Identifiers, Privacy and Trust in the Internet of Services

Leonardo A. Martucci, Sebastian Ries, and Max Mühlhäuser

Technische Universität Darmstadt, CASED
Mornewegstr. 32, DE-64293, Darmstadt, Germany
{firstname.lastname}@cased.de

Abstract. This paper presents a solution for the problem of merging privacy-friendly identifiers with trust information without support or assistance from central authorities during the operation phase. Trust information is dynamic and associated to the pseudonyms. Our solution is constructed using role-based pseudonyms that are associated to an arbitrary number of different contexts. Moreover, the presented scheme provides inherent detection and mitigation of Sybil attacks. Finally, we present an attacker model and evaluate the security and privacy properties and robustness of our solution.

1 Introduction

The Internet of the Services depicts the Internet as a conglomerate of interconnected services that interact and cooperate to fulfill tasks provided by the users. The rise of cloud computing environments is arguably the most visible face of the Internet of Services and there is a growing number of service providers that offer web-based services with a range of applications going from electronic bookstores to social networks.

In a service-oriented environment, where anybody is allowed to offer services, there will be numerous competing service providers of offering services with similar nature. Whenever a customer has the choice between services, e.g., offering books, music, or online web space, the quality of a service becomes important. The concepts of trust and reputation have been shown to be promising concepts to support the customers in such situations in selecting a high quality service [7, 8, 14]. However, building up trust and reputation usually requires long-term identifiers which can be link over numerous transactions. At a first glance, this seems to be in conflict with the protection of the users' privacy, as unlinkability is a key term when referring to privacy properties.

We address such conflicting goals by presenting a system architecture for generating role-based pseudonyms that are bound to a given set of services, called a service context. Furthermore, our solution is based on Sybil-proof pseudonyms [9, 10] and it is independent from central authorities or trusted third parties during operation phase.

The paper is organized as follows: Section 2 presents scenario and the objective of our approach. Section 3 introduces the basic building blocks. Section

4 presents the identity management scheme designed after the definition of the system requirements. The security evaluation is presented in Section 5. Finally, Section 6 concludes the paper.

2 Application Scenario and Objectives

In the first part of this section we introduce the service-oriented scenario and the notation. Then, we define the system objectives regarding privacy and trust.

2.1 Internet of Services Scenario & Notation

In an Internet of Services scenario an arbitrary number of *service providers* $s_i \in \mathcal{S} \,|\, 1 < i < z$ (with $z = |\mathcal{S}|$) offer their services to a set of users $u_i \in \mathcal{U} \,|\, 1 < i < x$ (with $x = |\mathcal{U}|$). Furthermore, we introduce the concept of *service contexts* $\mathcal{C}^i \,|\, 1 < i < y$ (with $y = |\mathcal{C}|$), and we group all services with similar nature, e.g., sellers of books or online web space providers in one *service context*[1]. Setting up service contexts is a natural consequence of an Internet of Services environment, where services competing for users are published, i.e., listed, in service directories. Within each service context $\mathcal{C}^i$, there are two sets of identifiers; the set $\mathcal{S}^i$ refers to the identifiers of the service providers available in this context, and the set $\mathcal{P}^i$ refers to the identifiers of the customers that want to use services in this context. For the customers u we propose identity management scheme, that allows each customer to create a unique pseudonym[2] $p_u^{\mathcal{C}^i}$ per service context $\mathcal{C}^i$. The relationship between the sets $\mathcal{U}$, $\mathcal{S}$, $\mathcal{C}^i$, $\mathcal{P}^i$, and $\mathcal{S}^i$ is illustrated in Fig. 1.

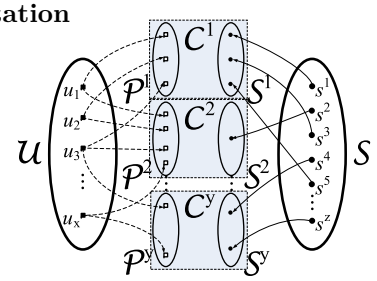

Fig. 1: Identifiers and service contexts

2.2 Objectives

The objective of our proposal is to offer a privacy-friendly identity management scheme with support to evidence-based trust or reputation systems for service environments. The requirements such a scheme are threefold:

(*i*) Providing unique, long-term identifiers as a basis for a trust or reputation model. These identifiers are needed as a basis for building histories on the quality of used services and on the behavior of others users when providing recommendations about service providers within a certain service context.

