Partial Identities as a Foundation for Trust and Reputation

Jose M. Such, Agustin Espinosa, Ana Garcia-Fornes, Vicent Botti

Departament de Sistemes Informàtics i Computació
Universitat Politècnica de València, Camí de Vera s/n, València, Spain
{jsuch,aespinos,agarcia,vbotti}@dsic.upv.es

Abstract

This paper explores the relationships between the hard security concepts of identity and privacy on the one hand, and the soft security concepts of trust and reputation on the other hand. We specifically focus on two vulnerabilities that current trust and reputation systems have: the change of identity and multiple identities problems. As a result, we provide a privacy-preserving solution to these vulnerabilities which integrates the explored relationships among identity, privacy, trust and reputation. We also provide a prototype of our solution to these vulnerabilities and an application scenario.

Keywords: Trust, Reputation, Privacy, Security

1. Introduction

Security-related studies in the Multiagent Systems (MAS) research field have been increasing over the last few years, as have intelligent autonomous agents and MAS based applications. This is mainly due to the fact that an understanding of the actual risk when using these sorts of applications is needed, since an agent's incorrect or inappropriate behavior may cause non-desired effects such as money and data loss.

Rasmusson and Jansson [1] first introduced the difference between two approaches to security in information systems, i.e., what they called hard and soft security. On the one hand, the term hard security is used for traditional security mechanisms like authentication, authorization, integrity, confidentiality, etc. On the other hand, the term soft security is used for social control mechanisms in general.

A major difference between these two approaches is related to how they deal with intruders in a system. Hard security mechanisms - such as identity management - aim to prevent intruders from joining the system so that the system is supposedly intruder free. Soft security mechanisms - such as trust and reputation - expect, and even accept, the presence of intruders in the system, so they attempt to identify the intruders and prevent them from harming the other actors in the system.

We strongly encourage research transversal to these two approaches to security. This is due to the fact that when relationships between these two approaches are not taken into account, some vulnerabilities can emerge which otherwise would not. The agent community in particular has not been taking the relationships between these two approaches into account.

Current Trust and Reputation systems are based on the assumption that identities are long-lived, so that ratings about a particular entity from the past are related to the same entity in the future. However, when such systems are actually used in real domains this assumption is no longer valid. For instance, an entity which has a low reputation due to its cheating behavior may be really interested in changing her identity and restarting her reputation from scratch. This is what Jøsang et al. [2] called the *change of*

[☆] NOTICE: this is the author's version of a work that was accepted for publication in Engineering Applications of Artificial Intelligence. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published: Jose M. Such, Agustin Espinosa, Ana Garcia-Fornes, and Vicent Botti. Partial Identities as a Foundation for Trust and Reputation. Engineering Applications of Artificial Intelligence, Vol. 24 N. 7 pp. 1128-1136 (2011). `http://www.sciencedirect.com/science/article/pii/S0952197611001084` *identities* problem. This problem has also been identified by other researchers under different names (e.g. *whitewashing* [3]).

The work of Kerr and Cohen [4] shows that Trust and Reputation Systems exhibit multiple vulnerabilities that can be exploited by attacks performed by cheating agents. Among these vulnerabilities, the *re-entry* vulnerability exactly matches the *change of identities* problem exposed by Jøsang et al. They propose a simple attack that takes advantage of this vulnerability: An agent opens an account (identity) in a marketplace, uses her account to cheat for a period, then abandons it to open another.

Kerr and Cohen [4] also point out the fact that entities could create new accounts (identity in the system) at will, not only after abandoning their previous identity but also holding multiple identities at once. This is known as the *sybil* attack [5]. An example of this attack could be an agent that holds multiple identities in a marketplace and attempts to sell the same product through each of them, increasing the probability of being chosen by a potential buyer.

