AN ANALYSIS ON RELIABLE TECHNIQUES FOR KEYWORD BASED SEARCH AND DATA SHARING IN CLOUD COMPUTING

ABSTRACT

Today data sharing and maintaining its security is major challenge. User in the data sharing system upload their file with the encryption using private key. This property is especially important to any large scale data sharing system, as any user leak the key information then it will become difficult for the data owner to maintain security of the information. In this paper provide a concrete and efficient instantiation of scheme, prove its security and provide an implementation to show its practicality. There are lots of challenges for data owner to share their data on servers or cloud. There are different solutions to solve these problems. These techniques are very much critical to handle key shared by the data owner. This paper will introduce the trusted authority to authenticate user those who have the access to the data on cloud. SHA algorithm is used by the trusted authority to generate the key and that key will get share to user as well as the owner. The trusted authority module receives encrypted file using AES Algorithm from the data owner and computes hash value using MD-5 algorithm. It stores key in its database which will be used during the dynamic operations and to determine the cheating party in the system. Trusted authority send file to CSP module to store on cloud.The resulting key sets are shown to have a number of desirable properties that ensure the confidentiality of communication sessions against collusion attacks by other network nodes.

INTRODUCTION

One possible solution is to migrate character sequences to public cloud computing platforms and to request that Cloud Service Providersprocess sequence comparisons. At present, primary sequence comparison algorithms are deployed as a universal outsourcing service on public clouds. But at the same time, its security and privacy issues are increasingly emerging. The outsourced data stored as plaintext could easily be exposed to malicious external intruders and internal attackers in the CSP, and the individual private information carried by character sequences (e.g., personal identification, financial transaction records,genetic markers for some diseases, information that is used to identify paternity or maternity, etc.) could more or less be disclosed or abused. Therefore, secure outsourcing is designed to protect the privacy of character sequences, and toensurethatthescheduledcomputingrequests arenormally performed on the cloud servers.

For this purpose, we present a scheme called Encrypted Sequence Comparison basedonasingle-server model.Somenovelsaltedhashandencryptiontechniquesare employed to allow public clouds to perform sequence comparisons directly on the character sequences outsourced as ciphertext. Overall, E-SC achieves a user-controlled reliable storage and a collusion-resistant outsourcing service, which plays an important role in the trade-off between security and execution performance. Our scheme is easy in deployment, efficient in processing and controllable in overhead. Thecontributionsofthispaperaremainlyinthefollowing four aspects.

1) Based on the universal model of a public cloud outsourcing, we propose an overall architecture for E-SC. This architecture is built on the end user

and the unmodified CSP. Its overall system model, whichhas been demonstrated to be secure under the threat model, is user-friendly and implementation-friendly.

2) A salted hash algorithm is improved to hash the character sequences and the indexes of cost matrices, so as to defend against statistical attacks. An additive order preserving encryption algorithm is designed to encrypt the elements of cost matrices. Also, this algorithm can achieve an indistinguishability under additive ordered chosen-plaintext attack with linear time complexity. 3) Asinglecloudserverworksforthefirsttimetoprovide a privacy-preserving computable outsourcing service to effectively resist collusion attacks from the cloud. Withpre-processingmodulesofpadding,partition,and expansion, there is no need to decrypt any outsourced data in the non-interactive sequence comparison stage. 4) Simulation results show that the overall execution performance of our E-SC is negatively correlated with its security.

LITERATURE SURVEY

Literature survey is the most important step in software development process. Before developing the tool it is necessary to determine the time factor, economy and company strength. Once these things are satisfied, then the next step is to determine which operating system and language can be used for developing the tool. Once the programmers start building the tool the programmers need lot of external support. This support can be obtained from senior programmers, from book or from websites. Before building the system the above consideration are taken into account for developing the proposed system.

The major part of the project development sector considers and fully survey all the required needs for developing the project. For every project Literature survey is the most important sector in software development process. Before developing the tools and the associated designing it is necessary to determine and survey the time factor, resource requirement, man power, economy, and company strength. Once these things are satisfied and fully surveyed, then the next step is to determine about the software specifications in the respective system such as what type of operating system the project would require, and what are all the necessary software are needed to proceed with the next step such as developing the tools, and the associated operations.