(*ii*) Providing unlinkability between the users' behaviors in different service contexts. This protects the users' privacy, in the sense that they can show different behaviors in different service contexts; but others cannot link a user's

[1] Note that the parameters x, y, and z may change over time, however, in the paper we treat them as static parameters for the simplicity of the notation.

[2] A pseudonym is an identifier of a subject other than one of subject's real names [11].

pseudonyms across service contexts (as long as the user does not create Sybil identifiers).

(*iii*) Being able to detect *Sybil identifiers* [4]. Sybil attacks are a threat to privacy and to trust models. They allow attackers to reduce the entropy of the anonymity set. Furthermore, they allow attackers to increase their influence on the trust system by misusing the Sybil entities to provide misleading recommendations by seemingly independent entities. Finally, a Sybil attack would also allow attackers to erase a bad history as they could use a newly created identifier when ever they want to. However, in the context of trust models the latter type is usually referred to as *white washing*.

3 Basic Concepts and Background

In this section, we briefly introduce the basic concepts that we propose as a basis for the trust establishment and for the identity management scheme.

3.1 Trust

In the setting introduced above trust and reputation models are important means for supporting users when selecting a service provider. For a definition of trust we refer to the definition of *reliability trust* in [8]. Trust and reputation are similar concepts and in computational models both are often based on history of past interactions. In this paper, we focus on trust as a user's subjective expectation about a service provider, and not on reputation, which is considered to be a more objective value that would be shared by all entities in a community. In the following, we present the building blocks of a (distributed) trust system.

We refer to the participants, i.e., users and service providers, in the system as *entities*. *Interactions* are actions between entities, i.e., the usage of a service or a capability that is offered by a service provider, e.g., buying goods or information. Thus, the type of interaction specifies the *service context*, in which a user wants to interact with a service provider. Whenever, an entity A is in the role of the initiator of an interaction, i.e., entity A has to select a service provider from a set of available service providers, it may evaluate the trustworthiness of the available service providers a basis for the selection. Hereby, entity A uses its direct evidence from previous interactions and recommendations (also called indirect evidence). Having collected direct evidence and recommendations about one or multiple service providers, the trust model can be used for aggregating the evidence – removing or giving lower weight to recommendations from unreliable sources – and deriving trust values for the service providers, which then can be the basis for the decision whether to interact with one of the available service providers at all, and which service provider to select.

Going along with the definition of trust, we argue for the use of probabilistic trust models, as in those models the trust value has a clear semantics, and in addition it can be used in order to judge whether it is rational to interact at all – given the possible benefits and the possible costs – based on the expected utility of

the interaction [6,12,14]. Here, we like to especially refer to Bayesian trust models as they naturally allow for the interpretation of trust as a subjective probability, which allows for the consideration of personal preferences and context-dependent parameters (for details see e.g. [2, 7, 13, 16–18])[3].

3.2 Cryptographic Tools

Different cryptographic systems can be used to create unlinkable and unique pseudonyms (see objective (ii) in Section 2.2). As long as the identification of "double-spent" pseudonyms is not an issue, such pseudonyms can be realized based on the so-called epoch number of direct anonymous attestation [1]. By binding a different tag to every identity domain, k-times anonymous authentication can be used to create unique pseudonyms.

To achieve the objectives (ii) and (iii) presented in Section 2.2, we use a cryptographic construction for creating self-certified Sybil-Free pseudonyms [9, 10]. Self-certified Sybil-free pseudonyms are obtained by a non-interactive publicly verifiable variant of a special signature scheme originally introduced for periodic n-times spendable e-tokens [3]. In our approach we use $k = 1$, so each user is represented by at maximum one pseudonym per service context. In addition, a freshly generated public key is bound to each pseudonym. Moreover, Sybil-free pseudonyms are produced through a mechanism of self-certification.

