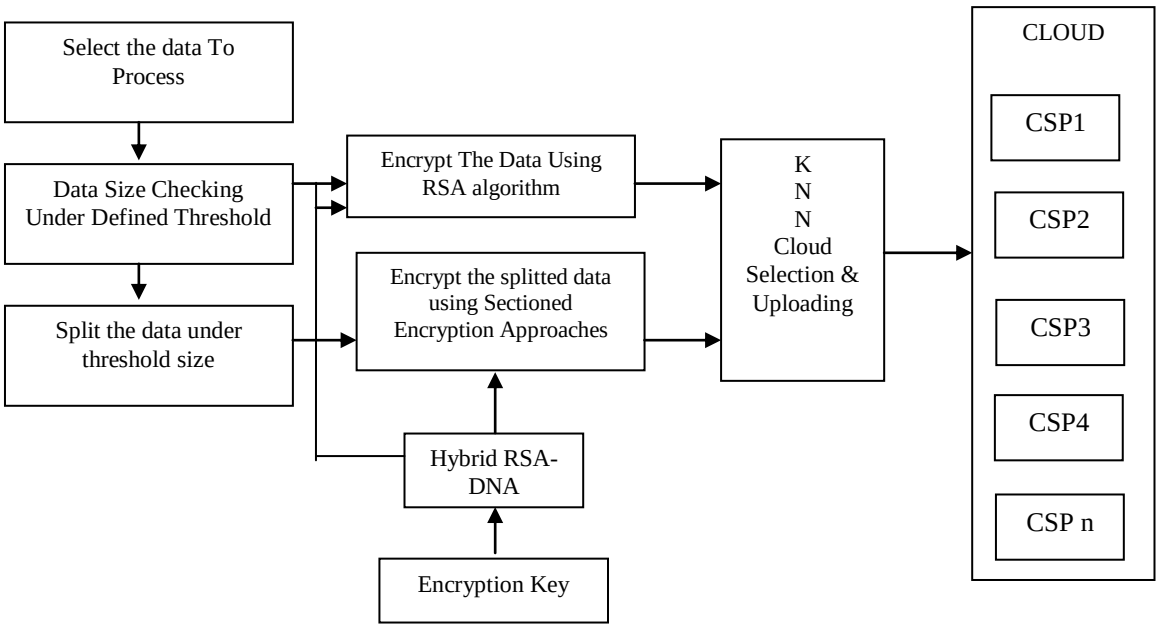

Figure 1. Block Diagram of proposed work

4. **RSA-DNA encryption**

After allotting the data next step is to apply Hybrid RSA-DNA encryption to the dataset in order to secure the data. The proposed work provides multi key security dataset by firstly applying RSA and then RSA encrypted data is further encrypted by applying DNA encryption.

5. **KNN Query Handling**

This protocol needs every service provider to handle the user's queries by using KNN by maintaining its confidentiality while assuring the user about the decryption of encrypted data such as using clouds service provider. In this step the KNN is applied for handling the incoming queries. Only the input queries will be forwarded to the KNN for pattern matching instead of whole data. In this the KNN query handling computation is evaluated.

IV. IMPLEMENTATION

This section portrays the results that are obtained after implementing the hybrid security mechanism with external KNN to the dataset of the wine. Two datasets are considered on the basis of wine quality under this work for the purpose of simulation i.e. Red wine and White wine. The datasets are acquired from UCI Machine Learning Repository and it comprised of 1599 instances, 12 attributes, 11 class labels and 4898 instances, 12 attributes, 11 class labels respectively.
The purposed work initiates by fetching the dataset and then a pre-processing is applied on the dataset in order to format the data. After formatting the dataset the data is allotted to each and every cloud randomly. In this study we consider total 5 clouds and each cloud has some upper limit for data handling. This is done to enhance capacity of data handling of the clouds and to reduce the communication cost of the network. All the processing is done on the service provider level.
The results are evaluated in the following term:

1. Communication Cost: Communication cost is evaluated in terms of Mb. This parameter describes the amount of data incurred for completing a communication.
2. Computation Time: Computation Time is measured in minutes. This parameter is used to depict the time that is taken by the cloud to compare the cipher text with external kNN.
3. Data Owner Computation: It is measured in seconds. It refers to the time that is taken by the owner for a small portion of the all evaluations.

