UNIVERSIDAD DE MURCIA

FACULTAD DE INFORMÁTICA

Enhancing User-Centric Identity Management Systems with Reputation Models in Distributed Environments

Mejora de Sistemas de Gestión de Identidades Centrados en el Usuario mediante Modelos de Reputación en Entornos Distribuidos

D. Ginés Dólera Tormo

2014

The following Thesis is a compilation of the next published articles, being the PhD student the main author in all of them:

1. Ginés Dólera Tormo, Félix Gómez Mármol, Gregorio Martínez Pérez, "Towards the Integration of Reputation Management in OpenID", Computer Standards & Interfaces, Special Issue on Secure Mobility in Future Communication Systems under Standardization, vol. 36, no. 3, pp. 438-453, March 2014
   http://dx.doi.org/10.1016/j.csi.2013.08.018

2. Ginés Dólera Tormo, Félix Gómez Mármol, Joao Girao, Gregorio Martínez Pérez, "Dynamic and Flexible Selection of a Reputation Mechanism for Heterogeneous Environments", Future Generation Computer Systems, Special Issue on Trustworthy Data Fusion and Mining in Internet of Things, 2014
   http://dx.doi.org/10.1016/j.future.2014.06.006

3. Ginés Dólera Tormo, Félix Gómez Mármol, Joao Girao, Gregorio Martínez Pérez, "Identity Management: In privacy we trust. Bridging the trust gap in e-Health environments", IEEE Security & Privacy, Special Issue on Health IT Security and Privacy, vol. 11, no. 6, pp. 34-41, 2013
   http://dx.doi.org/10.1109/MSP.2013.80

Table of Contents

Abstract

I Motivation and Goals

In the last years, due to the great acceptance of communication systems in the Internet, its users are increasing the amount of exchanged information, among which sensitive and personal data are included. However, these users sometimes are not aware of how their personal data are being handled, or they do not know who can actually have access to such information.

Likewise, gathering knowledge about users is considered more and more appealing, becoming even a target for some organizations with commercial purposes. The collected information is mostly used within advertisement goals, or it is aimed to develop advanced attacks on specific targets extracted from such acquired data. Several services offered in the Internet are willing to collect information of the users, for instance using sign up forms.

It is fairly common nowadays to have to deal with sign up forms even for services which will be likely used only once, e.g., commenting an entry in a blog. Data asked for using these services is usually personal, like email address or birthdate; sometimes even more private data is asked, which might seem irrelevant for the provision of the service itself, such as telephone number or hobbies.

That not only results in having to remember many different usernames and passwords for each service, or expose ourselves to receive spam, but it also jeopardizes the privacy of the users. In many cases, when users provide their personal data, they do not really know how this data is going to be handled, who will be released this data to, or whether it could be used on marketing campaigns outside the service they are signing up, for instance.

Both privacy protection and control over the information collected by other entities about oneself are characteristics more and more sought by the users of any communication system. Besides, these issues are considered a right of the users in certain geopolitical environments, such as the European Union. In this context, we can find users that do not want to relate their private life as they interact with different web services, or users that want to avoid the services to collect information about their preferences or to build usage profiles. For instance, journalists willing to report controversial circumstances without being concerned about possible reprisals, militaries that cannot reveal their geographical location, or just as an additional security measure for any user of the communication systems in the Internet.

An elegant solution to these concerns has begun to spread through the usage of the identity management systems. Identity management systems establish trust relationships between different organizations, in such a way that the information of the users is handled by a trusted

entity, usually known as identity provider (for example, their city council, university, or any other trustworthy organization).

When users access a service in the Internet (e.g., an online shopping site, subscription to a news feed, commenting an entry in a blog, etc.), the authentication methods and users' data management is delegated to the identity provider. Thus, the identity provider is in charge of preserving the privacy of the users, as the users do not need to sign up and perform authentication directly into the service. Additionally, identity management systems provide Single Sign-On, since they allow the users to make use of a unique account (the one they have in the identity provider) to access different services repeatedly.

Diverse solutions and standards have been developed in order to define the communication between identity providers and Internet services. As representative examples, we can consider SAML or OpenID. However, these systems present numerous shortcomings even nowadays, especially in the aspects related to the control that the users have on their own information. Current solutions barely inform the users on where their personal information is going and, to a lesser extent, they allow the users to decide which data may or may not be transferred to other entities.

Furthermore, in those cases where the users are informed about which services are requesting their personal data, they do not know how much they can trust those services. In other words, the users do not know how their information is going to be handled, or if the requested service will fulfill their expectations.

These difficulties are increased even further when identity management systems have to be deployed in distributed environments, such as P2P networks, where trust relationships cannot be established through static centralized servers. In such cases, establishing static contracts in order to set up the quality of service between a service provider and the clients, such as Service Level Agreements (SLA), is hardly applicable. Such limitation raises the necessity of deploying additional mechanism to manage trust.

An alternative that is being successfully deployed in recent years in order to manage trust in distributed environments is introduced by reputation management systems. In these systems, the trust that can be placed on a given entity (such as a network node, a service, or even a user), is not established by static agreements, but it is rather based on past experiences that others have been having with such an entity.