It is worth mentioning that this is not an authenticity problem. Interactions among entities are assured, i.e, an agent holding an identity is sure of being able to interact with the agent that holds the other identity. However, there is nothing which could have prevented the agent behind that identity from holding another identity previously or holding multiple identities at once. For instance, let us take a buyer agent and a seller agent in an e-marketplace. The buyer has an identity in the e-marketplace under the name of *buy1* and the seller two identities in the e-marketplace *seller1* and *seller2*. Authentication in this case means that if *buy1* is interacting with *seller1* she is sure that she is interacting with who she wants. However, *buy1* has no idea that *seller1* and *seller2* are the same entity.

These vulnerabilities can be more or less harmful depending on the final domain of the application. However, these vulnerabilities should be, at least, considered in domains in which trust and reputation play a crucial role. For instance, in e-marketplaces these vulnerabilities can cause users being seriously damaged by losing money. Another example can be a social network like Last.fm[1] in which users can recommend music to each other. A user who always fails to recommend good music to other users may gain a very bad reputation. If this user creates a new account in Last.fm (a new identity in Last.fm) her reputation starts from scratch, and she is able to keep on recommending bad music. Users may be really bothered with such recommendations and move to other social networks. In this case, the one seriously damaged is the social network itself by losing users.

As far as we are concerned, the two vulnerabilities presented are partially due to the lack of a clear definition of identity and its relationship to trust and reputation. In this sense, we introduce in the next section the concept of partial identity and relate this concept to trust and reputation later on in sections 3 and 4. In section 5 we introduce what we call the *Partial Identity Unlinkability* Problem (PIUP) which is a generalization of these two vulnerabilities. As a result, a solution to PIUP is proposed in section 6, taking into consideration partial identities and their relation to trust and reputation.

As the concepts of identity, trust and reputation are based on information about entities, the privacy of the entities may be compromised. This is because, nowadays, everything is inter-connected anytime and everywhere so that users are constantly exposed to personal data collection and processing without even being aware of it [6]. For this reason, our solution to PIUP is based on privacy-enhancing identity management [7], as explained in section 6. Finally, section 7 presents an implementation of a prototype of our solution to PIUP and an application scenario, and section 8 presents some related works.

2. Identity and Partial Identities

The identity and partial identity terms are broadly used in identity management literature such as [7], [8] and [9]. However, there is a lack of clear and formal definitions of these two terms. In this section, we propose formal definitions of both identity and partial identity.

We assume that an entity can be: a legal person (a human being, a company, etc.) or a software entity (an intelligent agent, a virtual organization, etc.).

We also assume that entities are described by attributes attached to them. Attributes can describe a great range of topics [9]. For instance, entity names, biological characteristics (only for human beings), location (permanent address, geo-location at a given time), competences (diploma, skills), social charac-

[1] Last.fm `http://www.last.fm`

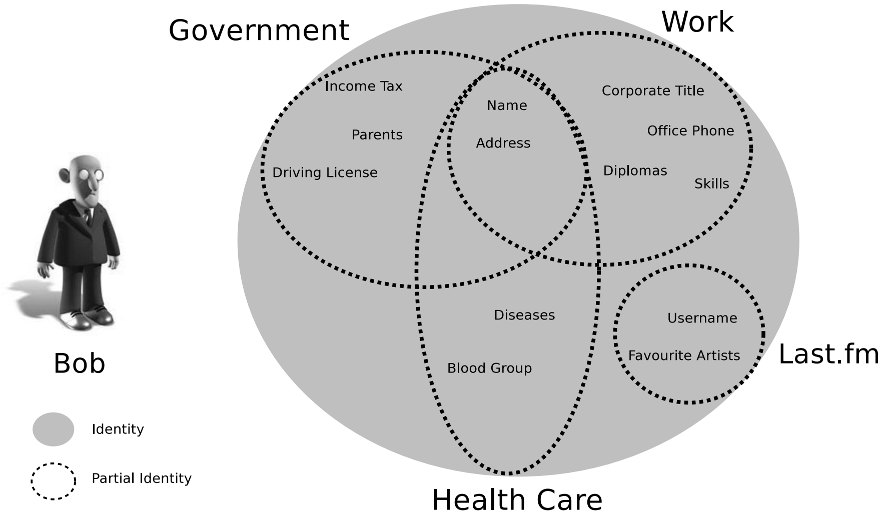

Figure 1: Identity and Partial Identities of Bob

teristics (affiliation to groups, friends), and even behaviors (personality or mood).

