# PERFORMANCE IMPROVEMENT OF CLOUD DATA USING KEYWORD SEARCH

**[1]Khushboo Patel [2]Shriya Chole [3]Abhay Kumar [4]Rohini Padole [5]Anup Bhange**
[1234] U. G. Student, Dept. of Computer Technology. KDKCE Nagpur, Maharashtra, India
5Assistant Professor Dept OF Computer Tech KDKDCE, Nagpur

***Abstract* — Data Mining has wide applications in many areas such as banking, medicine, scientific research and among government agencies. Classification is one of the commonly used tasks in data mining applications. For the past decade, due to the rise of various privacy issues, many theoretical and practical solutions to the classification problem have been proposed under different security models. However, with the recent popularity of cloud computing, users now have the opportunity to outsource their data, in encrypted form, as well as the data mining tasks to the cloud. Since the data on the cloud is in encrypted form, existing privacy preserving classification techniques are not applicable. In this paper, we focus on solving the classification problem over encrypted data. In particular, we propose a secure hybrid k-NN classifier over encrypted data in the cloud. The propose hybrid k-NN protocol protects the confidentiality of the data, user's input query, and data access patterns. To the best of our knowledge, our work is the first to develop a secure k-NN classifier over encrypted data under the semi-honest model. Also, we empirically analyze the efficiency of our solution through various experiments***

***Keywords*—k-NNclassifier,cloud,encryption ,mining,privacy preservation.**

## 1. INTRODUCTION

Today’s digital infrastructure supports innovative ways of storing, processing, and disseminating data. In fact, we can store our data in remote servers, access reliable and efficient services provided by third parties, and use computing power available at multiple locations across the network. Furthermore, the growing adoption of portable devices (e.g., PDAs, mobile phones) together with the diffusion of wireless connections in home and work environments have led to a more distributed computing scenario. These advantages come at a price of higher privacy risks and vulnerabilities as a huge amount of (private) information is being circulated and stored, often not under the direct control of its owner. Be that as it may, when information are encoded, independent of the fundamental encryption plan, performing any information mining assignments turns out to be extremely difficult without ever unscrambling the information. There are other sample security concerns, shown by the accompanying Data mining over encrypted data (denoted by DMED) [3] on a cloud also needs to protect a client's record when the record is a part of a data mining process. However cloud can also abstract useful and sensitive information about the outsource data items by observing the data access patterns even if the data are encrypted. Therefore, the privacy/security requirements of the DMED problem on a cloud are of three types: (1) privacy of the encrypted data, (2) privacy of a user’s query record, and (3) hiding data access patterns.

## 1.2 OUR CONTRIBUTIONS

In this paper, we propose a novel PPkNN protocol, a securek-NN classifier over semantically secure encrypted data. Inour protocol, once the encrypted data are outsourced to thecloud, Alice does not participate in any computations.Therefore no information is revealed to Alice. In additionour protocol meets the following privacy requirements:Contentsof D or any intermediate results should notbe revealed to the cloud.Bob’s query q should not be revealed to the cloud.$c_q$should be revealed only to Bob. Also, no otherinformation should be revealed to Bob.Data access patterns, such as the records corresponding to the k-nearest neighbors of q, should not berevealed to Bob and the cloud (to prevent any inference attacks).We emphasize that the intermediate results seen by the cloud in our protocol are either newly generated randomized encryptions or random numbers. Thus, which data records correspond to the k-nearest neighbors and the output class label are not know to the cloud In addition, after sending his encrypted query record to the cloud, Bob does not involve in any computations. Hence, data access patterns are further protected from Bob (see Section 5 for more details).The rest of the paper is organized as follows. We discuss the existing related work and some concepts as a back-ground in Section 2. A set of privacy-preserving protocols and their possible implementations are provided in Section. The formal security proofs for the mentioned privacy-pre-serving primitives are provided in Section The proposed PPkNN protocol is explained in detail in Section 5. Section 6 discusses the performance of the proposed protocol under different parameter settings. We conclude the paper along with future work in Section

