Towards Incremental End-to-End Privacy Preserving Data Classification

Rania Talbi
INSA Lyon - LIRIS
Lyon, France
rania.talbi@insa-lyon.fr

Abstract

Nowadays, machine learning algorithms are used in a myriad of domains, such as medical diagnosis, fraud detection and user behavior analysis. However, in some of these cases, it is important that the data manipulated by these algorithms remain confidential. In order to address this paradox, the domain of privacy preserving machine learning has emerged. In this work, we describe the design principles of an outsourced incremental data classification scheme that satisfies this privacy constraint, while maintaining good classification accuracy and computational performance.

1 Context and Research Problem Statement

With the pervasiveness of computer devices and digital services, huge amounts of data are nowadays continuously generated and collected. Machine learning is an increasingly popular set of tools which are used to extract hidden yet valuable knowledge from this data. These tools can be quite beneficial in many application domains, such as medical diagnosis, fraud detection, user behavior analysis and many others. However, using its conventional methods over sensitive data may lead to serious individual or collective privacy breaches by leaking sensitive information about data owners. In this paper, we are interested in privacy preserving machine learning methods that aim to efficiently extract useful knowledge from encrypted data.

2 Related work

In this initial phase of our study, we are only interested in privacy preserving data classification methods. It is well known that as a supervised machine learning problem, data classification generally consists of two phases : a learning phase and a prediction phase. In a structured literature review, we have noticed that most of the state of the art's privacy preserving data classification solutions only focus on one of these phases and rarely consider both. In the first case, the objective is to protect training data, as well as the learnt classification model, which is the case of the solution proposed by Kikuchi et al. [3]. On the other hand, methods interested in the second classification phase only focus on protecting the input data used for prediction, as well as the corresponding output, such as the work proposed by Bost et al. [1]. As far as we know, the only state of the art solution that provides a scheme for preserving privacy during the entire classification process was proposed by Liu et al. [5]. However, this solution does not entirely protect the classification model, where some of its parameters are revealed in their plaintext form to an external service provider.

The privacy guarantees described above can be achieved using either non-cryptographic techniques, like data randomization [7] or cryptographic techniques, such as homomorphic encryption [1]. The first class of techniques ensures good computational performance, but somehow deteriorates the classifier's precision, contrarily to second class that provides robust privacy grantees and maintains good classification accuracy, but causes high computational overhead. This overhead becomes even more visible in the case of solutions that rely on secure multiparty computation protocols with a high number of collaborators. One way to reduce this communication overhead is to outsource the privacy preserving computations required during the learning or prediction phases to a single external party, as in [5].

Another interesting research area in this domain is Privacy Preserving Incremental Data Classification. Very few research works considered this problem, such as the proposal presented by Samet et al. [6]. The main challenge in the case, is to maintain a good performance in terms of execution time, in order to be able to handle high rate incoming training data and use them to properly update a classification model in a privacy preserving manner.

As a first step towards the ambitious goal of implementing an efficient privacy preserving machine learning framework, we were interested in designing an outsourced incremental data classification scheme that ensures privacy during the entire classification process and maintains in the same time a good computational performance and classification efficiency. As far as we know, there exists no previous works that were interested in addressing all of the aspects described above simultaneously.

3 Proposed Approach

In order to address the research problem discussed above, we propose the DAPPLE system or DynAmic Privacy Preserving machine LEearning Framework. Although, DAPPLE is intended to implement multiple Privacy Preserving Machine LEearning solutions on the long term, its first version only focuses on efficiently implementing an outsourced privacy preserving incremental data classification scheme that relies on the VFDT algorithm [2]. On one hand, this system offers a privacy preserving incremental decision tree learning service over encrypted training data continuously sent by multiple data owners. On the other hand, the learnt classifier is used to respond to encrypted classification queries, without having access neither to their content, nor to the resulting classification output. To make this possible, The DAPPLE system involves five basic components which are the following :

Data owners (DO_i) These parties continuously outsource chunks of encrypted training data to the DAPPLE system, in order to build and incrementally update a global classification model.

Classification Queriers (Q_j) these are the final users of the system. They send encrypted classification queries to the DAPPLE system that predicts the corresponding class label using a private classification model trained earlier, without knowing the plaintext values of these queries, nor the obtained classification output.

Computation units (U_1,U_2) : Our system has two semi-honest computation units U_1 and U_2. These two collaborate in a set of secure two-party computation protocols to establish the outsourced privacy preserving classification service over encrypted data.

Key Management Unit (*KMU*) : This component is a trusted unit which generates and manages all cryptographic keys used in the DAPPLE system.

In order to enable computations over encrypted data, we have used an additively homomorphic cryptosystem called DT-PKC that was initially proposed by Liu et al. [4]. This cryptosystem provides a distributed double trapdoor decryption mechanism using two partial strong decryption keys SK^1 and SK^2 which are held by the system's computation units U_1 and U_2.