Definition 1. *Given a finite set of attributes* $A = \{a_1, \ldots, a_n\}$ *each one with a finite domain* $V_{a_i} = \{v_1, \ldots, v_k\}$*, a set of entities* E *and the entity* $e \in E$*, a* partial identity *of the entity* e *is a vector* $I_e = (i_1, \ldots, i_n)$*, satisfying* $i_j \in V_{a_j}$ *and* $\forall d[d \in E \setminus \{e\} \rightarrow \forall I_d(I_d \neq I_e)]$*.*

The set of attributes A, the set of values for each attribute a denoted as V_a and the set of entities E are context-dependent. Therefore, a partial identity I_e of an entity $e \in E$ sufficiently identifies (represented by the second constraint in the definition) the entity e within the set E considering A and V_a. For instance, let a human being be registered with a given profile in the Last.fm social network. This profile is a partial identity because it does sufficiently identify the human being among all of the different entities registered in Last.fm.

Although each partial identity usually identifies the entity in a specific context or role, the same partial identity can identify the entity in *different contexts*. For instance, a driver license identifies an entity in the context of operating a motorized vehicle but it also identifies an entity in the context of accessing a disco only for adults.

Definition 2. *The* identity *of an entity* e *is* $\mathcal{I}_e = \bigcup_j I_e^j$*.*

The identity $\mathcal{I}_e$ of an entity e is the union of all of the partial identities I_e^j of e. In this sense, an identity of an entity is composed of many partial identities. In order for the reader to better understand the identity and partial identity concepts, Figure 1 shows the identity and some of the partial identities of an individual person called Bob. Four partial identities are shown regarding four contexts: government, work, health care and social networking (Last.fm). For the sake of clarity, we only show some attributes that make up each of the partial identities represented. It is easily observed that the name and address of Bob are shared by three partial identities but are not used in the partial identity he uses in Last.fm.

2.1. Real Identities

We also consider an special type of partial identities: real identities. A real identity is a partial identity that sufficiently identifies an entity within the set of all of the *legal persons* — entities that can be liable for their acts in front of the law, such as human beings, companies, etc. As described later on in section 6, we use real identities for accountability concerns such as law enforcement. For this reason, real identities are restricted to only legal persons. A real identity would be for example: *Bob Andrew Miller, born in Los Angeles, CA, USA on July 7, 1975.*

Software entities (intelligent agents, virtual organizations, etc.) cannot have real identities because, up to now, they cannot be liable for their acts in front of the law. However, this may change in the future if they finally achieve some kind of legal personhood, as suggested by [10] and [11]. In this sense, they may be part of the set of all of the legal persons and will have a real identity.

3. Trusting Entities through Partial Identities

In this section, we propose partial identities as a foundation to build trust relationships. In this sense, we first introduce the concept of trust.

According to Gambetta [12], trust is "*the subjective probability by which an individual, A, expects that another individual, B, performs a given action on which its welfare depends*".

Most of the trust models proposed by the agent community are based on Gambetta's definition and treat trust as a probability. Different grounding theories are used to build these models. Although most of them are based on Game Theory (for a survey refer to [13]) there are other probabilistic approaches like [14], in which Sierra and Debenham use Information Theory.

Agent community has also developed *cognitive* models which treat trust differently. For instance, Castelfranchi and Falcone [15] define trust as "*a mental state, a complex attitude of an agent x towards another agent y about the behaviour/action relevant for the result (goal) g*".

Both probabilistic and cognitive models share that trust is established from a *trustor* (the one who trusts) to a *trustee* (the one who is trusted). Thus, we focus on trust as a directed relationship between two entities. In this sense, a primary requirement is that the trustor is able to recognize the trustee when they interact with each other.