## 2. RELATED WORK AND BACKGROUND

Due to space limitations, here we briefly review the existing related work and provide some definitions as a background.Please refer to our technical report [5] for a more elaborated related work and background.At first, it seems fully homomorphic cryptosystems (e.g., [6]) can solve the DMED problem since it allows a third-party (that hosts the encrypted data) to execute arbitrary functions over encrypted data without ever decrypting them. However, we stress that such techniques are very expensive and their usage in practical applications have yet to be explored. For example, it was shown in [7] that even for weak security parameters one “bootstrapping” operation of the homo morphic operation would take at least 30 seconds on a high performance machine.It is possible to use the existing secret sharing techniques in SMC, such as Shamir’s scheme [8], to develop a PPkNN protocol.l However, our work is different from the secret sharing based solution in the following aspect. Solutions based on the secret sharing schemes require at least three parties whereas our work require only two parties. For example, the constructions based on Sharemind [9], a well-known SMC framework which is based on the secret sharing scheme, assumes that the number of participating par-ties is three. Thus, our work is orthogonal to Sharemind and other secret sharing based schemes.

## 2.1 QUERY PROCESSING OVER ENCRYPTED DATA

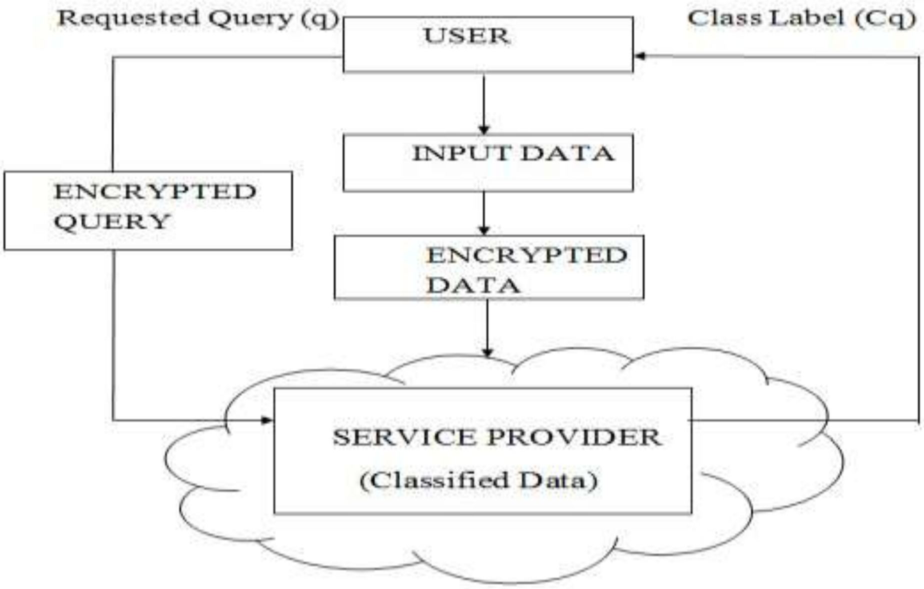
Various techniques related to query processing over encrypted data have been proposed, e.g., [19], [20], [21]. However, we observe that PPkNN is a more complex problem than the execution of simple kNN queries over encrypted data [22], [23]. For one, the intermediate k-nearest neighbors in the classification process, should not be disclosed to the cloud or any users. We emphasize that the recent method in [23] reveals the k-nearest neighbors to the user. Second, even if we know the k-nearest neighbors, it is still very difficult to find the majority class label among these neighbors since they are encrypted at the first place to prevent the cloud from learning sensitive information. Third, the existing work did not addressed the access pat-tern issue which is a crucial privacy requirement from the user's perspective.In our most recent work [24], we proposed a novel secure k-nearest neighbor query protocol over encrypted data that protects data confidentiality, user's query privacy, and hides data access patterns. However, as mentioned above, PPkNN is a more complex problem and it cannot be solved directly using the existing secure k-nearest neighbor techniques over encrypted data. Therefore, in this paper, we extend our previous work in [24] and provide a new solution to the PPkNN classifier problem over encrypted data.

## 2.2 THREAT MODEL

We adopt the security definitions in the literature of secure multi-party computation [25], [26], and there are three common adversarial models under SMC: semi-honest, covert and malicious. In this paper, to develop secure and efficient protocols, we assume that parties are semi-honest. Briefly, the following definition captures the properties of a secure protocol under the semi-honest model [27], [28].Definition 1. Let $a_i$ be the input of P's executionof image of the protocol p and $b_i$ be the output for party $P_i$ computed from p. Then, p is secure if P can be simulated from $a_i$ and $b_i$ such that distribution of the simulated image is computationally indistinguishable from $P_i$In the above definition, an execution image generally includes the input, the output and the messages communicated during an execution of a protocol. To prove a protocol is secure under semi-honest model, we generally need to show that the execution image of a protocol does not leak any information regarding the private inputs of participating parties [28].

