Proposition de recherche doctorale

Mécanismes cryptographiques pour la sécurité et la protection des données dans le Cloud Computing

Mots clés :

- **Directeur de thèse** : refik MOLVA
- **Co-encadrant(s)** :
- **Unité de recherche** : Laboratoire de recherche d'EURECOM
- **Ecole doctorale** : École Doctorale Informatique, Télécommunications, Électronique de Paris
- **Domaine scientifique principal**: Divers

Résumé du projet de recherche (Langue 1)

In distributed computing scenarios such as web applications, social networks, and cloud computing, data that is originally provided by users are handled by a number of potentially untrusted parties such as application servers, communication service providers, cloud service providers, etc. Most of these scenarios are motivated by significant advantages in terms of cost reduction, performance increase, and a variety of new service offers at a low cost or for free. Yet these advantages come at the expense of serious violations for users' privacy and security that are akin to current distributed computing scenarios. Classical security mechanisms such as data encryption unfortunately fall short of countering these privacy and security violations in the original setting whereby the performance advantages and the resulting extensive service offerings can still be kept. The goal of this thesis is to come up with solutions for user security and privacy that would not impact the advantages of current distributed computing applications in terms of increased performance, reduced cost, and extensive service offering at low cost or for free.

Résumé du projet de recherche (Langue 2)

The work will initially focus on two generic use cases that are part of popular distributed computing scenarios: data lookup and data aggregation. Privacy preserving data lookup In any privacy preserving data management system, the very first privacy requirement for the user is data confidentiality. A user initially encrypts the data it wishes to outsource and uploads it to the (distributed) system. While classical encryption algorithms are considered a good candidate for this requirement, they often obstruct the second advantage of a powerful data management system which is computation outsourcing. Indeed, once its data uploaded into the system, a user may want to query the data for either a lookup or a retrieval of some information while still assuring privacy. The user may of course download the whole encrypted data, decrypt it and perform the required operation but such a naïve solution unfortunately does not take advantage of the computational capabilities of the overall data management system. Current solutions based on either Private information retrieval (PIR) and/or searchable encryption have received a lot of attention. In PIR (private information retrieval)[1,2,3,4,5], a user retrieves a specific data from a database located in a server. The only privacy goal in PIR is access privacy whereby the server should not discover which data a user is interested in. Note that PIR does not ensure privacy of data in the database. Furthermore, searchable encryption [6,7,8,9,10,11] allows a user to verify whether some specific "keywords" exist in the remote data. With such techniques, user privacy is guaranteed thanks to the encryption of the queries and the stored data. Most of the previously mentioned PIR and searchable encryption techniques target single server settings. Therefore, existing PIR and searchable encryption techniques will be revisited in order to design dedicated primitives for a distributed environment. Furthermore, the only "privacy" goal in PIR is access privacy whereby the server should not discover which data a user is interested in. Therefore PIR alone is not sufficient to assure storage privacy. Similarly, in existing searchable encryption solutions the result (which is a Boolean result) originating from a search query is known to the adversary. Hence, the node who processes the query can obtain some knowledge on users queries. In a dedicated security model, the adversary should sometimes not even learn anything about queries or results. Finally, the majority of privacy preserving lookup schemes do not consider the data management system as being malicious while executing the required processing operation. The new privacy preserving primitives proposed in the thesis will give the user the ability to verify the correctness of operations. Privacy preserving data aggregation Secure data aggregation has been studied in the context of wireless sensor networks [12,13] whereby nodes aggregate their data to the current intermediate aggregate they receive in order to forward the resulting value to the next hop towards the sink. Thanks to the use of homomoprhic encryption algorithms, nodes can perform correct aggregation operations over encrypted data. Homomorphic encryption [14,15,16] allows a third party to perform meaningful computation over encrypted data. Most of the very well-known homomorphic encryption techniques support very limited operations such as simple addition [14] or multiplication [15] or both [16]. Current secure aggregation techniques rely on the existence of a trusted central node (the sink) and assume the correct behavior of all participating nodes. In this thesis, aggregation operations have to be distributed to all nodes among which some of them may be malicious. Therefore techniques such as secure multi-party computation (SMC) [17,18,19,20] will be investigated in order assure the correctness of the resulting aggregate even with the existence of some misbehaving nodes. Finally, the newly proposed privacy preserving and verifiable aggregation mechanisms will consider the trade-off between security and efficiency by considering more realistic adversary models.