Reputation management systems attempt to predict the behavior of an entity from the behavior that such an entity has had in the past. The reputation is usually computed from the feedback provided by other entities or users that have already interacted with the given entity. In this way, when a user wants to interact with an unknown entity, the user can be informed about the behavior of the entity, deciding whether continuing or not with the interaction.

Even though both identity management and reputation management have been positively developed, merging both worlds is not straightforward in distributed environments, since it raises a number of challenges that have to be taken into consideration. For instance, some solutions have been proposed in order to endow the system defined by the OpenID standard with reputation management mechanisms. Nevertheless, those solutions are based on establishing trustworthy and static centralized services, actually not fitting in distributed environments which, indeed, OpenID is aimed to.

In distributed environments, it is not only that trustworthy centralized entities are not applicable, but also that any user can deploy several entities. If appropriate measures are not undertaken, collaborative attacks might be introduced, where a vast amount of nodes could be deployed with malicious purposes, such as unfairly increasing or decreasing the reputation of other entities.

Moreover, system conditions could be highly volatile in terms of amount of users, amount

of deployed nodes, their participation, amount of malicious users, etc. Such scenario requires the reputation systems to be highly dynamic, in a way that they should be able to adapt to the changes of the environments where they are deployed.

Additionally, in order to make the reputation management systems work, the users have to supply recommendations about the services they have been using. In this way, those recommendations are aggregated by using different mechanisms, which could even provide customized reputation values taking into account the users' preferences.

However, that would imply the service in charge of aggregating those recommendations to know the recommendations supplied by each user, the service they have been accessing and even their preferences. That would be against the goal of identity management systems, whose aims to protect the privacy of the users.

Due to the challenges of enhancing user-centric identity management systems with reputation models in distributed environments, the goals pursued within this thesis are the following:

- Analyze the current state of the art regarding reputation management systems applicable to identity management systems scenarios.
- Design and suggest solutions allowing the integration of reputation management systems with current user-centric identity management systems, in such a way that the users are properly informed before using the offered services, and the trust is properly managed in distributed environments.
- Perform a deep analysis of the behavior of such solutions, making use of different mechanism to aggregate recommendations, and considering malicious users and entities, too.
- Propose and analyze solutions aimed to enhance the adaptability of the current reputation management systems in dynamic environments.
- Explore and propose solutions to improve the privacy of current reputation management systems within the context of user-centric identity management systems.

II Methodology

After analyzing the shortcomings of trust management within the context of user-centric identity management systems in distributed environments, it seemed reasonable to think that such shortcomings could be addressed by reputation management systems.

This PhD dissertation sets its starting point at analyzing the state of the art on reputation management systems, within the context of identity management systems, and taking the intrinsic characteristics of distributed environments as a reference. The idea was to improve identity management systems in order to properly inform the users before enjoying the services they want to consume, including environments where trust cannot be handled with static agreements.

Nevertheless, we soon realized that the integration between both fields was not straightforward, and numerous considerations have to be taken into account, in order to achieve a proper behavior. Most of the existing work in that direction proposes mechanism to manage the reputation by using centralized services, being hardly applicable in more dynamic contexts.

Thus, we decided to design and analyze a reputation management model (Chapter 1) applied to a user-centric identity management system, distributed and currently widely spread, which is the case of OpenID. To the best of our knowledge, this was one of the first solutions proposing a reputation management system applied to such a standard.

As part of the analysis of the aforementioned solution, we proposed a number of mechanisms to aggregate the collected users' recommendations, differing in its complexity, requirements to

make them work, and capabilities to prevent ill-intentioned users. In this way, we did not only analyze whether the proposed solution is feasible, but also different ways of obtaining the reputation values, whose behavior depends on the system conditions (e.g. number of users, users' participation, percentage of malicious users, etc.) and the expected performance measurements (e.g. computational resources, network resources, accuracy in the reputation values, etc.).

The results of such analysis were positive in order to demonstrate the feasibility of the model. However, the analysis of the distinct mechanisms used to aggregate users' recommendations made us realize the variability of the results depending on the chosen mechanism. That would imply difficulties on choosing the aggregation mechanism to apply, since there is not a model that works in an optimal manner under all circumstances.

Furthermore, in highly dynamic environments, where the system conditions tend to change frequently, swapping the aggregation mechanism could be commonly required, which is not an easy task in current reputation management systems. After an analysis of the state of the art regarding solutions trying to address that issue, we found that some of the current works propose configurable or tunable reputation management models, at most. This solution may adjust some of the parameters defining the internal behavior of the given reputation model, but they do not provide enough flexibility to be applicable in dynamic environments.

In order to address that problem, we came up with a solution able to select and activate the most appropriate reputation management model on-the-fly, depending on the system conditions and desired performance measurements (Chapter 2). Hence, a number of reputation management models are available, although only one of them remains active computing reputation values.

If the system detects that there is a model currently more appropriate to compute the reputation, which is determined from a number of pre-defined rules, the latter becomes active. Besides, the solution incorporates mechanisms to allow a smooth transition between the current active model and the one to become active, preventing inconsistencies in the bootstrapping period required by the initialized model.