## 3. PRIVACY PRESERVING DATA MINING (PPDM)

Privacy Preserving Data Mining (PPDM) is defined as the process of extracting/deriving the knowledge about data without compromising the privacy of data [3, 41, 48]. In the past decade, many privacy-preserving classification techniques have been proposed in the literature in order to protect user privacy. Agrawal and Srikant [3], Lindell and Pinkas [40] introduced the notion of privacy-preserving under data mining applications. In particular to privacy preserving classification, the goal is to build a classifier in order to predict the class label of input data record based on the distributed training dataset without compromising the privacy of data. 1. Data Perturbation Methods: In these methods, values of individual data records are perturbed by adding random noise in a such way that the distribution of perturbed data look very different from that of actual data. After such a transformation, the perturbed data is sent to the miner to perform the desired data mining tasks. Agrawal and Srikant [3] proposed the first data perturbation technique to build a decision-tree classifier. Since then many other randomizationbased methods have been proposed in the literature such as [5]. However, as mentioned earlier in Section 1, data perturbation techniques cannot be applicable for semantically secure encrypted data. Also, they do not produce accurate data mining results due to the addition of statistical noises to the data. 2. Data Distribution Methods: These methods assume the dataset is partitioned either horizontally or vertically and distributed across different parties. The parties later can collaborate to securely mine the combined data and learn 4 the global data mining results. During this process, data owned by individual parties is not revealed to other parties. This approach was first introduced by Lindell and Pinkas [14] who proposed a decision tree classifier under two-party setting. Since then much work has been published using secure multiparty computation techniques [1, 15].



1 Data Perturbation Methods: In these methods, values of individual data records are perturbed by adding random noise in a such way that the distribution of perturbed data look very different from that of actual data. After such a transformation, the perturbed data is sent to the miner to perform the desired data mining tasks. Agrawal and Srikant [3] proposed the first data perturbation technique to build a decision-tree classifier. Since then many other randomizationbased methods have been proposed in the literature such as [5]. However, as mentioned earlier in Section 1, data perturbation techniques cannot be applicable for semantically secure encrypted data. Also, they do not produce accurate data mining results due to the addition of statistical noises to the data. 2. Data Distribution Methods: These methods assume the dataset is partitioned either horizontally or vertically and distributed across different parties. The parties later can collaborate to securely mine the combined data and learn 4 the global data mining results. During this process,

