Journal of Network and Computer Applications xxx (2018) 1–13

Contents lists available at ScienceDirect

Journal of Network and Computer Applications

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

Interoperable, dynamic and privacy-preserving access control for cloud data storage when integrating heterogeneous organizations

Christian Esposito

Department of Electrical Engineering and Information Technology, University of Naples Federico II, Via Claudio 21, 80125 Napoli, Italy

ARTICLE INFO

Keywords:
Access control
Cloud data storage
Ontology
Ontology matching
Trust management
Privacy-preserving authentication
Pseudonym management

ABSTRACT

Cloud computing is extensively used as an integration means in varies application domains, spanning from the healthcare to the manufacturing, aiming at achieving an easy-to-access and elastic data storage and exchange among heterogeneous and geographically sparse organizations. This cloud-based integration poses crucial security issues related to the data protection from unauthorized access to the outsourced data, which calls for a proper access control solution. However, the heterogeneity among the organizations exacerbates this problem, demanding an interoperable authorization scheme, where multiple access control models must co-exist. The current literature is rich of academic solutions and standards to have an interoperable exchange of security policies and definition of authorization rules, but lacks an effective support to let different access control models to fully coexist. Moreover, the possibility of stealing authentication credentials and authorization claims paves the way to conducting masquerading attacks that cannot be treated by traditional static authorization solutions, but more dynamic approaches are needed. Last but not least, the continuous interaction of users with the cloud over the time has the vulnerability of exposing personal information to malicious adversaries and to let them trace the user activities. In this work, we propose to solve these three issues by having an ontology-based access control solution, to encompass trust within the authorization process and to use pseudonyms to preserve the user privacy.

1. Introduction

The problem of obtaining higher computing power and data storage capabilities on demand without having to own a proper computing hardware has been central in the last decades, with the profusion of research efforts and the advent of technological evolutions within the computer science and engineering (Dhar, 2012). In fact, over the years, we have witnessed a tremendous increase of the data to be managed and the progressive dematerialization of documents in every sector, from healthcare to the manufacturing. Despite the reduction of the costs for the data storage and computing hardware, a larger portion of the budget of several companies and organizations, where ICT is more and more pervasive and crucial, is starting to be progressively eaten by the increasing costs of managing the internal ICT capabilities. In fact, the traditional approach of having and maintaining commodity clusters is quite expensive to operate. On the contrary, outsourcing such capabilities, by having centralized facilities operated by third-party utilities providing the demanded computing and storage resources by using an Internet connection, is the approach that such companies are looking for in order to lower their operational costs for their ICT capabilities without giving up their demanding data management and analytics, which are vital for their mission and business. At the beginning, there were multiple technologies, such as grid computing, cluster computing or utility computing, that from different perspectives, have tried to tackle such a problem. After Web 2.0, the cloud computing has rapidly imposed itself as the most suitable solution to the elastic provisioning of computing and storage resources. Cloud computing (Pallis, 2010) overlaps with many existing previous technologies for computing virtualization and provisioning; however, it realized a progressive shift from an infrastructure-based to an application-based approach, which motivated its success and widespread adoption. In fact, the cloud computing moved from a mere solution that mimics and makes accessible over the Internet the traditional physical computing hardware (*i.e.*, the Infrastructure-as-a-Service (IaaS) mode that is similar to the one of grid computing), to more advanced provisioning models, where more abstract resources and services are delivered to the customers over the Internet (*e.g.*, Platform-as-a-Service (PaaS) or Software-as-a-Service (SaaS)). The widespread availability of the Internet connection due to

E-mail address: christian.esposito@unina.it.

https://doi.org/10.1016/j.jnca.2018.01.017
Received 31 May 2017; Received in revised form 12 January 2018; Accepted 27 January 2018
Available online XXX

the next generation of cellular and wireless networking, and the existence of cloud platforms with a massive amount of computing resources (which can be further enlarged thanks to the possibility of seamlessly federating multiple clouds (Grozev and Buyya, 2014; Esposito et al., 2013)) is currently paving the way for a radical rethinking of multiple traditional ICT systems, where the cloud plays the crucial role, such as the sensory networks (Esposito et al., 2017), the infrastructures for healthcare-related data management (Casola et al., 2016), or the manufacturing processes (Esposito et al., 2016a), just to cite the most prominent ones.