At the same time, we realized the lack of reputation management systems sensitive on preserving users' privacy. Within the context of identity management, reputation management systems attempt to gather users' recommendations about the services or other users (or even recommendations representing the trust that the services have amongst each other). These recommendations are considered private information, and freely distributing those recommendations opposes the principles of identity management systems, which aim to protect users' privacy.

Therefore, after an analysis of the related work in that direction, and after studying the applicability of advanced cryptographic mechanisms, we proposed a method towards solving that shortcoming (Chapter 3). Using homomorphic encryption techniques, the proposed method allows computing recommendations provided by the users, yet preserving the privacy of those recommendations.

III Results

The results of this thesis are essentially exposed in the articles that compose it. Hence, the article entitled "Towards the integration of reputation management in OpenID" defines a reputation management method and how it can be applied to an extended version of the protocol defined by the OpenID standard.

The solution is based on the idea that the users can provide recommendations about a service, in such a way that those feedbacks can be aggregated by an OpenID Provider. The result of such aggregation can be delivered to the users willing want to use the service. In this way, the

users would be properly informed about the trust they can place in the service before actually interacting with it. Such informed decision would increase the level of satisfaction of the users about the usage of the systems based on the user-centric identity management standard defined by OpenID.

Experiments were performed to analyze the behavior of the system, and to demonstrate that the solution is feasible, even when considering malicious entity or users.

Additionally, one of the most interesting results coming from the analysis of the proposed model was to realize that there is not a method for aggregation the recommendations offering better results that the rest in any scenario, but the result rather depends on the system conditions and expected performance measurements.

That is complemented with the outcomes of the article entitled "Dynamic and Flexible Selection of a Reputation Mechanism for Heterogeneous Environments". In that document, we present a mechanism able to dynamically and automatically select the most suitable reputation model. The selection is based on a number of pre-defined rules using fuzzy sets, in order to ease the rule definition for administrators. The rules are set to represent the behavior of the reputation management models depending on the current system conditions and the expected performance measurements.

Hence, it is no longer a matter of having a configurable model, but a pool of idle ones instead, and activate the most appropriate model at each moment (chosen according to the defined rules), without needing to stop or reconfigure the system. Moreover, the model swapping is smooth, in order to prevent inconsistencies caused by the recently activated model, as it would need a bootstrapping period until initializing properly.

A set of experiments was performed in order to validate the proposal. Thereof, it was proven that the provided reputation values are more accurate than those provided by traditional reputation management models, where a unique mechanism to compute the reputation is defined. Furthermore, the importance of performing the previously commented smooth transition was also analyzed.

Finally, in the article entitled "Identity Management: In Privacy We Trust. Bridging the Trust Gap in eHealth Environments", we address the lack of privacy that reputation system described in the current literature present. Recommendations provided by the users, or the opinion that the entities have among each other, could be considered as private information in the context of identity management.

In that document, we propose a solution aimed to have the advantages of sharing that information in order to feed the reputation management systems, while keeping it private. By using homomorphic encryption, that solution shows how users' recommendations and other required parameters to perform the reputation aggregation can be computed, but without revealing those values to potential attackers.

IV Conclusions and Future Work

User-centric identity management systems are of paramount importance to provide authentication preserving the privacy of the users, while enhancing interoperability between multiple domains. Those systems are designed as a solution to the Single Sign-On process, providing methods to share users' information between different entities.

By establishing trust relationships between different providers, the users are able to access different services making use of a unique identity, as a trustworthy entity is in charge of preserving their privacy. Nonetheless, identity management systems have shortcomings related to the trust management in distributed environments, where establishing static agreements is no longer an option.

Reputation management systems have been applied in the last years as an alternative to handle trust in this kind of environments. Trust and reputation models attempt to gather recommendations from different sources in order to predict the behavior of a given entity. However, the integration between reputation management systems and identity management systems is not straightforward, and several considerations and challenges have to be taken into account

One of the main goals of identity management systems is to protect the privacy of the users, and this may seem to conflict with the functionality of reputation management systems, which aim to collect as much amount of users' recommendations as possible. Nevertheless, as we have introduced in this thesis, researchers have already begun to find alternatives to address this kind of problem.

Moreover, we have seen that it is difficult to find a reputation management model that is suitable in any situation or appropriate for any of the scenarios that pretend to be deployed. This leads to find numerous reputation management proposals, focused on specific scenarios. This thesis aims to establish a starting point for unifying different solutions, yet there are still many interesting challenges ahead in this direction.

In this PhD thesis, a set of solutions have been proposed and analyzed in order to enhance user-centric identity management systems with trust and reputation models, taking the peculiarities of distributed environments as a reference.

Certainly, integrating reputation management models with identity management systems is a very interesting research field, although there is still much work to do. In our opinion, this thesis can be used as a reference guide for researchers willing to focus on this field.

As future work, we foresee a research line in order to propose the mechanisms, which result from the articles composing this thesis, in a standardization body. That would ease the integration between reputation management systems and identity management systems in a near future. The idea behind would be to also provide a set of best-practices, use cases and a recommendations guide to be followed by a designer of this kind of systems as reference.

An interesting research line, also resulting from this thesis, would consist on enhancing the proposed solutions in order to assist administrators on managing this kind of systems. Even though a mechanism aimed to automatically select the most appropriate reputation management model at each moment according to some defined rules has been proposed, that still would present some problems for the administrators to define such rules.

