Summary of Practical Techniques for Searches on Encrypted Data

Summary -

The concept of storing data in the cloud has been of prime interest in the software industry. Once the data is in the cloud, it can be analyzed/searched with high computing power machines provided by the cloud. Many has considered it as the "Next Big Thing". An important question which arises is - ' Shall we trust the cloud server ? '. In the past, there have been incidents where the users data has been compromised but what if the data saved in the cloud is encrypted ?. Well then even if the data is accessible to an adversary, it is meaningless to him/her. Techniques like basic scan method and indexed search method are discussed here. This concept introduces another problem which is - How to perform searches on this encrypted data. Various techniques resolving this issue have been discussed in this paper. These algorithm uses the symmetric key encryption which is fast. One assumption made in this paper is that all the words are of equal length which is very rare in real life. The solutions given to overcome this assumption is not explained in details and lacks clarity.

Opinion -

I my opinion, the searching techniques for encrypted data ensures that the search parameters and the results are secure from the adversaries but it comes at a cost. Indexing the documents with some pre defined keys will for sure increase the efficiency but only for some suitable cases. When the results or encrypted documents are returned to the client, they are decrypted and for this reason there is some client side computation required. The paper also mentions some kind of padding of words which will increase the size of documents. For large data set this will create a problem. Why ? The reason is that cost of saving data in cloud is directly proportional to the size of data. The algorithms also talks about minimizing the communication overhead by limiting the data being shared between the client and server which plays an important role in cloud computing. Overall this concept introduced some interesting facts about data encryption and searching the same.

Pros and Cons -

The advantages of using these algorithms are many. Firstly, adversary or the server cannot gain any knowledge from the search key as it is encrypted. Secondly, adversary or the server cannot determine any information about the documents or document's content or even cannot distinguish between documents or learn anything about it. Lastly, these algorithm are efficient with small documents as the complexity is O(n) for sequential scan algorithm which can be improved significantly by indexing the document. Well there some limitations of this algorithm. For example if there is a large data set of the documents like a million documents then the sequential scan method won't work. Similarly the indexing is based on a set of pre defined values Which provides quick results only when the search keyword involves the pre defined values but not otherwise. Also the index has to be updated regularly in order to meet the new requirements of search criteria. Updating of the keys is not secure.

Thanks