In an increasing number of application domains, cloud computing is extensively used as an integration means (Buyya et al., 2009; Li et al.,), in order to allow the seamless data flow and service provision among heterogeneous organizations and/or systems that have to collaborate among themselves. In such contexts, the cloud plays a critical role in the effective realization of the collaboration, both because it contains sensitive information related the users and/or companies that are interconnected, or because the data is valuable for the achievement of the mission of the applications running on top of the realized collaboration. As a concrete example, healthcare data can contain private information on patients, such as HIV test outcomes, psychological profiles or social security numbers, whose exposure can compromise the reputation, finance and/or life of the patients. Cloud manufacturing vehicles business critical data of the interconnected firms, such as confidential and copyrighted information on a particular manufacturing design, production plan or commercialization strategy, that malicious employees may use for blackmailing their employers, or competitors may be willing to obtain in order to copy innovative upcoming products or improve their own products to the detriment of the competitor's products. Last, the sensing data may reveal habits of the users so as to let thefts to plan a house burgled. In addition to the protection of the data confidentiality, our daily activities are tightly coupled with the successful behavior of the cloud platforms, which must be protected against possible cyber-attacks aiming at compromising their availability and/or their correct behavior. As a practical example, a Denial of Service attach can target a cloud platform hosting the management services of a healthcare provider, making them unavailable so that doctors are not able to retrieve their patients' documents for a certain time window, or a solution of cloud manufacturing may be compromised causing a sudden stop of the production and shipping activities of the affected firms. Another example is the injection of false data or the tampering of real sensing data so that the application running within the IoT may take the wrong decisions, with the effect of causing losses of human lives, of money and/or the application reputation. Therefore, security and privacy of the cloud computing is starting to be demanding (Li et al.,), since the data hosted in the cloud is sensitive, and the cloud is itself important for the successful execution of several critical processes.

When integrating multiple organizations, the cloud data storage solution has to deal with multiple kinds of security and privacy issues. First of all, each organization is typically characterized by a proper access control model and relative authorization policies, which must coexist and interoperate with the ones of the other organizations, in order to achieve an effective collaboration. It is impossible to impose a single access control model, such as a role-based or an attribute-based one. This is mainly due to the fact that there is no agreement on the most suitable and effective access model when integrating multiple organizations. Moreover, having a single model implies to rethink the internal access control rules of the integrated organizations and is not reasonable or profitable to undergo. Second, the interactions with the cloud occur outside the protected context of an organization and typically conveyed by the public Internet. This causes the rise of multiple vulnerabilities that can be exploited by hackers, so as to intercept the exchanged messages and steal valuable personal information to be used to architect and perform masquerading attacks. The traditional static access control models, which are only based on the provided claims and the beforehand agreed policies, are inefficient to cope with such a kind of attacks, even if coupled with encryption (Li et al., 2018). This calls out for a more dynamic solution, where the previous behavior of the users must be considered when deciding about the authorization of incoming requests. Last, during the authorization of a request, the exchanged authentication data can be used to trace the user activities or to obtain personal data. Privacy must be preserved also during such a process, so as to avoid possible vulnerabilities and prevent misuses.

Our work has the ambition of resolving such problems with three contributions that can be summarized as follows:

1. Access control policies and the relative satisfaction have been formalized according to an ontological formalisms, by means of the approach presented in (Esposito et al., 2016b), and multiple ontologies (relative to the multiple access models of the interconnected organizations) have been matched, so as to let them be interoperable;
2. When outsourcing data to the cloud a criticality level must be assigned, and each requesting user must have associated a trust degree that is continuously updated based on their behavior. These two information are used in the proposed solution to realize a second stage of authorization to grant or deny a request that already passed the satisfaction control of the security policies.
3. Authentication and reputation management is not relative to the true identity of the users, but to pseudonyms properly managed so as to avoid the exposure of personal data or the traceability of the user habits and activities.

In the rest of the paper, the problem tackled by this work will be illustrated and some related work on the topic of access control will be briefly overviewed (in Section 2). Later, the proposed approach will be described in Section 3, where we have dedicated a given subsection for each of the above-mentioned contributions. Section 4 presents our prototype and an empirical evaluation of the quality of our solution, while Section 5 concludes such a paper with some lesson learnt and a plan for future work.

2. Background and related work

As mentioned, cloud computing is being increasingly used as an integration means among collaborating users belonging to different organizations, spread across a given region, country or even around the world, thanks to the easiness in accessing the hosted resources through the Internet. Fig. 1 depicts such a use, where each organization is characterized by a given internal information system, used by users to store their documents. Such an internal information system is characterized by a given set of security policies for the secure access of the stored documents. Specifically, such policies are verified against the user claims so as to determine if a received request can be granted or not, and what kind of operation (*e.g.*, a read or a write) the user is allowed to do on the accessed document. When these organizations need to exchange data among themselves, a cloud computing platform can be used to hold the data to be shared. Such a platform can be used directly by the users or we can have the internal information systems transparently mediating such an interaction. The first problem is that an outsourced datum is subject to the satisfaction of the security policies of its organization of reference (in the figure, these are the ones of the first organization), while the requesting user provides its security claims according to the security policies established by its employing organization (in the figure, these are the ones of the second organization). Typically, the two sets of policies are not the same, but diverge both by assuming a different access control model, but also by using different terms and semantics in expressing the policies. So, the first issue is related to have a technical, syntactical and semantic interoperability among these different security policies, when data are exchanged by means of the cloud among heterogeneous organizations.

The current literature is rich of methods, and relative standards, in order to have an interoperable exchange of the security policies based