This mechanism uses different cryptographic building blocks and primitives, such as anonymous credentials and group signatures, for generating an arbitrary number of pseudonyms $p \in \mathcal{P}$ from one initial identifier $u \in \mathcal{U}$, which is obtained from a trusted third party (TTP) in the bootstrapping phase. The generation of the self-certified Sybil-free pseudonyms also produces a certificate associated with the self-certified Sybil-free pseudonym that has the following uses [9, 10]:

- to bind a freshly generated public key to the pseudonym p. This operation is similar to the binding of public keys to X.509 certificates;
- to verify the pseudonym p and its binding to the aforementioned public key;
- to disclose the initial identifier u, which is obtained from the TTP, and revocation of the certificates obtained from it, if the user that owns it creates more than one pseudonym p for a given service context $\mathcal{C}^i$.

4 Identity Management Scheme

The identity management scheme, which we propose in this paper, is the point where the Internet of Service scenario, the trust model, and the self-certified Sybil-free identifiers come together (see also Figure 1). There are four main steps in the proposed system: the bootstrapping, which is the initial step for any participant in the proposed system; the setting up of service contexts, which

[3] In the most simple version, the aggregation of direct evidence and recommendations would lead to a number of positive evidence units r and a number of negative evidence units s, and the trust value of a service provider would be calculated as $\frac{r+1}{r+s+2}$.

is usually performed by the service providers; the creation of pseudonyms for different service contexts; and the following use of such pseudonyms.

Bootstrapping: At first, we assume that each user and service provider who wants to participate in the system owns a unique, initial identifier, which is obtained at the bootstrapping phase from a party that is trusted by all involved parties (i.e., users, service directory provider, and service providers). For the service providers we assume that each provider is represented the identifier obtained in this bootstrapping phase, i.e., each service provider is represented by a single identifier across all service contexts. It is also possible to create a pseudonym for each service provider per service context in the same way as for the users.

Setting up service contexts: In principle, any party can set up service contexts. In an Internet of Services scenario, it can be carry out by the party that publishes the directories with the different services or by a set of service providers that offer services with similar nature. Setting up a service context requires a *unique identification tag* for each context. Such tags can be created from different sources, but for usability reasons the should at least provide a meaningful name for the service context, like mp3-downloads, online books, etc. A user-friendly option is to use an XML tag with context information, such as the name of the service context, region where the services are available, and validity time. The tag is then hashed into a unique value and used as input to the creation of the self-certified Sybil-free pseudonyms.

Creating user pseudonyms: The pseudonyms of the users are bound to the service contexts and are created by the users themselves. User u_i issues a pseudonym $p_{u_i}^{\mathcal{C}^j}$ that is valid in the service context $\mathcal{C}^j$ using as input her identifier originally issued by a trusted third party, a freshly generated public-private key pair, and the unique information tag associated to the service context. The pseudonym is a tuple: newly generated public key, a serial number, and a certificate that proves the correctness of the operation (for details see [9, 10]).

Using the pseudonyms: Whenever a user u_i wants to interact with a service provider in a given service context $\mathcal{C}^j$, or when she wants to evaluate the trustworthiness of a service provider in the context $\mathcal{C}^j$, she uses the pseudonym $p_{u_i}^{\mathcal{C}^j}$, which was created for this context. Thus, the real identity of the costumers is not revealed to the service providers. A service provider can only recognize whether he has already interacted with a costumer in the service context $\mathcal{C}^j$, as the customer has only a single pseudonym per context.

Expiration of pseudonyms & service contexts: A service context is valid until the validity time of the service context expires – if specified in its *unique identification tag*. When a service context expires, all pseudonyms bound to this service context become invalid. Users can also delete pseudonyms that are associate to them, but they are not able to create a new pseudonym for a service context that they were already part of – in such a case, users would need to restore the pseudonym that they had created for this service context before.

Furthermore, this pseudonym is also used for the exchange of recommendations about the behavior of the service providers between the users in the context $\mathcal{C}^j$. The differentiation between the trust relationships with regard to the dif-

ferent service contexts is as natural as important for the application scenario since a user u_1 may trust u_2 in service context $\mathcal{C}^1$, but not in service context $\mathcal{C}^2$. In our approach, both users u_1 and u_2 are identified through their (unlinkable) pseudonyms $p_{u_1}^{\mathcal{C}^1}$, $p_{u_1}^{\mathcal{C}^2}$, $p_{u_2}^{\mathcal{C}^1}$, and $p_{u_2}^{\mathcal{C}^2}$ in the service contexts $\mathcal{C}^1$ and $\mathcal{C}^2$. Thus, the users can establish trust between each other and learn who (in the sense of the owner of which pseudonym) provides accurate recommendations. Furthermore, it's also possible to sign recommendations using the private key obtained during the creation of a pseudonym.

