Automatica 96 (2018) 314–325

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Privacy preserving distributed optimization using homomorphic encryption[☆]

Yang Lu, Minghui Zhu [*]

School of Electrical Engineering and Computer Science, Pennsylvania State University, 201 Old Main, University Park, PA, 16802, USA

ARTICLE INFO

Article history:
Received 29 June 2017
Received in revised form 9 May 2018
Accepted 30 May 2018

Keywords:
Distributed optimization
Privacy
Homomorphic encryption

ABSTRACT

This paper studies how a system operator and a set of agents securely execute a distributed projected gradient-based algorithm. In particular, each participant holds a set of problem coefficients and/or states whose values are private to the data owner. The concerned problem raises two questions: how to securely compute given functions; and which functions should be computed in the first place. For the first question, by using the techniques of homomorphic encryption, we propose novel algorithms which can achieve secure multiparty computation with perfect correctness. For the second question, we identify a class of functions which can be securely computed. The correctness and computational efficiency of the proposed algorithms are verified by two case studies of power systems, one on a demand response problem and the other on an optimal power flow problem.

1. Introduction

In the last decades, distributed optimization has been extensively studied and broadly applied to coordinate large-scale networked systems (Bertsekas & Tsitsiklis, 1997; Bullo, Cortés, & Martínez, 2009; Zhu & Martínez, 2015). In distributed optimization, the participants collectively achieve network-wide goals via certain mechanisms driven by data sharing between the participants. Such data sharing, however, causes the concern that private information of legitimate entities could be leaked to unauthorized ones. Hence, there is a demand to develop new distributed optimization algorithms that can achieve network-wide goals and simultaneously protect the privacy of legitimate entities.

The problem of interest is closely relevant to the one where a group of participants aim to compute certain functions over their distributed private inputs such that each participant's inputs remain private after the computation. To protect privacy, two questions need to be answered (Lindell & Pinkas, 2009): (i) How to securely compute given functions with distributed private inputs so that the computation process does not reveal anything beyond the function outputs? This question is referred to as secure multiparty computation (SMC) (Cramer, Damgård, & Nielsen, 2015; Hazay & Lindell, 2010). (ii) Which functions should be computed in the first place so that the adversary cannot infer private inputs of benign participants from function outputs? In this paper, we refer to this question as input–output inference (IOI).

Literature review on homomorphic encryption. For the question of SMC, homomorphic encryption is a powerful tool, and has been applied to various problems including, e.g., statistical analysis (Shi, Chan, FxPal, Chow, & Song, 2011) and data classification (Yang, Zhong, & Wright, 2005). It is because homomorphic encryption allows certain algebraic operations to be carried out on ciphertexts, thus generating an encrypted result which, when decrypted, matches the result of operations performed on plaintexts. A detailed literature review on homomorphic encryption is provided in Section 9.1 of the complete version of this paper (Lu & Zhu, 2018). It is worthy to mention that private key fully (and surely partially) homomorphic encryption schemes and public key partially homomorphic encryption schemes can be efficiently implemented (see Section 9.1 of the complete version Lu & Zhu, 2018). Recently, homomorphic encryption has been applied to control and optimization problems, e.g., potential games (Lu & Zhu, 2015b), quadratic programs (Shoukry, Gatsis, Alanwar, Pappas, Seshia, Srivastava, & Tabuada, 2016), and linear control systems (Farokhi, Shames, & Batterham, 2017; Kogiso & Fujita, 2015). Most existing homomorphic encryption schemes only work for binary or non-negative integers. In contrast, states and parameters in most distributed optimization problems are signed real numbers.

Literature review on IOI. SMC alone is not enough for data privacy. Even if a computing scheme does not reveal anything beyond the outputs of the functions, the function outputs themselves could

[☆] This work was partially supported by NSA grant H98230-15-1-0289 and NSF grant CNS-1505664. The material in this paper was partially presented at the 5th IFAC Workshop on Distributed Estimation and Control in Networked Systems, September 10–11, 2015, Philadelphia, PA, USA. This paper was recommended for publication in revised form by Associate Editor Shreyas Sundaram under the direction of Editor Christos G. Cassandras.

[*] Corresponding author.
E-mail addresses: yml5046@psu.edu (Y. Lu), muz16@psu.edu (M. Zhu).

https://doi.org/10.1016/j.automatica.2018.07.005
0005-1098/© 2018 Elsevier Ltd. All rights reserved.

tell much information about the private inputs. This raises the second question of IOI. A representative set of papers, e.g., Clarkson, Myers, and Schneider (2009), Mardziel, Magill, Hicks, and Srivatsa (2011), Mardziel, Hicks, Katz, and Srivatsa (2012), adopted the method of belief set. In particular, each participant maintains a belief which is an estimate of the other participant's knowledge. When the belief is below certain threshold, the participant continues computation. Otherwise, the participant rejects the computation. This approach is only applicable to discrete-valued problems with constant private inputs. To the best of our knowledge, the issue of IOI in real-valued time series, e.g., generated by dynamic systems, has been rarely studied.

Contributions. This paper investigates how a system operator and a set of agents execute a distributed gradient-based algorithm in a privacy preserving manner. In particular, on the one hand, each agent's state and feasible set are private to itself and should not be disclosed to any other agent or the system operator; on the other hand, each of the agents and the system operator holds a set of private coefficients of the component functions and each coefficient should not be disclosed to any participant who does not initially hold it. The privacy issue of the gradient-based algorithm is decomposed into SMC and IOI.