The graph in figure 2 represents the contrast among DSEkNN, OCkNN and PPkNN on the basis of cloud computation time. In the graph the x axis represents the number of counts for KNN which ranges between 5 and 25. The y axis calibrates the data in minutes to depict the computation time which starts from 0 minutes and ends at 3000 minutes. The following evaluations are observed from the graph of figure 2.

Table 1: Computational time

S. No.	Techniques	Values(min.)
1.	DSEkNN	153
2.	OCkNN	1010
3.	PPkNN	2900

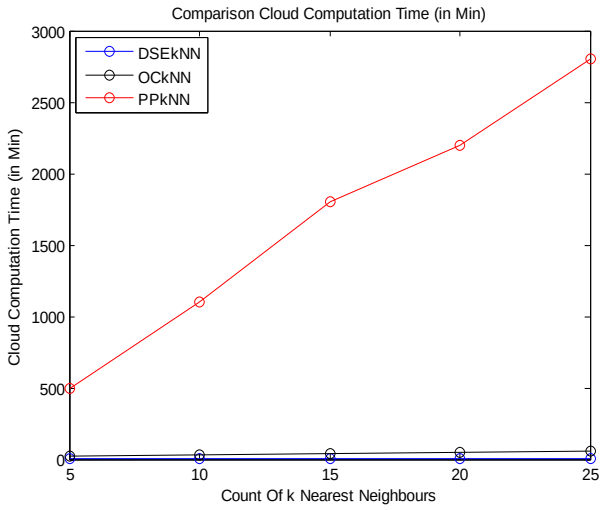

Figure 2 Computational Times of DSEkNN, OCkNN and PPkNN.

The values in the above table 1 are derived on the basis of the graph that is shown in figure 2. From the values that are given in above table, it can be said that the DSEkNN (Data Split Encrypted kNN) has the lowest computational time as compare to OCkNN and PPkNN. This proves that the proposed work has less complexity which decreases the computational time and increases its efficiency.
The figure 3 comprised of graph that illustrates the contrast among proposed and traditional work to show the efficacy of the proposed work in the terms of communication cost. In the graph the y axis calibrates the data for communication cost in Mbs and x axis depicts the number of counts for external kNN.
The table 2 is derived on the basis of the graph of figure 3. It shows a contrast between DSEkNN, OCkNN and PPkNN on the basis of communication time. The values in the table proves that the DSEkNN has the least communication time (153 Mb) in comparison to the OCkNN and PPkNN. Hence it proves the proficiency of the proposed work over traditional techniques.

Table 2. Communication Time

S. No.	Techniques	Values(Mb.)
1.	DSEkNN	153
2.	OCkNN	1010
3.	PPkNN	2900

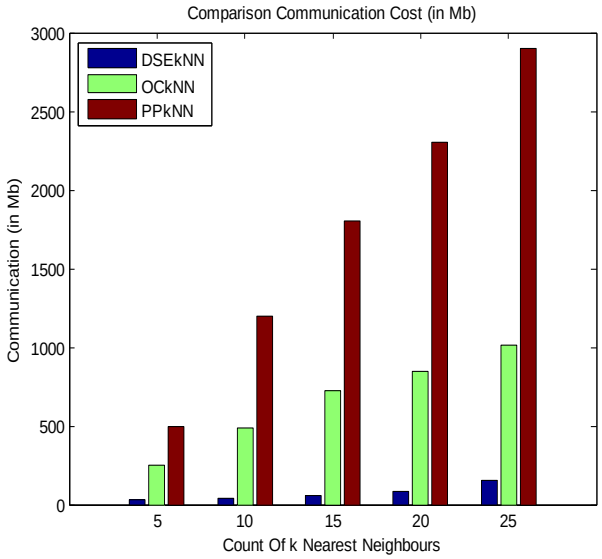

Figure 3 Computational Times of DSEkNN, OCkNN and PPkNN

The computation cost of DSEkNN is 153 Mb, OCkNN is 1010 Mb and PPkNN is 2900 Mb (Figure 3). The graph depicts that the computation cost increases with the increment in count.

Table 3 Data Owner's Cost

S. No.	Techniques	Values(sec.)
1.	DSEkNN	12
2.	OCkNN	38.5
3.	PPkNN	70.5

The table 3 represents the values that are observed from the graph of figure 4. It represents that the DSEkNN produces almost half of the workload of the client in comparison to the OCkNN and PPkNN. The data gathered in table 3 also proves that the burden of data owner is decreased with the implementation of outsourced kNN and hybrid RSA-DNA encryption. Figure 4 explains the comparison of proposed work and traditional work on the basis of data owner's computation.