In some scenarios, where the analysis of the intrinsic properties of the reputation model would be quite laborious, the idea would be defining a mechanism able to assist the administrators in that process. Furthermore, rule definition would be complemented with artificial intelligence techniques, in such a way that the system would be able to analyze the behavior of the different reputation management models, and adapt the rules accordingly.

Resumen

I Motivación y Objetivos

En los últimos años, debido a la gran acogida de los sistemas de comunicaciones a través de Internet, sus usuarios intercambian cada vez más cantidad de información, entre la que se incluyen datos sensibles y personales. Pero estos usuarios, en ocasiones, no están informados de cómo están siendo tratados sus datos personales, o desconocen quién podrá tener realmente acceso a éstos.

De la misma manera, tener información sobre estos usuarios se considera cada vez más valioso, llegando incluso a convertirse en un objetivo para ciertas organizaciones con intereses comerciales. Esta información se utiliza especialmente con fines publicitarios, o con intención de desarrollar ataques avanzados sobre objetivos concretos en base a la información recopilada de dichos usuarios. Una gran cantidad de servicios ofrecidos a través de Internet intentan recopilar información de los usuarios, por ejemplo, a través de formularios de registro.

Es muy habitual hoy día tener que rellenar formularios de registro incluso para servicios que probablemente se utilizarán una sola vez, como por ejemplo añadir un comentario en un blog. Los datos requeridos suelen ser personales, como dirección de email o fecha de nacimiento, o incluso datos más privados que parecen innecesarios para la provisión del servicio en sí, como número de teléfono, aficiones, etc.

Esto no sólo deriva en tener que recordar distintos nombres de usuario y contraseñas para cada servicio, o exponernos a recibir correo no deseado, sino que también se pone en juego la privacidad de los usuarios. Cuando los usuarios proporcionan sus datos personales, en muchas ocasiones no conocen realmente cómo estos datos van a ser gestionados, a quién serán cedidos, o si, por ejemplo, podrán ser utilizados en campañas de marketing externas al servicio en el cual se están dando de alta.

Tanto la privacidad, así como tener el control sobre la información que las otras entidades pueden obtener de uno mismo, son características cada día más buscadas por los usuarios de cualquier sistema de comunicaciones. Además, estos temas están considerados en ciertos entornos geopolíticos como la Unión Europea, como un derecho de los usuarios. En este contexto se encuentran aquellos usuarios que no quieren relacionar su vida privada con sus interacciones en los distintos sitios web, o aquellos que no quieren que se recopile información sobre sus preferencias y perfiles de uso. Por citar algunos ejemplos, periodistas que quieren denunciar situaciones comprometedoras sin verse implicados en posibles represalias, militares que no pueden o deben revelar su situación geográfica, o simplemente como una medida más de seguridad para cualquier

usuario de los sistemas de comunicaciones a través de Internet.

Una solución elegante para estos dilemas ha comenzado a extenderse mediante la utilización de sistemas de gestión de identidades. Los sistemas de gestión de identidades establecen relaciones de confianza entre distintas organizaciones, de manera que la información de los usuarios es gestionada por una entidad confiable, conocida normalmente como proveedor de identidad (por ejemplo, su ayuntamiento, universidad o alguna organización conocida).

Cuando los usuarios acceden a un servicio en Internet (por ejemplo, una web de compra en línea, subscripción a un servicio de noticias, comentar una entrada de un blog, etc.), las funciones de autenticación y gestión de datos de los usuarios es delegada al proveedor de identidad. Así, ese proveedor de identidad se encarga de mantener la privacidad de los usuarios, al no tener éstos que registrarse y autenticarse directamente en el servicio. Adicionalmente, los sistemas de gestión de identidades proporcionan Single Sign-On, esto es, permiten a los usuarios utilizar una única cuenta (la de su proveedor de identidad) para acceder a distintos servicios repetidamente.

Se han desarrollado diversas soluciones y estándares para definir la comunicación entre los proveedores de identidad y los servicios de Internet. Como ejemplos representativos podemos considerar SAML u OpenID. Sin embargo, estos sistemas, presentan numerosas carencias aún hoy día, especialmente en lo que se refiere al control que tiene el usuario sobre su propia información. Las soluciones actuales escasamente informan a los usuarios acerca de hacia dónde viaja su información personal, y en menor medida, permiten a los usuarios decidir qué datos pueden ser o no cedidos a otras entidades.

Es más, incluso cuando los usuarios son informados de qué servicio está solicitando su información, éstos desconocen cuánto pueden confiar en dicho servicio. En otras palabras, los usuarios no saben cómo su información va a ser tratada, o si el servicio solicitado cumplirá con sus expectativas.

Estas dificultades se ven incrementadas cuando dichos sistemas de gestión de identidades han de ser desplegados en entornos distribuidos, como redes P2P, donde las relaciones de confianza no pueden ser gestionadas a través de servidores centrales estáticos. En estos casos, el establecimiento de contratos estáticos para fijar la calidad de servicio entre un proveedor de servicio y los clientes, como acuerdos de nivel de servicio (conocidos como SLA por sus siglas en inglés), son de difícil aplicación. Esto hace necesario desplegar otro tipo de mecanismos para gestionar la confianza.