In the real world, an individual can recognize other individuals by means of identity documents such as a passport. However, inter-personal meetings are also carried out without the needing for such documents. For instance, a trustor is able to recognize a trustee from past interactions by recognizing her face.

In the digital world there is no physical contact, all of the interactions between entities are carried out through online networks and most of them across the Internet. The increase in global connectivity increases the number of entities taking part in the digital world and also the number of interactions they carry out. In this scenario, recognizing an entity in an interaction usually means authenticating it using technologies like Kerberos[2], OpenID[3], and so on. Entities are authenticated using such technologies according to a partial identity that they hold.

We consider trust relationships to be established between two entities through some of their partial identities. Moreover, these partial identities represent part of the context where the trust relationship is established.

Partial identities are key parts in order to build trust relationships. There are attributes of a partial identity of an entity that clearly describe important features of an entity. For instance, a corporate title (such as chief executive officer) is an attribute which is part of the partial identity of an employee of a company. When this employee interacts with other entities in a business context, his corporate title is an important attribute that the rest of the entities in that context will consider valuable to trust in him.

Figure 2 shows an example of a trust relationship established between two entities through partial identities. The entity with the username antoine trusts (represented as a directed arrow) the entity with the username JohnyFM (Adam John Wilkes). This trust relationship is contextualized in Last.fm. Moreover the favorite artist of both partial identities plays a crucial role in the trust relationship. In this sense, JohnyFM has as favorite artist Arturo Sandoval and antoine has Clifford Brown as his favorite artist. Both Arturo Sandoval and Clifford Brown are trumpet players. By knowing this, antoine may consider music recommendations from JohnyFM to be relevant for him, because they like the same kind of music players.

4. Reputation through Partial Identities

In the previous section, we stated how trust relationships can be built through partial identities. In this section, we state how partial identities relate to reputation.

We understand reputation in the same way as Sabater et al. in their Repage Model [16]. In this sense, reputation is a social evaluation of a target entity attitude towards socially desirable behavior which circulates in the society (and can be agreed on or not by each one of the entities in the society).

Reputation, just like trust, is known to be context dependent [13]. For instance, a lawyer can have a great reputation defending digital criminals while having a bad reputation making cakes.

[2] Kerberos `http://web.mit.edu/Kerberos/`

[3] OpenID `http://openid.net/`

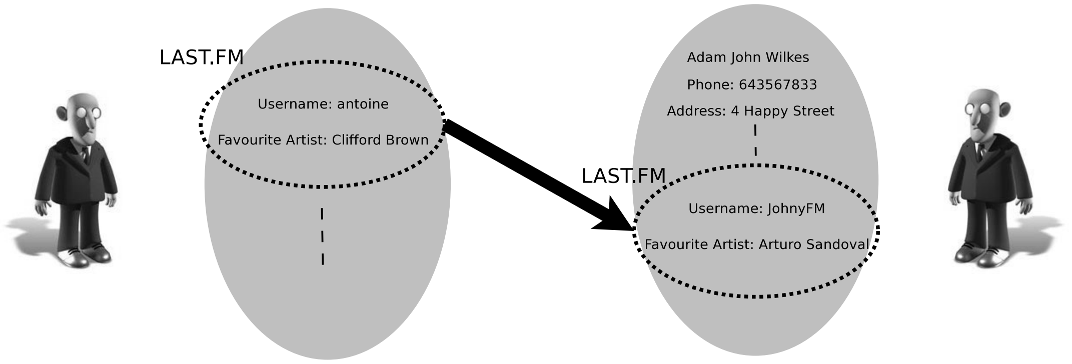

Figure 2: Trust Through Partial Identities

Unlike trust, reputation also relates to anonymity. The anonymity concept is defined by Pfitzmann and Hansen in [8] as: *"Anonymity of a subject means that the subject is not identifiable within a set of subjects"*. Reputation, as a social evaluation circulating in the society, is *anonymously* assigned to an entity. Therefore, the social evaluation any entity has about other entities remains private (whenever she does not communicate her social evaluation to others in a non-anonymous fashion).