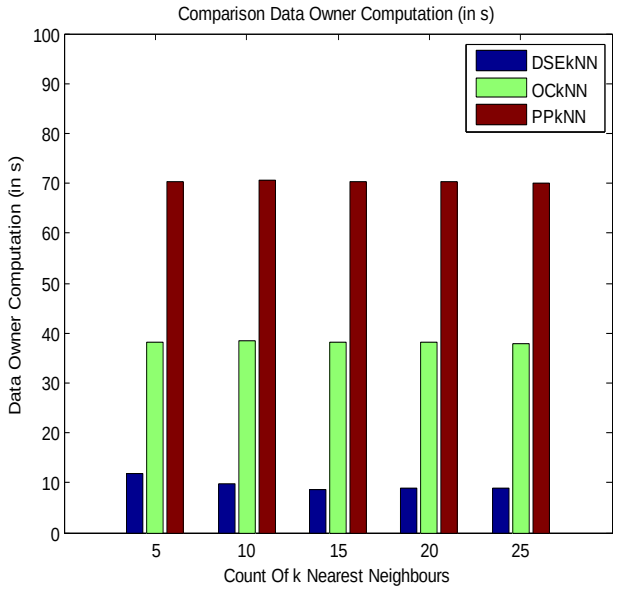

Figure 4 Data Owner's Computation of DSEkNN, OCkNN and PPkNN

V. CONCLUSION

This study provides a hybrid mechanism for securing the data over cloud services by using RSA-DNA encryption mechanism along with kNN classifiers. The proposed work is implemented by using 2 datasets regarding the quality of wine i.e. Red Wine and White Wine. Both have different attributes. The Dataset is pre-processed in order to perform pattern matching on it by using kNN. The dataset is randomly distributed at the 5 clouds. In this each and every cloud has its own data handling limit which prevents the clouds from getting overloaded with data. After allotting the data to the clouds the proposed work also performs a check to evaluate that whether the limit of a cloud did not exceeds. If a cloud is located with the higher amount of allotted data as compare to its handling limit then the extra data is fetched from 0that cloud and further divided to the rest of clouds randomly where the data allotment is low so that the data load can be balanced over the clouds. Then multiple key level securities is applied to the dataset by using RSA and DNA. The proposed approach for security enhances the security level and also reduces the complexity of the multiple key handling process. Then the encrypted data is uploaded as file. Then kNN is applied for performing pattern matching among single query and dataset on the cloud server.

It can be concluded that the proposed work provides a multi level security to the dataset without increasing the complexity for key handling mechanism. It also reduces the data communication which directly reduces the data communication rates. Query handling is performed only by forwarding the query for pattern matching instead of whole data which reduces the computation time and data owner's computation time also.

REFERENCES

[1] Ashish Singh et al, "Cloud Security issues and challenges : A Survey", ELSEVIER, Vol 79, Issue 1, Pp 88-115, 2017
[2] Luigi Coppolino et al, "Cloud Security: Emerging threats and current solution", ELSEVIER, Vol 59, PP 126-140, 2017
[3] Sushil Kr Saroj et al, "Threshold Cryptography Based Data Security in Cloud Computing", IEEE, 2015
[4] Gururaj Ramchandran et al, "A Comprehensive Survey on Security in Cloud Computing", ELSEVIER, Vol 110, Pp 465-472, 2017
[5] Salman Iqbala et al, "On cloud security attacks: A taxonomy and intrusion detection and prevention as a service", ELSEVIER, Vol 74, Pp 98-120, 2016
[6] Hong Rong et al, "Efficient Integrity Verification of Secure Outsourced kNN Computation in Cloud Environments". IEEE, Pp 236-243, 2016
[7] Sunoh Choi et al, "Secure kNN Query Processing in Un-trusted Cloud Environments", IEEE, Vol 26, Issue 11, Pp 2818-2831, 2014
[8] Ni Zhang et al, "A Research on Cloud Computing Security", IEEE, PP 370-373, 2013,
[9] Rabi Prasad Padhi et al, "Cloud Computing: Security Issues and Research Challenges", IRACST, Vol 1, Issue 2, Pp 136-146, 2011
[10] Suruchi V. Nandgaonkar et al, "A comprehensive study on cloud computing", IJCSMC, Vol 3(4), Pp 733-739, 2014
[11] Monjur Ahmed et al, "Cloud Computing And Security Issues in The Cloud", IJNSA, Vol 6, Issue 1, Pp 25-36, 2014