Una alternativa que viene implantándose exitosamente, en los últimos años, para gestionar la confianza en entornos distribuidos viene dada por los sistemas de gestión de reputación. En estos sistemas, la confianza que se puede depositar en una entidad dada (como por ejemplo, un nodo de la red, un servicio o incluso un usuario) no viene establecida por contratos fijos, sino que está basada en experiencias pasadas que se han tenido con dicha entidad.

Los sistemas de gestión de reputación intentan predecir el comportamiento de una entidad a partir del comportamiento que ha tenido ésta en el pasado. La reputación se calcula normalmente a partir de la opinión que tienen otras entidades o usuarios que ya hayan interactuado con la entidad dada. De esta manera, cuando un usuario quiere interactuar con una entidad que no conoce, el usuario puede ser informado sobre el comportamiento que tiene dicha entidad, y éste decidirá si continuar con la interacción o no.

Aunque se ha avanzado favorablemente tanto en la gestión de identidades como en la gestión de reputación, la integración de ambos conceptos no es inmediata en entornos distribuidos, y plantea una serie de retos que han de ser tenidos en cuenta. Por ejemplo, se han propuesto soluciones para dotar al sistema definido por el estándar OpenID de mecanismos de gestión de reputación. Sin embargo, esas soluciones están basadas en el establecimiento de servicios centrales fijos y confiables, que no encajarían en entornos distribuidos, en los que OpenID están basado precisamente.

En entornos distribuidos, no sólo no se puede contar con entidades centrales de confianza, sino que también permiten a cualquier usuario desplegar entidades. Si no se toman las medidas apropiadas, esto puede introducir ataques colaborativos, en los que se despliegan una gran cantidad de nodos con objetivos maliciosos, como por ejemplo aumentar o disminuir la reputación de otras entidades malintencionadamente.

Es más, las condiciones del sistema pueden ser muy cambiantes en cuanto a la cantidad de usuarios, número de nodos desplegados, participación de los mismos, cantidad de usuarios maliciosos, etc. Esto requiere que los sistemas de gestión de reputación tengan que ser altamente dinámicos, de manera que se puedan adaptar a las variaciones del entorno.

Adicionalmente, para hacer funcionar los sistemas de gestión de reputación, los usuarios deben proporcionar recomendaciones sobre los servicios que han estado utilizando. De esta manera, las recomendaciones son agregadas utilizando diversos mecanismos, que incluso pueden producir valores de reputación personalizados a partir de las preferencias de los usuarios.

Sin embargo, eso podría implicar que el servicio encargado de agregar esas recomendaciones conociera las recomendaciones proporcionadas por cada uno de los usuarios, los servicios que han estado accediendo e incluso sus preferencias. Esto contradiría el fundamento de los sistemas de gestión de identidades, que están destinados principalmente a proteger la privacidad de los usuarios.

Dados los retos que conlleva mejorar los sistemas de gestión de identidades centrados en el usuario mediante modelos de reputación en entornos distribuidos, en esta tesis se pretende:

- Analizar el estado del arte en cuanto a los sistemas de gestión de reputación aplicables a escenarios de gestión de identidades.
- Diseñar y sugerir soluciones que permitan integrar sistemas de gestión de reputación en sistemas actuales de gestión de identidades centrados en el usuario, de manera que se informe a los usuarios apropiadamente antes de utilizar los servicios, y se gestione la confianza en entornos distribuidos.
- Realizar un profundo análisis del comportamiento de dicha soluciones, utilizando distintos mecanismos para agregar las recomendaciones, y considerando usuarios y entidades maliciosas.
- Proponer y analizar soluciones para mejorar la capacidad de adaptación de los sistemas de gestión de reputación actuales en entornos dinámicos.
- Estudiar y plantear soluciones para aumentar la privacidad de los sistemas de gestión de reputación en el ámbito de los sistemas de gestión de identidades centrados en el usuario.

II Metodología

Tras haber analizado las carencias de gestión de confianza de los sistemas de gestión de identidades centrados en el usuario en entornos distribuidos, parecía razonable pensar que estas carencias podían ser solventadas mediante sistemas de gestión de reputación.

Esta tesis doctoral tuvo como punto de partida el análisis del estado del arte de los sistemas de gestión de reputación, dentro del ámbito de los sistemas de gestión de identidades, y tomando como referencia las características intrínsecas de los entornos distribuidos. La idea era mejorar los sistemas de gestión de identidades para que estos informen a los usuarios apropiadamente antes de utilizar los servicios a los que dichos usuarios quieren acceder, incluyendo entornos donde no es posible la gestión de confianza mediante contratos estáticos.

Sin embargo, pronto nos dimos cuenta de que la integración entre ambos campos no era inmediata, y habíamos de tener en cuenta numerosas consideraciones para un correcto funcionamiento. La mayor parte de los trabajos existentes en este ámbito basaban la gestión de la reputación en servicios centralizados, siendo difícil su aplicación en entornos más dinámicos.

De este modo, decidimos diseñar y analizar un modelo de gestión de reputación (Capítulo 1), aplicado en un sistema de gestión de identidades centrado en el usuario, distribuido y ampliamente extendido actualmente, como es el caso de OpenID. Hasta donde pudimos comprobar, se trataba de una de las primeras soluciones que propusieran un sistema de reputación distribuido aplicado a dicho estándar.