data owned by individual parties is not revealed to other parties. This approach was first introduced by Lindell and Pinkas [14] who proposed a decision tree classifier under two-party setting. Since then much work has been published using secure multiparty computation techniques [1, 15]. Classification is one important task in many applications of data mining such as health-care and business. Recently, performing data mining in the cloud attracted significant attention. In cloud computing, data owner outsources his/her data to the cloud. However, from user's perspective, privacy becomes an important issue when sensitive data needs to be outsourced to the cloud. The direct way to guard the outsourced data is to apply encryption on the data before outsourcing. Unfortunately, since the hosted data on the cloud is in encrypted form in our problem domain, the existing privacy preserving classification techniques are not sufficient and applicable to PPkNN due to the following reasons. (i) In existing methods, the data are partitioned among at least two parties, whereas in our case encrypted data are hosted on the cloud. (ii) Since some amount of information is loss due to the addition of statistical noises in order to hide the sensitive attributes, the existing methods are not accurate. (iii) Leakage of data access patterns: the cloud can easily derive useful and sensitive information about users' data items by simply observing the database access patterns. For the same reasons, in this paper, we do not consider secure k-nearest neighbor techniques in which the data are distributed between two parties (e.g., [12]). 2.2 Query processing over encrypted data Using encryption as a way to achieve the data confidentiality may cause another issue at the cloud during the query evaluation. The question here is "how can the cloud perform computations over encrypted data while the data stored are in encrypted form?" Along this direction, various techniques related to query processing over encrypted data have been proposed, e.g., [12]. However, we observe that PPkNN is a more complex problem than the execution of simple KNN queries over encrypted data [13]. For one, the intermediate k-nearest neighbors in the classification process , should not be disclosed to the cloud or any users. We emphasize that the recent method in [14] reveals the k-nearest neighbors to the user. Secondly, even if we know the k-nearest neighbors, it is still very difficult to find the majority class label among these neighbors since they are encrypted at the first place to prevent the cloud from learning sensitive information. Third, the existing work did not addressed the access pattern issue which is a crucial privacy requirement from the user's perspective. In our most recent work [12], we proposed a novel secure k-nearest neighbor query protocol over encrypted data that protects data confidentiality, user's query privacy, and hides data access patterns. However, as mentioned above, PPkNN is a more complex problem and it cannot be solved directly using the existing secure k-nearest neighbor techniques over encrypted data. Therefore, in this paper, we extend our previous work in and provide a new solution to the PPkNN classifier problem over encrypted data. More specifically, this paper is different from our preliminary work in the following four aspects. First, in this paper, we introduced new security primitives, namely secure minimum (SMIN), secure minimum out of n numbers (SMINn), secure frequency (SF), and proposed new solutions for them. Second, the work in did not provide any formal security analysis of the underlying sub-protocols. On the other hand, this paper provides formal security proofs of the underlying sub-protocols as well as the PPKNN protocol under the semi-honest model. Additionally, we demonstrate various techniques through which the proposed protocol can possibly be extended to a protocol that is secure under the malicious model. Third, our preliminary work in addresses only secure KNN query which is similar to Stage 1 of PPKNN. However, Stage 2 in PPKNN is entirely new. Finally, our empirical analyses in Section VI are based on a real dataset whereas the results in are based on a simulated dataset. In addition, new results are included in this paper. As mentioned earlier, one can implement the proposed protocols under secret sharing schemes. By doing so, we need to have at least three independent parties. In this work, we only concentrate on the two party situation ; thus, we adopted the Paillier cryptosystem. Two-party and multi-party (three or more parties) SMC protocols are complement to each other, and their applications mainly depend on the number of available participants. In practice, two mutually independent clouds are easier to find and are cheaper to operate. On the other hand, utilizing three cloud servers and secret sharing schemes to implement the proposed protocols may result more efficient running time. We believe both two-party and multi-party schemes are important. As a future work, we will consider secret sharing based PPKNN 5 implementatio

## 4. SCURITY UNDER THE MALICIOUS MODEL

The next step is to extend our PPkNN protocol into a secure protocol under the malicious model.Under the malicious model, an adversary (i.e., either $C_1$ or $C_2$) can arbitrarily deviate from the protocol to gain someAdvantages (e.g., learning additional information about inputs) over the other party. The deviations include, as example,for $C_1$ (acting as a malicious adversary) to instantiate the PPkNN protocol withmodified inputs (say E $_{\text{and}}$p) and to abort the protocol after gaining partial information.However, in PPkNN, it is worth pointing out that neither $C_1$ nor $C_2$knows the results of Stages 1 and 2. In addition, allintermediate results are either random or pseudo-random values. Thus, even when an adversary modifies the intermediate computations he/she cannot gain any additional information. Nevertheless, as mentioned above, the adversary can change the intermediate data or perform computations incorrectly before sending them to the honest party which may eventually result in the wrong output. Therefore, we need to ensure that all the computations per-formed and messages sent by each party are correct.Remember that the main goal of SMC is to ensure the honest parties to get the correct result and to protect their private input data from themalicious parties. Therefore, under the two-party SMC scenario, if both parties are malicious, there is no point to develop or adopt an SMC protocol at the first place. In the literature of SMC [30], it is the norm that at most one party can be malicious under the two-party scenario. When only one of the party is malicious, the standard way of preventing the malicious party from misbehaving is to let the honest party validate the other party's work using zero-knowledge proofs [31]. However, checking the validity of operations at each step of PPkNN can significantly increase the cost .An alternative approach, as proposed in [32], is to instantiate two independent executions of the PPkNN protocol by swapping the roles of the two parties in each execution. At the end of the individual executions, each party receives the output in encrypted form. This is followed by an equality test on their outputs. More specifically, suppose E1 and $E_2$ be the outputs received by $C_1$ and $C_2$ respectively, where $pk_1$ and $pk_2$ are their respective public keys. Note that the outputs in our case are in encrypted format and the corresponding cipher texts (resulted from the two executions) are under two different public key domains. Therefore, we stress that the equality test based on the additive homomorphic encryption properties which was used inis not applicable to our problem. Nevertheless, $C_1$ and $C_2$can perform the equality test based on the traditionalgarbled-circuit technique [33].