[12] Yunchuan Sun et al, "Data Security and Privacy in cloud computing", HINDAWI, Vol 2014, Pp 1-9, 2014
[13] Keiko Hashizume et al, "An Analysis if security issues for cloud computing", SPRINGER, Vol 4, Issue 5, 2013
[14] Michael armbrust et al, "A view on Cloud computing", communication of the ACM, Vol 53(4), 2009
[15] Ling Qian et al, "Cloud computing-An overview", Springer, Pp 626-631, 2009
[16] S. Poonkodi et al, "Providing a Secure Data Forwarding in Cloud Storage System Using Threshold Proxy Re-Encryption Scheme", IJETAE, Vol 3, Issue 1, Pp 468-472, 2013,
[17] Pradnyesh Bhisikar et al, "Security in Data Storage and transmission in cloud computing", IJARCSSE, Vol 3, Issue 3, Pp 410-415, 2013,
[18] Shankar Nayak Bhukya, et al, "Data Security in Cloud Computing and Outsourced Databases", IEEE, Pp: 2458-2462, 2016.
[19] S. V. Nandgaonkar et al, "A comprehensive study on cloud computing", IJCSMC, Vol 3(4), Pp 733-739, 2014
[20] Mrinal Kanti Sarkar, et al, "A Framework to Ensure Data Storage Security In Cloud Computing", IEEE, Pp: 1-4, 2016.
[21] Ahmed Albugmi et al, "Data Security in Cloud Computing", IEEE, Pp: 55-59, 2016.
[22] K. B. Priya Lyer, et al, "Analysis of Data Security in Cloud Computing", IEEE, Pp: 540-543, 2016.
[23] C. Linda Hepsiba, et al, "Security Issues in Service Models of Cloud Computing", IJCSMC, pp: 610-615, 2016.
[24] Mrinal Kanti Sarkar, et al, "A Framework to Ensure Data Storage Security In Cloud Computing", IEEE, Pp: 1-4, 2016.
[25] R. Velumadhava Rao, et al, "Data Security Challenges and its Solutions in Cloud Computing", ELSEVIER, Pp: 204-209, 2015.
[26] Kamal Kumar Chauhan, et al, "Homomorphic Encryption for Data Security in cloud Computing", IEEE, Pp: 206-209, 2015.
[27] Ashok Kote, et al, "Cloud Data Security Challenges and its Solutions", IJCCER, 2015.
[28] Pin Zhang, et al, "Access Control Research on Data Security in Cloud Computing", IEEE, Pp: 873-877, 2015.
[29] Sushil Kr Saroj, et al, "Threshold Cryptography Based Data Security in Cloud Computing", IEEE, Pp: 202-207, 2015.
[30] Aws Naser Jaber, et al, "A study in Data Security in Cloud computing", IEEE, Pp: 367-371, 2014.
[31] M. Sugumaran, et al, "An Architecture for Data Security in Cloud Computing", IEEE, Pp: 252-255, 2014.
[32] Manas M N, et al, "Cloud Computing Security Issues and Methods to Overcome", IJARCCE, Pp: 6306-6310, 2014.
[33] Minhaj Ahmed Khan et al, "A survey of security issues for cloud computing", ELSEVIER, Vol 71, PP 11-29, 2016
[34] T V Sathyanarayana, et al, "Data Security in Cloud Computing", IEEE, Pp; 822-827, 2013.
[35] Huda Elmogazy, et al, "Towards healthcare data security in Cloud Computing", IEEE, Pp: 363-368, 2013.
[36] Ms. Disha H. Parekh, et al, "An Analysis of Security Challenges in Cloud Computing", IJACSA, Pp: 38-46, 2013.
[37] Du Meng, et al, "Data Security in Cloud Computing", IEEE, Pp: 810-813, 2013.
[38] Prashant Rewagad, et al, "Use of Digital Signature with Diffie Hellman Key Exchange and AES Encryption Algorithm to Enhance Data Security in Cloud Computing", IEEE, Pp: 437-439, 2013.
[39] Abhinay B. Angadi, et al, "Security Issues with Possible Solutions in Cloud Computing-A Survey", IJARCET, Pp: 652-661, 2013.
[40] Nidal M. Turab, et al, "Cloud Computing Challenges and Solutions", IJCNC, Pp: 209- 216, 2013.