5 Security Evaluation

We assume that an attacker tries to manipulate the trust value of a service provider or to attack the users' privacy, i.e., the attacker aims to establish relationships between pseudonyms from different pseudonym sets associated to different service contexts. Hereby, we concentrate on the attacks which have a relation to the identity management scheme described in Section 4.

The attacker model allows attackers to participate in the system and to provide both misleading or correct recommendations to other users. The attacker can eavesdrop all communications between the service context and the pseudonym. We assume that the attacker can (try to) build relationships between pseudonyms only from the pseudonyms themselves, but not from the other sources of identification, such as the network layer information, i.e., IP addresses. Thus, we assume that an anonymous communication mechanism is used the link between users and services.

5.1 Attacks on Trust Systems

When the trust value of an entity is evaluated, the main factors that are considered are direct evidence and recommendations. This leads to two basic types of attacks. On the one hand, an entity can attack the model in the role of an interactor, e.g., it starts to build trust in order to exploit it later. This type of attack should be dealt with by the trust model itself, e.g., by considering the age of evidence [7,13]. As this attack has no relation to the identification scheme, we do not further evaluate it. On the other hand, an attacker can try to influence a trust value in the role of a recommender, i.e., by providing misleading recommendations, either false praise or false accusation – again this type of attack should be dealt with in the trust model, e.g., by considering the trustworthiness of the recommenders [16]. However, both kinds of attacks are susceptible to *whitewashing* [5], i.e., the attacker repeatedly joins the community as new entity in order to get rid of a bad history, and Sybil attacks [4], i.e., an attacker creates a group of seemingly independent entities that collude.

The proposed identity management scheme prevents both types of attacks. At first, whitewashing for service providers is not possible as they have only one identifier. This attack would also be prevented if service providers would be allowed to act pseudonymously per service context using the same type of

pseudonym as the users. Whitewashing for recommenders is also not possible because a user is only allowed to have one pseudonym per context. If a user creates a second pseudonym, this can be detected given the underlying cryptographic construction, which allows the detection of multiple pseudonyms generated from a same user u_i to a given service context $\mathcal{C}^j$ by a pairwise comparison of the known pseudonyms [9, 10], which means that Sybil attacks can be detected.

5.2 Pseudonym Unlinkability

The system architecture has strong unlinkability properties as the cryptographic properties of the k-times anonymous authentication ensure the algorithmic unlinkability of two pseudonyms $p_{u_1}^{\mathcal{C}^1}$, $p_{u_1}^{\mathcal{C}^2}$ generated for $\mathcal{C}^1$ and $\mathcal{C}^2$.

However, the attacker may still be able to make an educated guess on whether two arbitrary pseudonym certificates from different identity domains are related or not, since information that may identify a device can be acquired from different sources in the TCP/IP stack, such as the network or application layers (thus, the initial assumption regarding the anonymous communication mechanism). In a real world scenario, additional information sources, like the geographical location of the user, could help the attacker to make such a guess.

6 Conclusions

In this paper, we presented an identity management scheme for a service environment enabling users to establish trust, yet, preserving a user's privacy. The trade-off between privacy, asking for anonymity at best, and trust, requiring for long-term identifiers, is set by defining that each user is able to have only one pseudonym per context. However, those identifiers cannot be linked across service contexts.

The services context that are linked to service with similar nature; they can be directly derived from a service directory. The proposed identity management scheme supports the establishment of trust within each of those service context. Furthermore, when the trust model takes recommendations from third parties into account (as in e.g., [15, 17]), then a user can also learn whether the recommendations from a certain recommender tend to be correct or misleading. The trust models especially benefits from the proposed identifier scheme as the construction of the identifier prevents Sybil attacks, which are considered to be a major threat to trust models in distributed systems.