### 4.1 COMPLEXITY ANALYSIS

The total computation complexity of Stage 1 is boundedby O l $\log_2$ n encryptions and exponentiations. On the other hand, the total computation complexity of Stage 2 is bounded by $O\log_{2\ W}$ encryptions and exponentiations. Due to space limitations, we refer the reader to [5] for detailed complexity analysis of PPkNN. In general, as w n, the computation cost of Stage1 should be significantly higher than that of Stage 2. This observation is further justified by our empirical results given in the next section.

## 5. DATA SET AND EXPERIMENTAL RESULT

For our experiments, we used the Car Evaluation dataset from the UCI KDD archive [34]. It consists of 1,728 records (i.e., n¼1;728) and six attributes (i.e., m¼6). Also, there is a separate class attribute and the dataset is categorized into four different classes (i.e., w¼4). We encrypted this dataset attribute-wise, using the Paillier encryption whose key size is varied in our experiments, and the encrypted data were stored on our machine. Based on our PPkNN protocol, we then executed a random query over this encrypted data. For the rest of this section, we do not discuss about the performance of Alice since it is a one-time cost. Instead, we evaluate and analyze the performances of the two stages in PPkNN separately.

### 5.1 EMPIRICAL RESULTS

In this section, we discuss some experiments demonstrating the performance of our PPkNN protocol under different parameter settings. We used the Paillier cryptosystem [4] as the underlying additive homomorphic encryption scheme and implemented the proposed PPkNN protocol in C. Various experiments were conducted on a Linux machine with an Intel Xeon Six-Core CPU 3.07 GHz processor and 12 GB RAM running Ubuntu 12.04 LTS. To the best of our knowledge, our work is the first effort to develop a secure k-NN classifier under the semi-honest model. There is no existing work to compare with our approach. Hence, we evaluate the performance of our PPkNN protocol under different parameter settings.

## 6. LITERATURE SURVEY

[1]S. De Capitani di Vimercati, S. Foresti, and P. Samarati, "Managing and accessing data in the cloud: Privacy risks and approaches," in CRiSIS, pp. 1 –9, 2012

In this Data mining is the extraction of interesting patterns or knowledge from huge amount of data. In recent years, with the explosive development in Internet, data storage and data processing technologies, privacy preservation has been one of the greater concerns in data mining. A number of methods and techniques have been developed for privacy preserving data mining. This paper provides a wide survey of different privacy preserving data mining algorithms and analyses the representative techniques for privacy preserving data mining, and points out their merits and demerits. Finally the present problems and directions for future research are discussed.

[2]P. Williams, R. Sion, and B. Carbunar, "Building castles out of mud: practical access pattern privacy and correctness on untrusted storage," in ACM CCS , pp. 139–148, 2008.

This paper presents We present the reticent statistical zero-knowledge protocols to prove statements such as: A committed number is a prime . • A committed (or revealed) number is the product of two safe primes, i.e., primes p and q such that (p − 1) =2and (q − 1) =2areprime.

• A given integer has large multiplicative order modulo a composite number that consists of two safe prime factors.

[3]P. Paillier, "Public key cryptosystems based on composite degree residuosity classes," in Eurocrypt , pp. 223–238, 1999

For the past decade, query processing on relational data has been studied extensively, and many theoretical and practical solutions to query processing have been proposed under various scenarios. With the recent popularity of cloud computing, users now have the opportunity to outsource their data as well as the data management tasks to the cloud. However, due to the rise of various privacy issues, sensitive data (e.g., medical records) need to be encrypted before outsourcing to the cloud. In addition, query processing tasks should be handled by the cloud; otherwise, there would be no point to outsource the data at the first place. To process queries over encrypted data without the cloud ever decrypting the data is a very challenging task. In this paper, we focus on solving the k-nearest neighbor (KNN) query problem over encrypted database outsourced to a cloud: a user issues an encrypted query record to the cloud, and the cloud returns the k closest records to the user. We first present a basic scheme and demonstrate that such a naive solution is not secure. To provide better security, we propose a secure kNNprotocol that protects the confidentiality of the data, user's input query, and data access patterns. Also, we empirically analyze the efficiency of our protocols through various experiments. These results indicate that our secure protocol is very efficient on the user end, and this lightweight scheme allows a user to use any mobile device to perform the KNN query

