Chapter 39
Privacy Preserving and Efficient Outsourcing Algorithm to Public Cloud: A Case of Statistical Analysis

Malay Kumar
Department of Computer Science and Engineering, National Institute of Technology, Raipur, India

Manu Vardhan
Department of Computer Science and Engineering, National Institute of Technology, Raipur, India

ABSTRACT

The growth of the cloud computing services and its proliferation in business and academia has triggered enormous opportunities for computation in third-party data management settings. This computing model allows the client to outsource their large computations to cloud data centers, where the cloud server conducts the computation on their behalf. But data privacy and computational integrity are the biggest concern for the client. In this article, the authors attempt to present an algorithm for secure outsourcing of a covariance matrix, which is the basic building block for many automatic classification systems. The algorithm first performs some efficient transformation to protect the privacy and verify the computed result produced by the cloud server. Further, an analytical and experimental analysis shows that the algorithm is simultaneously meeting the design goals of privacy, verifiability and efficiency. Also, found that the proposed algorithm is about 7.8276 times more efficient than the direct implementation.

1. INTRODUCTION

The cloud computing has established and become a reliable technology. It provides convenient, on-demand and economical computing resources to the clients (Mell, & Grance, 2011). It enables every entity to execute, analyse and store large amount of data without even setting up own IT infrastructure. However, outsourcing of data and computations to the third-party cloud server beyond the physical

DOI: 10.4018/978-1-7998-8954-0.ch039

reach of the client increases the risk of exposing data to multiple security and privacy threats (Reed, Rezek & Simmonds, 2011). These potential threats come from the malicious behaviour of cloud server. For example, a malicious cloud server might increase or decrease the pace of execution and terminate the execution of problem prematurely and return some random arbitrary result (Lei et al., 2013). The cloud server not only produces wrong results, it might secretly record the private information about the computation. Later, an adversary may exploit these vulnerabilities and attempt to access private and confidential information. Therefore, the client become reluctant to harbour its data and computation to the cloud server. Subsequently, confidentiality of data and integrity of the result must be addressed before outsourcing the secret and valuable data (Kumar et al., 2017b). Further, Diffie has quoted that the "…cryptography will not be the solution of cloud computing due to economic reasons…" (Chen & Brook, 2010). In fact, cryptography only provides a partial solution of the outsourcing problem. Cryptographic solution secures the privacy of data, but make the computation on cipher highly complex and expensive. Further, (Kerschbaum, 2011) said that the non-secure local computation is more economical than cryptographically protected outsourced computing.

In this paper, the authors have made an attempt to provide secure and efficient solution for the classification problem. Basically, the authors computed covariance matrix, which could be used as a basic building blocks in various classification problems. There are mainly three approaches, which are widely used for classification, that are statistical, machine learning and neural network. In this paper, our point of focus is statistical classification. In literature many algorithms exist, which perform secure execution of classification problem such as (Du, Chen, & Han, 2004). But they use complex cryptographic primitives for securing the privacy of data. Also, they execute the problem in multiparty framework where the data are shared among the multiple parties, who have participated into the computation. Each party computes their part of the computation and produces the result. The final result is the union of the results from all participants. Moreover, each party involves into computation must have similar computation complexity. Motivate by the fact that statistical classification is an important problem. Its secured and efficient solution is required by a variety of clients across all domains. Also, no existing work offers a secure and efficient solution of the statistical analysis problem on the cloud in the best of our knowledge. Therefore, this paper proposes a new algorithm where any entity, which is resource-constrained or unable to procure or maintain their own IT infrastructure able to securely outsource the statistical classification problem to cloud server. The cloud server executes the problem on the client's behalf and produces the result.

The major contribution of the paper is presented in following points,

1. A new secure outsourcing framework has been proposed for the outsourcing of classification problem. In this framework the complex primitives which take polynomial time for execution will be executed on the cloud server, while rest the operations, which run in sub-quadratic or linear time executed on the client system.
2. The proposed algorithm introduces a privacy preserving efficient transformation operation. This operation secures the input and the output computation. The proposed algorithm also verifies the result executed on the cloud servers.
3. Finally, the algorithm is implemented, and result analysis is presented, which demonstrates significant time saving as compared to the direct implementation.