Como parte del análisis de la solución anterior, investigamos varios mecanismos para agregar las recomendaciones recolectadas de los usuarios, diferenciándose en su complejidad, requerimientos para hacerlos funcionar y en la capacidad para evitar usuarios malintencionados. De esta manera, no sólo analizamos si la solución propuesta es factible, sino también distintas formas de obtener los valores de reputación, cuyo comportamiento y precisión dependían de las condiciones del sistema (número de usuarios, participación de los mismos, porcentaje de usuarios maliciosos, etc.), así como de las medidas de rendimiento esperadas (recursos computacionales, consumo de recursos de red, precisión en los valores de reputación, etc.).

Los resultados del análisis fueron positivos para demostrar la viabilidad del modelo. Sin embargo, el análisis de los distintos mecanismos para agregar recomendaciones puso de manifiesto la variabilidad de los resultados dependiendo del mecanismo escogido. Esto supondría dificultades a la hora de elegir cual es el mecanismo de agregación que debía ser utilizado en cada momento, ya que no había un modelo que funcionara de manera óptima bajo todas las circunstancias.

Es más, en entornos altamente dinámicos, donde las condiciones del sistema cambian constantemente, podría ser necesario estar reemplazando el método de agregación muy a menudo, que ciertamente no es tarea fácil en los sistemas actuales. Tras un análisis del estado del arte sobre soluciones tratando de remediar este problema, nos encontramos con que algunos trabajos actuales proponían, como mucho, modelos de gestión de reputación configurables. Estas soluciones podían calibrar algunos de sus parámetros que definían su comportamiento interno, pero esos medios no proporcionan la flexibilidad suficiente para ser aplicados en entornos dinámicos.

Para solventar ese problema, ideamos y analizamos una solución capaz de seleccionar y activar el modelo de reputación más apropiado sobre la marcha, dependiendo de las condiciones del sistema y de las medidas de rendimiento deseadas (Capítulo 2). De esta manera, se tienen varios modelos de gestión de reputación disponibles, aunque solo uno de ellos permanece activo calculando los valores de reputación.

Si el sistema detecta, a partir de una serie de reglas definidas, que hay un modelo para obtener la reputación más apropiado que aquel que actualmente se encuentra en ejecución, dicho modelo pasará a estar activo. Además, la solución incorpora mecanismos para permitir que la transición entre el modelo activo y el que pasará a estar activo se lleve a cabo de manera suave, evitando posibles inconsistencias en el período de inicialización de los modelos.

Paralelamente, nos percatamos de la deficiencia de los sistemas de gestión de reputación a la hora de preservar la privacidad de los usuarios. Los sistemas de gestión de reputación, aplicados en el ámbito de la gestión de identidades, funcionan mediante la recopilación de las opiniones que los usuarios tienen sobre los servicios o sobre otros usuarios (o incluso la confianza que tienen los servicios entre sí). Esas opiniones son consideradas como información privada, y distribuir libremente las mismas contradice uno de los principios de los sistemas de gestión de identidades, que consiste precisamente en proteger la privacidad de los usuarios.

Así, tras un análisis del estado del arte en esta dirección, y tras estudiar la aplicabilidad de mecanismos de criptografía avanzada, propusimos un método orientado a la solución de dicha

carencia (Capítulo 3). Utilizando técnicas de encriptación homomórfica, el método propuesto permite el cómputo de las recomendaciones proporcionadas por los usuarios, pero preservando la privacidad de las mismas.

III Resultados

Los resultados de esta tesis han quedado esencialmente reflejados en los artículos que la componen. Así, el artículo titulado "Towards the integration of reputation management in OpenID" define un modelo de gestión de reputación y cómo éste puede ser aplicado a una versión extendida del protocolo definido por el estándar OpenID.

Dicho modelo está basado en la idea de que los usuarios puedan proporcionar recomendaciones sobre un servicio, de manera que éstas sean agregadas por un proveedor de identidad de OpenID. El resultado de tal agregación puede ser proporcionado a los usuarios que pretendan utilizar el servicio. De esta manera, los usuarios serían informados sobre la confianza que pueden depositar en el servicio antes de utilizarlo. Esto aumentaría el nivel de satisfacción de los usuarios con el uso de los sistemas basados en el estándar de gestión de identidades centrado en el usuario definido por OpenID.

Se llevaron a cabo experimentos para comprobar el comportamiento del sistema y demostrar que la solución propuesta es viable, incluso en presencia de entidades o usuarios maliciosos. Adicionalmente, una de las conclusiones más interesantes tras el análisis del modelo propuesto fue el hecho de que no existe un método para agregar las recomendaciones que ofreciera mejores resultados en cualquier escenario, sino que los resultados dependían de las condiciones del sistema, como las medidas de rendimiento esperadas.