For the question of SMC, we first propose a private key fully homomorphic encryption scheme to address the case where jointly computed functions are arbitrary polynomials and the system operator could only launch temporarily independent attacks, i.e., at each step of the gradient-based algorithm, the system operator only uses the data received at the current step to infer the agents' current states, but does not use previous data to collectively infer the agents' states in the past. We notice that the assumption of temporarily independent attacks is crucial for the private key homomorphic encryption setting and the security might be compromised if this assumption does not hold. Similar assumptions have been widely used in database privacy (Bhaskar, Bhowmick, Goyal, Laxman, & Thakurta, 2011; Dwork, Naor, Pitassi, & Rothblum, 2005; Zhu, Li, Zhou, & Yu, 2017). Please refer to Remark 3.2 for the detailed discussion on this assumption. To deal with real numbers, we propose mechanisms for transformations between real numbers and integers as pre and post steps of plain homomorphic encryption schemes for integers, and provide the condition that the key should satisfy to guarantee the correctness of decryption and post-transformation. The proposed technique to handle real numbers can be applied to arbitrary homomorphic encryption schemes. All the agents encrypt their states and coefficients by a key which is unknown to the system operator. The system operator computes the polynomial functions with the encrypted data and sends the results to the target agent that initiates the computation. The target agent can then perform decryption by the same key. We prove that the algorithm correctly computes the functions and meanwhile prove by the simulation paradigm that each agent does not know anything beyond what it must know (i.e., its inputs and the outputs of its desirable functions) and it is computationally hard for the system operator to infer the private data of the agents.

We then propose a public key partially homomorphic encryption computing algorithm to address the case where jointly computed functions are affine and the system operator could launch causal attacks. The secure computation of affine functions is carried out by the Paillier encryption scheme. Similar techniques are applied to handle real numbers. We prove that the proposed algorithm computes the correct values of the functions and after the computation, each participant does not learn anything beyond what it must know and it is computationally hard for the system operator to distinguish the private data of the agents.

For the question of IOI, we provide a control-aware definition which, informally, states that a function is secure to compute if, given the output of the function, the adversary cannot uniquely determine the private inputs of any participant and meanwhile the uncertainty about the inputs is infinite. The definition is inspired by the notion of observability in control theory and is consistent with the uncertainty based privacy notions in database analysis (Brodsky, Farkas, & Jajodia, 2000; Marks, 1996). For a class of quadratic functions, we derive sufficient conditions of IOI by exploiting the null vectors of the coefficient weight matrices and the constant terms.

The correctness and computational efficiency are verified by two case studies of power systems, one on a demand response problem and the other on an optimal power flow problem.

In our earlier paper (Lu & Zhu, 2015b), homomorphic encryption was applied to discrete potential games. In Lu and Zhu (2015b), the issue of IOI was not discussed and the proofs were omitted. The current paper studies a class of distributed algorithms which have been widely used to solve distributed optimization, convex games and stochastic approximation. In the current paper, both the issues of SMC and IOI are studied with formal proofs. For the issue of SMC, the current paper adopts two different homomorphic encryption schemes which are based on some computationally hard problems and achieve higher privacy level than that of Lu and Zhu (2015b). Due to space limitation, the proofs, case studies and detailed literature review on homomorphic encryption are omitted in the current paper and included in the complete version (Lu & Zhu, 2018).

Literature review on other secure computation techniques. Besides homomorphic encryption, several other techniques have been adopted to address privacy issues in control and optimization.

The first branch of works uses differential privacy (Dwork, 2006; Dwork & Roth, 2014) as the security notion, e.g., Hale and Egerstedt (2015), Han, Topcu, and Pappas (2017), Huang, Mitra, and Dullerud (2012), Ny and Pappas (2014) developed differentially private algorithms for distributed optimization, consensus and filtering problems, respectively. Differentially private schemes add persistent randomized perturbations into data to protect privacy (Geng & Viswanath, 2014). For control systems, such persistent perturbations could potentially deteriorate system performance. In contrast, homomorphic encryption schemes do not introduce any perturbation and are able to find exact solutions.

The second branch of works uses obfuscation techniques (Dreier & Kerschbaum, 2011) to protect coefficient confidentiality for optimization problems in cloud computing. Related works include, e.g., Borden, Molzahn, Lesieutre, and Ramanathan (2013), Borden, Molzahn, Ramanathan, and Lesieutre (2012), Wang, Ren, and Wang (2016), which studied optimal power flow and linear programming problems, respectively. Existing obfuscation techniques are only applicable to centralized optimization problems with linear or quadratic cost functions in order to have the property that inverting the linear obfuscation transformation returns the optimal solution of the original problem.

The third branch of works adopts the techniques of secret sharing, e.g., in our earlier work (Lu & Zhu, 2015a), the Shamir's secret sharing scheme (Shamir, 1979) was used to achieve state privacy for distributed optimization problems on tree topologies. In Lu and Zhu (2015a), the functions are restricted to be linear and the privacy of coefficients is not taken into account.

Notations. The vector $\mathbf{1}_n$ denotes the column vector with n ones. Given a finite index set Ω, let $[A_i]_{i\in\Omega}$ denote the column-wise stack of A_i for all $i \in \Omega$, where A_i's are matrices with the same number of columns. When there is no confusion in the context, we drop the subscript $i \in \Omega$ and use $[A_i]$. Given matrices $A_1, \dots, A_N$ with the same column number, let A_{-i} denote $[A_j]_{j\neq i}$. Given matrices $A_1, \dots, A_N$, denote by $\mathrm{diag}\{A_1, \dots, A_N\}$ the block diagonal matrix for which the sub-matrix on the ith diagonal block is A_i and all the off-diagonal blocks are zero matrices. For a matrix or vector