Finally, the users privacy is preserved in the way, that a users behavior cannot be tracked over the boundaries of contexts, but it can be tracked within a contexts. This allows a service providers to created a history of its customers in a certain context, e.g., music, which is clearly preferable for the service providers, as they can tailor their advertising to the profile of the customer. However, a service provider has no means to recognize a customer in another context. Especially, we like to emphasis, that due to the construction of the identifiers, the Sybil-freeness of any set of identifiers (within a given context) can be verified without the need for an online certification authority.

References

1. E. F. Brickell, J. Camenisch, and L. Chen. Direct Anonymous Attestation. In *Proceedings of the 11th ACM Conference on Computer and Communications Security (CCS 2004)*, pages 132–145. ACM, 2004.
2. S. Buchegger and J.-Y. Le Boudec. A Robust Reputation System for Peer-to-Peer and Mobile Ad-hoc Networks. In *P2PEcon 2004*, 2004.
3. J. Camenisch, S. Hohenberger, M. Kohlweiss, A. Lysyanskaya, and M. Meyerovich. How to Win the Clone Wars: Efficient Periodic n-Times Anonymous Authentication. In *Proceedings of the 13th ACM Conference on Computer and Communications Security (CCS 2006)*, 2006.
4. J. R. Douceur. The Sybil Attack. In *Peer-to-Peer Systems: Proceedings of the 1st International Peer-to-Peer Systems Workshop (IPTPS)*, volume 2429, pages 251–260. Springer-Verlag, 7–8 Mar 2002.
5. M. Feldman and J. Chuang. Overcoming Free-riding Behavior in Peer-to-Peer Systems. *SIGecom Exch.*, 5(4):41–50, 2005.
6. A. Jøsang. Trust-based Decision Making for Electronic Transactions. In *Proceedings of the 4th Nordic Workshop on Secure IT Systems (NORDSEC'99)*, 1999.
7. A. Jøsang and R. Ismail. The Beta Reputation System. In *Proceedings of the 15th Bled Conference on Electronic Commerce*, 2002.
8. A. Jøsang, R. Ismail, and C. Boyd. A Survey of Trust and Reputation Systems for Online Service Provision. *Decision Support Systems*, 43(2):618–644, 2007.
9. L. A. Martucci. *Identity and Anonymity in Ad Hoc Networks.* PhD thesis, Karlstad University, Jun 2009.
10. L. A. Martucci, M. Kohlweiss, C. Andersson, and A. Panchenko. Self-Certified Sybil-Free Pseudonyms. In *Proceedings of the 1st ACM Conference on Wireless Network Security (WiSec'08)*, pages 154–159. ACM Press, 2008.
11. A. Pfitzmann and M. Hansen. A Terminology for Talking about Privacy by Data Minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management v0.32, 18 Dec 2009. See http://dud.inf.tu-dresden.de/literatur/.
12. D. Quercia and S. Hailes. MATE: Mobility and Adaptation with Trust and Expected-utility. *International Journal of Internet Technology and Secured Transactions (IJITST)*, 1:43–53, 2007.
13. S. Ries. Extending Bayesian Trust Models Regarding Context-dependence and User Friendly Representation. In *Proceedings of the 2009 ACM Symposium on Applied Computing.* ACM Press, 2009.
14. S. Ries. *Trust in Ubiquitous Computing.* PhD thesis, Technische Universität Darmstadt, 2009.
15. S. Ries and E. Aitenbichler. Limiting Sybil Attacks on Bayesian Trust Models in Open SOA Environments. In *Proceedings of the The 1st International Symposium on Cyber-Physical Intelligence (CPI-09)*, 2009.
16. S. Ries and A. Heinemann. Analyzing the Robustness of CertainTrust. In *2nd Joint iTrust and PST Conference on Privacy, Trust Management and Security*, pages 51–67. Springer, 2008.
17. W. T. L. Teacy, J. Patel, N. R. Jennings, and M. Luck. TRAVOS: Trust and Reputation in the Context of Inaccurate Information Sources. *Autonomous Agents and Multi-Agent Systems*, 12(2):183–198, 2006.
18. A. Whitby, A. Jøsang, and J. Indulska. Filtering out Unfair Ratings in Bayesian Reputation Systems. *ICFAIN Journal of Management Research*, 4(2):48–64, 2005.