[4]B. K. Samanthula, Y. Elmehdwi, and W. Jiang, "k-nearest neighbor classification over semantically secure encrypted re-lational data." eprint arXiv:1403.5001, 2014.

Ensuring proper privacy and protection of the information stored, communicated, processed, and disseminated in the cloud as well as of the users accessing such information is one of the grand challenges of our modern society. As a matter of fact, the advancements in the Information Technology and the diffusion of novel paradigms such as data outsourcing and cloud computing, while allowing users and companies to easily access high quality applications and services, introduce novel privacy risks of improper information disclosure and dissemination. In this paper, we will characterize different aspects of the privacy problem in emerging scenarios. We will illustrate risks, solutions, and open problems related to ensuring privacy of users accessing services or resources in the cloud, sensitive information stored at external parties, and accesses to such information.

[5]C. Gentry and S. Halevi, "Implementing gentry's fully- homomorphic encryption scheme," in EUROCRYPT , pp. 129– 148, Springer, 2011.

Most of the cryptographic work in privacy-preserving distributed datamining deals with semi-honest adversaries, which are assumed to follow the prescribed protocol but try to infer private information using the messages they receive during the protocol. Although the semi-honest model is reasonable in some cases, it is unrealistic to assume that adversaries will always follow the protocols exactly. In particular, malicious adversaries could deviate arbitrarily from their prescribed protocols. Secure protocols that are developed against malicious adversaries require utilization of complex techniques. Clearly, protocols that can withstand malicious adversaries provide more security. However, there is an obvious trade-off: protocols that are secure against malicious adversaries are generally more expensive than those secure against semi-honest adversaries only. In this paper, our goal is to make an analysis of trade-offs between performance and security in privacy preserving distributed data mining algorithms in the two models. In order to make a realistic comparison, we enhance commonly used sub protocols that are secure in the semi-honest model with zero knowledge proofs to be secure in the malicious model. We compare the performance of these protocols in both models.

| Name of author | Algorithm used | Disadvantages |
|---|---|---|
| 1.S. De Capitani di Vimercati, S. Foresti, and P. Samarati, "Managing and accessing data in the cloud: Privacy risks and approaches," in | the reticent statistical zero-knowledge protocols | Privacy preservation not considered. |

| | | |
|---|---|---|
| CRiSIS | | |
| 2. P.Paillier, “Public key cryptosystems based on composite degree residuosity classes,” in Eurocrypt | Homomorphic encryption | Classification not consider |
| 3.B. K. Samanthula, Y. Elmehdwi, and W. Jiang, “k-nearest neighbor classification over semantically secure encrypted re-lational data | “k-nearest neighbor classification | Searching operation is time consuming. |
| 4.C. Gentry and S. Halevi, “Implementing gentry’s fully-homomorphic encryption scheme,” in EUROCRYPT , pp. 129– 148, Springer, 2011. | Fully homomorphic | Classification not consider. |

## 7. CONCLUSION AND FUTURE WORK

From the above literature survey it is clearly observed that Different type of privacy preserving classification has been introduced in past few years. This method is not applicable to outsourced databases. Here there is need to improve performance of data classification and searching over cloud encrypted data proposed system makes the system highly scalable and minimizes information leakage. Prevents overloads by ranking the files at the user side, reducing bandwidth and protects document frequencywe try to implement system is secure, scalable and accurate compared to the other ranked keyword search.

To protect user privacy, various privacy-preserving classification techniques have been proposed over the past decade. The existing techniques are not applicable to outsourced database environments where the data resides in encrypted form on a third-party server. This paper proposed a novel privacy-preserving k-NN classification protocol over encrypted data in the cloud. Our protocol protects the confidentiality of the data, user’s input query, and hides the data access patterns. We also evaluated the performance of our protocol under different parameter settings.Since improving the efficiency of $SMIN_n$ is an important first step for improving the performance of our PPkNN protocol, we plan to investigate alternative and more efficient solutions to the $SMIN_n$ problem in our future work. Also, we will investigate and extend our research to other classification algorithms.