Estos resultados vienen complementados con los obtenidos en el artículo titulado "Dynamic and Flexible Selection of a Reputation Mechanism for Heterogeneous Environments". En ese trabajo presentamos un mecanismo capaz de seleccionar dinámica y automáticamente el modelo de reputación más apropiado. La selección está basada en reglas definidas utilizando conjuntos difusos, para facilitar la definición de estas reglas por parte de los administradores. Las reglas representan el comportamiento de los modelos de gestión de reputación a partir de las condiciones del sistema actuales y las medidas de rendimiento deseadas.

De esta manera, no se trata de desarrollar un modelo configurable, sino de tener varios modelos disponibles en espera, y activar el modelo más apropiado en cada momento (elegido según las reglas definidas), sin necesidad de parar o reconfigurar el sistema. Es más, el cambio de modelo se hace de manera suave, para evitar posibles inconsistencias causadas por los modelos recién activados, ya que estos necesitan un tiempo hasta inicializarse correctamente.

Un conjunto de experimentos se llevó a cabo para validar esta propuesta. Mediante los mismos se probó que los valores de reputación proporcionados son más precisos que aquellos obtenidos por modelos de gestión de reputación tradicionales, los cuales solo definen un mecanismo para agregar las recomendaciones. También se analizó la importancia de realizar esa transición suave que comentábamos anteriormente.

Finalmente, en el artículo titulado "Identity Management: In Privacy We Trust. Bridging the Trust Gap in eHealth Environments" abordamos el problema de la falta de privacidad que presentan los sistemas de gestión de reputación de la literatura actual. Las recomendaciones proporcionadas por los usuarios, o la opinión que tienen las entidades entre sí, podrían ser consideradas como información privada, en el ámbito de la gestión de identidades.

En ese artículo planteamos una solución que permite las ventajas que tiene compartir esa información para alimentar los sistemas de gestión de reputación, pero manteniéndola privada. Utilizando encriptación homomórfica, esta solución muestra como las recomendaciones de los

usuarios y otros parámetros necesarios para realizar el cálculo de la reputación pueden ser computados, sin necesidad de revelar dichos valores.

IV Conclusiones y Trabajo Futuro

Los sistemas de gestión de identidades centrados en el usuario son de una importancia primordial para proporcionar autenticación preservando la privacidad de los usuarios, a la par que mejoran la interoperabilidad entre múltiples dominios. Estos sistemas están diseñados como solución a los procesos de Single Sign-On, facilitando métodos para intercambiar información de los usuarios entre las distintas entidades.

A través del establecimiento de relaciones de confianza entre los distintos proveedores, los usuarios pueden acceder a diferentes servicios utilizando una sola identidad, y sería una entidad confiable quien se encargaría de preservar la privacidad de dichos usuarios. Sin embargo, los sistemas de gestión de identidades presentan carencias a la hora de gestionar la confianza en entornos distribuidos, donde el establecimiento de contratos estáticos no es una opción viable.

Los modelos de gestión de reputación han estado aplicándose en los últimos años como una alternativa para gestionar la confianza en ese tipo de entornos. Los modelos de gestión de reputación suelen recopilan recomendaciones de distintas fuentes para predecir el comportamiento de una entidad dada. Sin embargo, la integración de los sistemas de gestión de reputación con los sistemas de gestión de identidad no es inmediata, debiéndose tener en cuenta diversas consideraciones.

Los sistemas de gestión de identidad están principalmente destinados a proteger la privacidad de los usuarios, y esto puede parecer que choca con la idea de los sistemas de gestión de reputación, que tratan de recopilar la mayor cantidad de recomendaciones de los usuarios posible. Sin embargo, como hemos introducido previamente, ya se ha comenzado a trabajar para encontrar alternativas que aborden este tipo de problemas.

Por otra parte, ya hemos visto que es difícil encontrar un modelo de gestión de reputación que sea idóneo en cualquier situación o para cualquiera de los escenarios en los que pretendan ser desplegado. Esto conlleva a encontrar numerosas propuestas de modelos de gestión de reputación, enfocadas a escenarios particulares. Esta tesis pretende establecer un punto de partida para aunar las distintas soluciones, aunque aún quedan muchos retos interesantes que resolver en esta dirección.

En esta tesis se han propuesto y analizado una serie de soluciones para mejorar los sistemas de gestión de Identidades centrados en el usuario mediante modelos de reputación, tomando como referencia las peculiaridades y retos de los entornos distribuidos.

Sin duda la integración de modelos de reputación en sistemas de gestión de identidades es un campo de investigación muy interesante, aunque aún queda mucho trabajo por hacer. En nuestra opinión, esta tesis puede servir de guía de referencia para los investigadores que pretendan profundizar más en este campo.

Como trabajo futuro, podría resultar interesante una línea de investigación para proponer los mecanismos resultantes de los artículos que componen esta tesis en un cuerpo de estandarización. Esto facilitaría la integración de sistemas de gestión de reputación en sistemas de gestión de identidades en un futuro. La idea sería también proporcionar un conjunto de buenas prácticas, casos de uso y guía de recomendaciones a seguir, que pudieran tomar como referencia los diseñadores de este tipo de sistemas.

Otra línea de investigación interesante también resultante de esta tesis, consistiría en mejorar las soluciones desarrolladas para ayudar a los administradores a gestionar este tipo de sistemas. Aunque se ha propuesto un mecanismo que selecciona automáticamente el modelo de gestión

de reputación más apropiado en cada momento a partir de reglas definidas, esto aún podría suponer algún problema para los administradores a la hora de definir dichas reglas.

