Secure Interpolation in the Cloud

Jordi Ribes González *

Department of Computer Engineering and Mathematics, Universitat Rovira i Virgili
Tarragona, Spain
jordi.ribes@urv.cat

1 Introduction

Cost-effective third-party (cloud) service providers offer very convenient data storage and computation services at a low cost, thus providing an attractive alternative to other forms of storage. However, outsourcing potentially sensitive datasets to an external cloud service provider poses many security and privacy concerns.

A natural approach to address these security concerns consists in applying cryptographic techniques. However, traditional symmetric-key encryption techniques fail to provide an efficient solution. A trivial option consists on encrypting all data using a symmetric-key encryption scheme and using the server only for Storage-as-a-Service. To perform a computation, all relevant data is retrieved, decrypted and computed on locally. Unfortunately, this solution may not be efficient, particularly if client devices have limited computational power or storage capacity, and requires a high bandwidth during queries. Alternative cryptographic schemes must be developed in order to overcome this obstacle.

In recent years there have been important advances in cryptographic techniques that allow to take advantage of the economical and functional benefits of cloud computing while securing the data. Two of these techniques are Homomorphic Encryption (HE) and Secure Multi-party Computation (SMC), both of which allow for remote computations over encrypted data.

Secure Multi-party Computation protocols are interactive protocols that allow a set of parties to jointly compute a function over their inputs. In SMC, parties keep their inputs private and engage in an interactive protocol with each other, so that at the end of the protocol each party learns the function evaluation and nothing else about inputs from other parties.

Homomorphic Encryption schemes allow computations to be performed directly on encrypted data, and they are classified according to the operations they support. Additive HE (such as [6]) and multiplicative HE schemes

* PhD advisor: Oriol Farràs Ventura

efficiently support a single operation on ciphertexts, that is, addition and multiplication respectively. Somewhat Homomorphic Encryption (SHE) schemes support any number of additions, but a limited number of multiplications. Fully Homomorphic Encryption (FHE) schemes support an arbitrary number of additions and multiplications on ciphertexts. Unfortunately, all known FHE schemes are computationally very expensive, which hinders their applicability in practice.

2 Secure Interpolation in the Cloud

Interpolation and regression techniques such as generalized least squares, polynomial regression and Spline interpolation are often used in practical applications, for example in order to predict values of some phenomena given a set of samples. They have a vast amount of applications, ranging from computer graphics to data analysis or experiment designs.

Outsourcing such computations to the cloud can offer numerous cost-saving and practical benefits, since applications often involve massive datasets and expensive computations. Data ubiquity is also very convenient, as such computations can involve data owned by multiple organizations, or they can be requested by multiple parties.

However, outsourcing computations to the cloud can pose security and privacy concerns, since applications usually involve potentially sensitive datasets. Therefore, we aim to provide practical solutions to enable clients to efficiently delegate an encrypted dataset to a semi-trusted server, in such a way that interpolation computations can be performed directly over encrypted data.

Following the previous discussion, we may look for a solution involving HE schemes. The main obstacle to applying this approach is that the considered computations often involve complex operations, requiring many additions and products. Some interpolation techniques involve computations that are currently challenging even when using FHE, including the computation of square roots, natural exponentiations or solving systems of linear equations. In order to overcome this obstacle, we look for tailored adaptations of the interpolation computations, so that we can apply HE schemes and enable the delegation of interpolation computations to the cloud.

3 Private Outsourced Kriging Interpolation

Kriging [1, 3, 5, 8] is a well-recognized form of linear interpolation widely used with datasets involving spatially correlated data. It aims at predicting the value of some phenomena at an unobserved location in a two-dimensional region. This interpolation method was designed with geo-statistical applications in mind (*e.g.* to predict the best location to mine within a region, based

on the mineral deposits found at previous boreholes), but has also found applications in a variety of settings including remote sensing, real-estate appraisal and computer simulations. Kriging has been identified as a good candidate process to be outsourced to the cloud, based on the practical and legislative requirements of industrial users [2, 4].

Based on a recent work carried out in conjunction with James Alderman, Benjamin Curtis, Oriol Farràs and Keith M. Martin, we present a method for the efficient private outsourcing of Kriging interpolation. The proposed solution uses a tailored modification of the Kriging algorithm in combination with additively homomorphic encryption, allowing crucial information relating to measurement values to be hidden from the cloud service provider. Moreover, with the exception of the high one-time cost of encrypting the dataset, the remaining client-side processes are very efficient. We evaluate the performance of our solution through an implementation in Python 3.4.3, using the PHE library [7].

Since the approach followed for Kriging interpolation is applicable to other interpolation techniques and statistical tools, a next step in this line of work is to develop solutions for other similar techniques.

The proposed results have been presented at the 5th Workshop on Encrypted Computing and Applied Homomorphic Cryptography (WAHC'17).

Acknowledgement. This work is partly supported by the European Comission through H2020-ICT-2014-1-644024 "CLARUS" and H2020-DS-2015-1-700540 "CANVAS", by the Government of Spain through TIN2014-57364-C2-1-R "SmartGlacis" and TIN2016-80250-R "Sec-MCloud", by the Government of Catalonia through Grant 2014 SGR 537, and by COST Action IC1306.

References

[1] J.-P. Chilès and P. Delfiner. Multivariate methods. *Geostatistics: Modeling Spatial Uncertainty, Second Edition*, pages 299–385, 1999.

[2] CLARUS: User centered privacy and security in the cloud. `http://clarussecure.eu`.

[3] N. Cressie. Statistics for spatial data. *Terra Nova*, 4(5):613–617, 1992.

[4] InGeoCloudS: inspired geo-data cloud services. `https://www.ingeoclouds.eu/`. Accessed: 11/12/2016.

[5] D. Krige. A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy*, 52(6):119–139, 1951.

[6] P. Paillier. Public-key cryptosystems based on composite degree residuosity classes. In *International Conference on the Theory and Applications of Cryptographic Techniques*, pages 223–238. Springer, 1999.

[7] python-paillier: a library for partially homomorphic encryption in python, Data61|CSIRO. `https://github.com/NICTA/python-paillier`, 2016. Accessed: 11/12/2016.

[8] H. Wackernagel. *Multivariate geostatistics: an introduction with applications*. Springer Science & Business Media, 2013.