TITLE: Efficient and verifiable outsourcing scheme of sequence comparisons

AUTHOR:Y.Feng,H.Ma,andX.Chen

With the rapid development of cloud computing, the techniques for securely outsourcing prohibitively expensive computations are getting widespread attentions in the scientific community. In the outsourcing computation paradigm, the clients with resource-constrained abilities can outsource the heavy computation workloads into the cloud server and enjoy unlimited computing resources in a pay-per-use manner. One of the most critical functionalities in outsourcing computation is the verifiability of the result. That is, the client should efficiently verify the validity of the result returned by the cloud servers. In this paper, we solve the problem of verifiable outsourcing computation of sequence comparisons by integrating the technique of Yao's garbled circuit with homomorphic encryption. Compared with the existing schemes, our proposed solution enables clients to efficiently detect the misbehavior of dishonest servers. Furthermore, our construction re-garbles the circuit only for malformed responses and thus is very efficient for real-world applications. Besides, we also present the formal analysis for our proposed construction.

TITLE:Secure outsourcing of sequence comparisons

AUTHOR: M. J. Atallah and J. Li

With the advent of cloud computing, secure outsourcing techniques of sequence comparisons are becoming increasingly valuable, especially for clients with limited resources. One of the most critical functionalities in data outsourcing is verifiability. However, there is very few secure outsourcing scheme for sequence comparisons that the clients can verify whether the servers honestly execute a protocol or not. In this paper, we tackle the problem by integrating the technique of garbled circuit with homomorphic encryption. As compared to existing schemes, our proposed solution enables clients to efficiently detect the dishonesty of servers. In particular, our construction re-garbles the circuit only for malformed responses and hence is very efficient. Besides, we also present the formal analysis for our proposed construction.

TITLE: Secure and private sequence comparisons

AUTHOR: M. J. Atallah, F. Kerschbaum, and W. Du

We give an efficient protocol for sequence comparisons of the edit-distance kind, such that neither party reveals anything about their private sequence to the other party (other than what can be inferred from the edit distance between their two sequences – which is unavoidable because computing that distance is the purpose of the protocol). The amount of communication done by our protocol is proportional to the time complexity of the best-known algorithm for performing the sequence comparison. The problem of determining the similarity between two sequences arises in a large number of applications, in particular in bioinformatics. In these application areas, the edit distance is one of the most widely used notions of sequence similarity: It is the least-cost set of insertions, deletions, and substitutions required to transform one string into the other. The generalizations of edit distance that are solved by the same kind of dynamic programming recurrence relation as the one for edit distance, cover an even wider domain of applications.

TITLE:Toward a practical data privacy scheme for a distributed implementation of the Smith-Waterman genome sequence comparison algorithm

AUTHOR: D. Szajda, M. Pohl, J. Owen, and B. Lawson

Volunteer distributed computations utilize spare processor cycles of personal computers that are connected to the Internet. The resulting platforms provide computational power previously available only through the use of expensive clusters or supercomputers. However, distributed computations running in untrustworthy environments raise a number of security concerns, including computation integrity and data privacy. This paper introduces a strategy for enhancing data privacy in some distributed volunteer computations, providing an important first step toward a general data privacy solution for these computations. The strategy is used to provide enhanced data privacy for the Smith-Waterman local nucleotide sequence comparison algorithm. Our modified Smith-Waterman algorithm provides reasonable performance, identifying most, and in many cases all, sequence pairs that exhibit statistically significant similarity according to the unmodified algorithm, with reasonable levels of false positives. Moreover the modified algorithm achieves a net decrease in execution time, with no increase in memory requirements. Most importantly, our scheme represents an important first step toward providing data privacy for a practical and important real-world algorithm.

TITLE:New algorithms for secure outsourcing of modular exponentiations

AUTHOR: X. Chen, J. Li, J. Ma, Q. Tang, and W. Lou

Modular exponentiations have been considered the most expensive operation in discrete-logarithm based cryptographic protocols. In this paper, we propose a new secure outsourcing algorithm for exponentiation modular a prime in the one-malicious model. Compared with the state-of-the-art algorithm [33], the proposed algorithm is superior in both efficiency and checkability. We then utilize this algorithm as a subroutine to achieve outsource-secure Cramer-Shoup encryptions and Schnorr signatures. Besides, we propose the first outsource-secure and efficient algorithm for simultaneous modular exponentiations. Moreover, we prove that both the algorithms can achieve the desired security notions.

EXISTING SYSTEM

Large-scale problems in the physical and life sciences are being revolutionized by Internet computing technologies, like grid computing, that make possible the massive cooperative sharing of computational power, bandwidth, storage, and data. A weak computational device, once connected to such a grid, is no longer limited by its slow speed, small amounts of local storage, and limited bandwidth: It can avail itself of the abundance of these resources that is available elsewhere on the network. An impediment to the use of “computational outsourcing” is that the data in question is often sensitive, e.g., of national security importance, or proprietary and containing commercial secrets, or to be kept private for legal requirements such as the HIPAA legislation, Gramm-Leach-Bliley, or similar laws. This motivates the design of techniques for computational outsourcing in a privacy-preserving manner, i.e., without revealing to the remote agents whose computational power is being used, either one’s data or the outcome of the computation on the data.