En ciertos escenarios, donde el análisis de las propiedades intrínsecas de cada modelo de reputación pudiera resultar muy laborioso, la idea sería definir algún mecanismo capaz de asistir a los administradores en ese proceso. Es más, la definición de reglas se podría complementar con técnicas de inteligencia artificial, de manera que el sistema fuera capaz de analizar el comportamiento de los diferentes modelos de gestión de reputación y adaptar las reglas en consecuencia.

Publications composing

the PhD Thesis

1

Towards the Integration of Reputation Management in OpenID

Title:	Towards the Integration of Reputation Management in OpenID
Authors:	Ginés Dólera Tormo, Félix Gómez Mármol, Gregorio Martínez Pérez
Type:	Journal
Journal:	Computer Standards & Interfaces, Special Issue on Secure Mobility in Future Communication Systems under Standardization
Publisher:	Elsevier
Volume:	36
Number:	3
Pages:	438-453
Year:	2014
Month:	March
DOI:	`http://dx.doi.org/10.1016/j.csi.2013.08.018`
State:	Published

Table 1: Towards the Integration of Reputation Management in OpenID

Abstract

OpenID is an open standard providing a decentralised authentication mechanism to end users. It is based on a unique URL (Uniform Resource Locator) or XRI (Extensible Resource Identifier) as identifier of the user. This fact of using a single identifier confers this approach an interesting added-value when users want to get access to different services in the Internet, since users do not need to create a new account on every website they are visiting. However, OpenID providers are normally used as a point to store certain personal attributes of the end users too, which might be of interest for any service provider willing to make profit from collecting that personal information. The definition of a reputation management solution integrated as part of the OpenID protocol can help users to determine whether a given service provider is more or less reliable before interacting with it and transferring their private information. This paper is providing the definition of a reputation framework that can be applied to the OpenID SSO (Single Sign-On) standard solution. It also defines how the protocol itself can be enhanced so OpenID providers can collect (and provide) recommendations from (to) users regarding different service providers and thus enhancing the users' experience when using OpenID. Besides the definition, a set of tests has been performed validating the feasibility of the framework.

2

Dynamic and Flexible Selection of a Reputation Mechanism for Heterogeneous Environments

Title:	Dynamic and Flexible Selection of a Reputation Mechanism for Heterogeneous Environments
Authors:	Ginés Dólera Tormo, Félix Gómez Mármol, Gregorio Martínez Pérez
Type:	Journal
Journal:	Future Generation on Computer Systems, Special Issue on Trustworthy Data Fusion and Mining in Internet of Things
Publisher:	Elsevier
Year:	2014
DOI:	http://dx.doi.org/10.1016/j.future.2014.06.006
State:	In Press, available online

Table 2: Dynamic and Flexible Selection of a Reputation Mechanism for Heterogeneous Environments

Abstract

Current trust and reputation management approaches usually offer rigid and inflexible mechanisms to compute reputation scores, which hinder their dynamic adaptation to the current circumstances in the system where they are deployed. At most, they provide certain parameters which are configurable or tunable. Yet, this is not enough for such heterogeneous and dynamic environments as the ones introduced by Internet of Things (IoT). In this paper we propose a rupture with this old philosophy and have therefore designed and prototyped a flexible mechanism to select the most suitable trust and reputation model to apply on-the-fly, amongst a pool of predefined ones, considering both the current system conditions and the selected performance measurements, which, to the best of our knowledge, is missing nowadays. Additionally, this mechanism guarantees a smooth transition between different computation engines avoiding abrupt changes in the computed reputation scores. Conducted experiments prove that our solution is able to identify and start up the most suitable trust and reputation model depending on the current system conditions (number of users, allocated resources, etc.) and expected performance measurements (accuracy, scalability, robustness, etc.).

3

Identity Management: In privacy we trust. Bridging the trust gap in e-Health environments

Title:	Identity Management: In privacy we trust. Bridging the trust gap in e-Health environments
Authors:	Ginés Dólera Tormo, Félix Gómez Mármol, Joao Girao, Gregorio Martínez Pérez
Type:	Journal
Journal:	IEEE Security & Privacy, Special Issue on Health IT Security and Privacy
Publisher:	Elsevier
Volume:	11
Number:	6
Pages:	34-41
Year:	2013
DOI:	http://dx.doi.org/10.1109/MSP.2013.80
State:	Published

Table 3: Identity Management: In privacy we trust. Bridging the trust gap in e-Health environments

Abstract

Healthcare systems are moving to decentralized scenarios, becoming a mesh of mobile services and care providers. Its rapidly growing number of users are not shy to take on new services whether in the comfort of their homes or on the move using their mobile phones and laptops. To ease use and control the data the users need to provide to individual care services, Identity Management (IdM) solutions are already being deployed. These solutions raise a number of concerns with who owns the users' data, how to control its spread and how to build trustworthy associations, between care providers. In addition to the trust and privacy issues in IdM, we present how reputation systems can enhance eHealth systems by bridging the gap between strong contractual agreements and first time domain exchanges. We further provide an example of how to preserve the privacy of feedbacks and recommendations which feed the reputation model.