## 8. REFERENCES

[1] D.Nurmi, R.Wolski, C.Grzegorczyk, G.Obertelli,S.Soman, L.Youseff and D.Zagorodnov, “The eucalyptus open-source cloud- computing system,” CCGRID 20009.9th IEEE/ACM International Symposium, 2009.

[2] S.S and A. Basu, “Performance of eucalyptus and open stack clouds on future grid,”International Journal of Computer Applications, vol. 80,no.13,pp.31-37, 2013.

[3] Z.Pantić and M. A.Babar, “Guidelines for Building a Private Cloud Infrastructure,“ IT University of Copenhagen, Denmark, Copenhagen, Denmark,2012.

[4] B. Beal, “Public vs. private cloud applications: twocriticaldifferences,”23May2012.[Online]. Available:http://searchcloudapplications.techtarget.com/feature/Public-vs-private-cloud-applicationsTwo-critical-differences.

[5] Tarik Moataz, Abdullatif Shikfa, “Boolean symmetric searchable encryption,” ASIA CCS '13 Proc. of the 8th ACM SIGSAC symposium on Information computer and communications security, .pp. 265- 276, NY, USA , 2013.

[6] R. Curtmola, J.A. Garay, S. Kamara, and R. Ostrovsky, “SearchableSymmetric Encryption: Improved Definitions and Efficient Constructions,”Proc. ACM 13th Conf. Computer and Comm. Security (CCS), 2006

[7] D. Song, D. Wagner, and A. Perrig, “Practical Techniques for Searches on Encrypted Data,” Proc. IEEE Symp. Security and Privacy, 2000.

[8] D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, “Public key encryption with keyword search,” in Proc. of EUROCRYP’04, volume 3027 of LNCS. Springer, 2004.

[9] S.Adhikari,G.Bunce,W.Chan,A.Chandramouly,D.Kamhout, B.McGeough,J.JonSlusser,C.Spence and B. Sunderland, “Best practices for building and enterprise private cloud,” Intel IT Centre,2011.

[10] B.Adler,”Designing Private and hybrid clouds: architectural best practices,” RightScaleInc.,2012.

[11] “Planning Guide: Virtualisation and cloud computing,” Intel IT Centre,2013

[12] Y.Wadia,“TheEucalyptusOpenSourcePrivateCloud. over Untrusted Data Cloud through Privacy Homomorphism,”Proc. IEEE 27th Int’l Conf. Data Eng. (ICDE), 2011.

[13] G.VonLaszewski,J.Diaz, F.WangandG.Fox, “Comparison of multiple cloud frameworks,” IEEE on Cloud computing(CLOUD),vol.734,no.741,pp.24-29,2012,5th International Conference

[14] F. Bao, R. Deng, X. Ding, and Y. Yang, “Private query on encrypted data in multi-user settings,” in Proc. of ISPEC 2008.

[15] A. Swaminathan, Y. Mao, G.-M. Su, H. Gou, A.L. Varna, S. He, M. 5u, and D.W. Oard, “Confidentiality-Preserving Rank-Ordered Search,” Proc. Workshop Storage Security and Survivability, 2007.

[16] Cong Wang, Ning Cao, Jin Li, Kui Ren, Wenjing Lou, “Secure ranked keyword search over encrypted cloud data,” IEEE 2010 30th International Conference 2010.

[17] S. Zerr, D. Olmedilla, W. Nejdl, and W. Siberski, “Zerber+r: Top-k Retrieval from a Confidential Index,” Proc. 12th Int’l Conf.Extending Database Technology: Advances in Database Technology (EDBT), 2009.

[18] P. Golle, J. Staddon, and B. Waters, “Secure Conjunctive Keyword Search over Encrypted Data,” over Untrusted Data Cloud through Privacy Homomorphism,”Proc. IEEE 27th Int’l Conf. Data Eng. (ICDE), 2011.

[19] L. Ballard, S. Kamara, and F. Monrose, “Achieving Efficient Conjunctive Keyword Searches over Encrypted Data.