To create a privacy preserving implementation of the VFDT algorithm, we have proposed a set of secure building blocks that enable computations such as : square root, logarithm and comparison over encrypted data. These building blocks are constructed using two-party secure computation protocols between the system's computation units. These protocols rely on the partial decryption mechanisms of the DT-PKC cryptosystem, as well as additive and multiplicative blinding techniques that aim to prevent the computation units from having access to the plaintext values of encrypted data using their strong partial decryption keys. These techniques simply add randomness to encrypted data using the homomorphic properties of the DT-PKC cryptosystem. The most frequently computed metric in our privacy preserving implementation of the VFDT algorithm is Entropy, which is used to determine node splitting criterion during the learning phase. A naive approach to compute this metric over encrypted data is to combine multiple elementary building blocks. However, the computational cost of this strategy would be extremely high due to the numerous interactions between the system's computation units. To reduce this time overhead, we have proposed a single inline building block that computes the following equivalent formula of entropy

$$E(S) = log_2(|S|) - \frac{1}{|S|} \cdot \sum_{i=1}^{i=K} (|C_i|.log_2(|C_i|)$$

using multiplicative blinding as well as DT-PKC's partial decryption mechanisms. We have also used parallel computing to simultaneously evaluate entropy for different data attributes and node splitting conditions to improve our solution's computational performance.

4 Experiments and Preliminary Results

In order to evaluate our solution, we have used a synthetic dataset for fraud detection in a B2B network. This dataset contains 1000 bank transaction with 9 attributes each and consists of 3 classes : fraudulent, suspect and non fraudulent transactions. The preliminary results of our experiments show that our solution maintains the classification accuracy of the original version of the VFDT algorithm that operates on plaintext data. In both cases we obtain a 92% classification accuracy. As for the evaluation of the computational performance of our solution, as shown in Figure 1, we were able to observe the effect of our optimization strategies in reducing training time overhead. We have noticed that using an inline building block to evaluate Entropy (OPT-1) instead of combining multiple elementary building blocks (NAIVE approach) reduces training time by 4%. In another hand, using parallel computing (OPT-2) improves training time by 6% in comparison to the naive approach. As for the prediction phase, we were able to compare our solution to the Ciphermed-DT solution proposed by Bost et al. [1]. As shown in Figure 2, our approach was twice as fast as Ciphermed-DT in terms of classification time. This is due to the fact that the latter combines multiple building blocks that make use of heterogeneous homomorphic cryptosytems and requires additional steps to convert ciphers from one cryptosystem to another to make this combination possible.

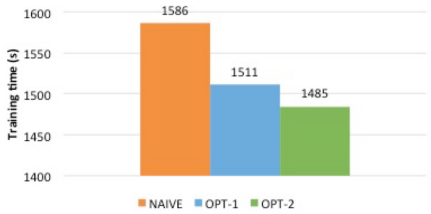

Figure 1. Training time comparison between our naive approach (NAIVE), our optimized solution that uses inline building block for Entropy computation (OPT-1) and the solution that uses parallel computing (OPT-2)

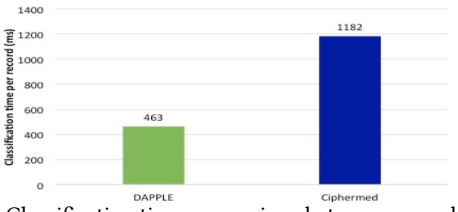

Figure 2. Classification time comparison between our solution and the CIPHERMED-DT solution [1]

5 Conclusion

In this paper, we presented DAPPLE, a framework for outsourced privacy preserving incremental data classification. Our preliminary results seem to be promising but still need to be confirmed using more consistent and complex datasets. During the rest of this PhD project, we intend to create an optimized implementation of DAPPLE using SPARK framework. We also consider extending our solution to cover other machine learning algorithms. The objective is to provide the first efficient fully equipped privacy preserving machine learning toolkit over encrypted data. Finally, we intend to test our proposal in practical use cases and make sure of its efficiency in real life applications.

Acknowledgments

This work is conducted under the supervision of Prof. Sara Bouchenak. It is supported by the French National Research Agency (ANR), through the SIBIL-Lab project (ANR-17-LCV2-0014).

References

[1] Raphael Bost, Raluca Ada Popa, Stephen Tu, and Shafi Goldwasser. 2015. Machine learning classification over encrypted data.. In *NDSS*.

[2] Pedro Domingos and Geoff Hulten. 2000. Mining high-speed data streams. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 71–80.

[3] Hiroaki Kikuchi, Kouichi Ito, Mebae Ushida, Hiroshi Tsuda, and Yuji Yamaoka. 2013. Privacy-Preserving Distributed Decision Tree Learning with Boolean Class Attributes. In *Advanced Information Networking and Applications (AINA), 2013 IEEE 27th International Conference on*. IEEE, 538–545.

[4] Ximeng Liu, Robert H Deng, Kim-Kwang Raymond Choo, and Jian Weng. 2016. An efficient privacy-preserving outsourced calculation toolkit with multiple keys. *IEEE Transactions on Information Forensics and Security* 11, 11 (2016), 2401.

[5] Ximeng Liu, Rongxing Lu, Jianfeng Ma, Le Chen, and Baodong Qin. 2016. Privacy-preserving patient-centric clinical decision support system on naive Bayesian classification. *IEEE journal of biomedical and health informatics* 20, 2 (2016), 655–668.

[6] Saeed Samet, Ali Miri, and Eric Granger. 2013. Incremental learning of privacy-preserving Bayesian networks. *Applied Soft Computing* 13, 8 (2013), 3657–3667.

[7] Lita Yang and Boris Murmann. 2017. Approximate SRAM for energy-efficient, privacy-preserving convolutional neural networks. In *2017 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*. IEEE, 689–694.