The anonymous nature of reputation is sometimes not taken into account, which leads to some problems. For instance, the eBay reputation system is not anonymous which leads to an average 99% of positive ratings [17]. This is due to the fact that entities in eBay do not negatively rate other entities for fear of retaliations which could damage their own reputation and welfare.

We consider reputation as an anonymous social evaluation of an entity in a given context through one of its partial identities. In this sense, the partial identity of the entity reputed is needed to define the context of a reputation. Moreover, if an entity has a reputation in a given context, all of the entities interacting with this entity in the same context can be aware of her reputation through her partial identity.

5. The Partial Identity Unlinkability Problem

After the definition of the partial identity concept and its relationships to trust and reputation has been given, we are now in a position to define what we call the *partial identity unlinkability problem* (PIUP).

In section 1 we described two vulnerabilities that affect trust and reputation systems: the multiple identities and the change of identities problems. As far as we are concerned, these two vulnerabilities are closely related to the *unlinkability* concept described by Pfitzmann and Hansen in [8]. They define *unlinkability* as *" Unlinkability of two or more items of interest (IOIs, e.g., subjects, messages, actions, ...) from an attacker's perspective means that within the system (comprising these and possibly other items), the attacker cannot sufficiently distinguish whether these IOIs are related or not"*.

We use this definition of *unlinkability* made by Pfitzmann and Hansen and our definition of partial identity to formulate the PIUP:

Definition 3. *The* partial identity unlinkability *problem (PIUP) states the impossibility that an entity, which takes part in a system, is able to sufficiently distinguish whether two partial identities in that system are related or not.*

It is easily observed that the *change of identities* problem is an instantiation of PIUP, i.e., an entity with an identity by which she is known to have a bad reputation, acquires another identity with a fresh new reputation so that other entities are unable to relate the entity to its former reputation. In a similar way, if an entity does not trust another entity, the latter can change her identity. Therefore, the former entity is unable to notice that the same entity which he used to trust (distrust) is behind the new identity, so the trust relationship is restarted.

Regarding multiple identities, a similar instantiation can be made, so that an entity holds several identities and has different reputations with each of them. Thus, another entity is unable to relate the different reputations that the entity has because it is unaware of all of the identities the entity has. PIUP relates to trust in the same way when multiple identities are considered. An entity can believe that she trusts multiple entities in a given system (such as a specific marketplace), but she may be trusting the same entity with different identities without being aware of it.

5.1. The Straightforward Solution

PIUP is obviously solved by forcing the entities taking part in a system to use their real identity. Historically, a real identity has been used to uniquely identify persons [9].

If an entity is not allowed to change its identity, then trust and reputation assessments of this identity cannot be removed. Although the changing of real identities has always been possible as a way of erasing reputation, these changes are not cost-free and do not completely erase the reputation. For instance, there are some companies that change their name in order to erase their previous reputation. However, a link with the previous reputation can be made (e.g. looking at its employees in order to find employees of the former company).

Due to the impossibility of completely erasing reputation, new online services are emerging related to the management of the online reputation of an entity with a real identity. For instance, ReputationDefender[4] and Mamba IQ[5] provide services to report the online reputation of an entity with a real identity (individuals or companies). These services usually find information related to an entity searching in blogs, social networks, and audio and video pages. These services also give the entities advice on improving their online reputation.

However, the solution of forcing entities to use their real world identities exposes a great disadvantage: privacy loss. Fisher-Hübner and Hedbom in [6] define *privacy* as "*the right to informational self-determination, i.e. the right of individuals to determine for themselves when, how, to what extent and for what purposes information about them is communicated to others*".

Nowadays, in the era of global connectivity (everything is inter-connected anytime and everywhere) privacy is a great concern regarding identity management in the digital world. While in the real world everyone decides (at least implicitly) what to tell other people about themselves (after considering the situational context and the role each person plays), in the digital world users have more or less lost effective control over their personal data. Users are therefore exposed to constant personal data collection and processing without being aware of it [6].