DISADVANTAGES OF EXIXTING SYSTEM

Secure outsourcing for widely applicable sequence comparison problems
Risk of Leak of Secret Information

PROPOSED SYSTEM

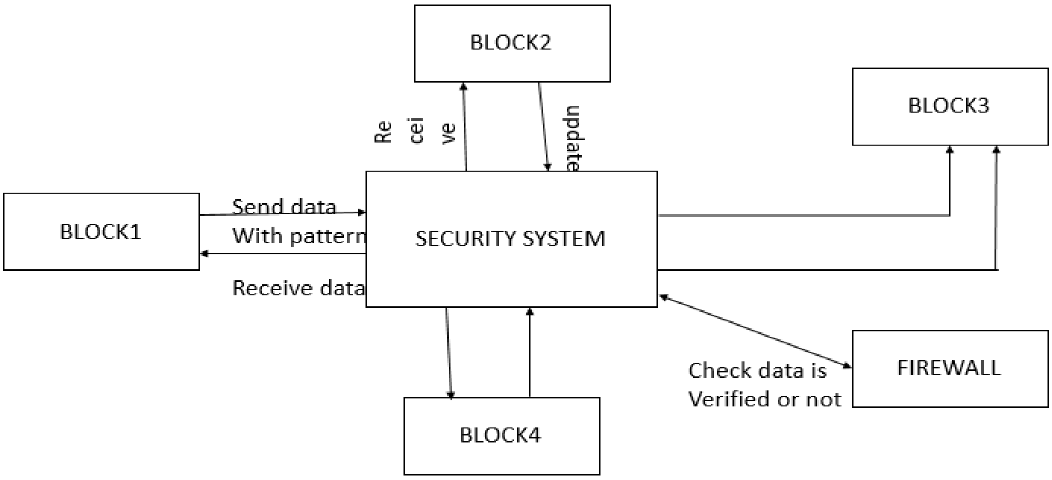
We propose a secure data sharing scheme, which can achieve secure key distribution and data sharing for dynamic group. We provide a secure way for key distribution without any secure communication channels. The users can securely obtain their private keys from group manager without any Certificate

Authorities due to the verification for the public key of the user. Our scheme can achieve fine-grained access control, with the help of the group user list, any user in the group can use the source in the cloud and revoked users cannot access the cloud again after they are revoked. We propose a secure data sharing scheme which can be protected from collusion attack. The revoked users can not be able to get the original data files once they are revoked even if they conspire with the untrusted cloud. Our scheme can achieve secure user revocation with the help of polynomial function. Our scheme is able to support dynamic groups efficiently, when a new user joins in the group or a user is revoked from the group, the private keys of the other users do not need to be recomputed and updated. We provide security analysis to prove the security of our scheme.

ADVANTAGES OF PROPOSED SYSTEM

- Power Means of Persuasion and control
- More Reliable
- It’s more secure and efficient.
- Data confidentiality

SYSTEM ARCHITECTURE



SYSTEM REQUIREMENTS

HARDWARE REQUIREMENTS:

- System - Pentium-IV
- Speed - 2.4GHZ
- Hard disk - 40GB
- Monitor - 15VGA color
- RAM - 512MB

SOFTWARE REQUIREMENTS:

- Operating System - Windows XP
- Coding language - Java
- IDE - Net beans
- Database -MYSQL

login activity again. Registered user then needs to login in order to access the app. Validations are applied on all the textboxes for proper functioning of the app. Like information in each textbox is must that is each textbox, either it is of name, contact, password or confirm password, will not be empty while registering. If any such textbox is empty app will give message of information is must in each textbox. Also data in password and confirm password fields must match for successful registration. Another validation is contact number must be valid one that is of 10 digits. If any such validation is violated then registration will be unsuccessful and then user needs to register again. Message that app will display when one of the field is empty. If all such information is correct user will be directed to login activity for login into the app.

Creation Storage and Instance

The data owner has not control over the data after it is uploaded on cloud.In this module, the original data get encrypted into two different values.The data in each slice can be encrypted by using different cryptographic algorithm's and encryption key before storing them in the Cloud.

Find Collusion Module

In this Module, Receiver can find collusion occurring or not using calculating a distance.

Find Third-Party Module

In this Module, receiver can also find third-parties. Third party refers to another company making software for the original vendors product.