6. A Privacy Preserving Solution for PIUP

After the definition of PIUP and the privacy issues of the straightforward solution, we provide a privacy preserving solution to PIUP so that trust and reputation systems can be used without PIUP and preserving users' privacy. Figure 3 shows our proposed architecture for trust and reputation systems. There are two layers that make up the architecture: the identity management layer and the trust and reputation model layer. The identity management layer is in charge of providing the entities taking part in a trust and reputation system with partial identity management. The trust and reputation model layer is in charge of providing the actual trust and reputation models being deployed in the system.

We assume that entities communicate to each other following a secure connection (such as TLS), so that the data they exchange in their interactions is provided with basic security features such as integrity and confidentiality.

6.1. Identity Management Layer

The technical systems supporting the process of management of partial identities are known as Identity Management Systems (IMSs) [9]. User-centric privacy-enhancing IMSs are supposed to enable a user to control the nature and amount of personal information disclosed [7]. These infrastructures are usually composed of three main parts:

- *Identity Service (IdS)* is composed of two kinds of services: Identity Providers (IdPs), that issue partial identities and validate these identities to the RPs; and Relying Parties (RPs), that are a set of APIs that allows services to check the identity of the entities that interact with them.

- The *Identity Selector (IS)* provides a simple way to manage partial identities and choose which partial identity to be used in a given context.

- *Attribute Service (AS)* include services that allow an entity to determine the access control rights of every other entity when accessing each

[4] `http://www.reputationdefender.com/`

[5] `http://www.mambaiq.com`

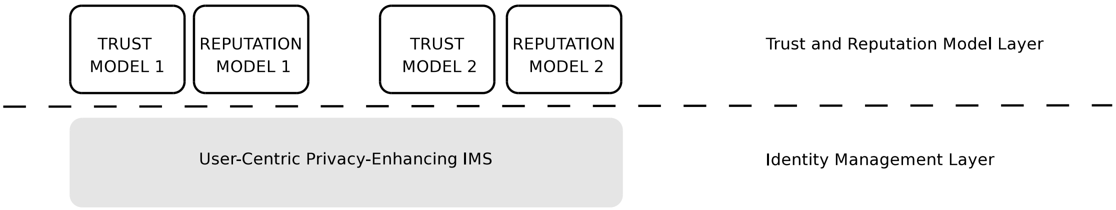
TRUST MODEL 1
REPUTATION MODEL 1
TRUST MODEL 2
REPUTATION MODEL 2
Trust and Reputation Model Layer
User-Centric Privacy-Enhancing IMS
Identity Management Layer

Figure 3: Two-layer architecture for Trust and Reputation without PIUP

> attribute of each partial identity she holds. Attributes can be managed and self-issued. Managed attributes are verified by IdPs and are reliable (and provable) information about an entity. Self-issued attributes contain information about what an entity claims about itself. IdPs can only verify that self-issued attributes are what the entities claim about themselves.

Our solution to PIUP is based on *once-in-a-lifetime* partial identities [18]. We propose that IdPs issue two kinds of partial identities: permanent partial identities (PPIs) and regular partial identities (RPIs). Entities can only hold one PPI in a system. RPIs do not pose any limitation. Although both kinds of partial identities enable trust and reputation relationships, only PPIs guarantee that PIUP is avoided. Then, entities will choose to establish trust and reputation through PPIs if they want to avoid PIUP. Our proposed identity management layer considers three main parties:

PIdP. The Permanent Identity Provider is an IdP (or a federation of IdPs[6]) that issues PPIs to the entities taking part in the specific system. Entities must register using a real identity which the PIdP will not reveal to others. The PIdP is also in charge of forcing one entity to only hold a PPI in this specific system.

IdPs. IdPs issue RPIs to the entities taking part in the specific system. Entities request RPIs providing either a real identity, or a PPI that IdPs will not reveal to others. There is no limitation in the number of IdPs per system as well as in the number of RPIs per entity and per system.

Entities. Entities, which are in a given trust and reputation system, select and manage their own partial identities using the IS. Moreover, entities also act as RPs that validate the partial identities of other entities through the PIdP and the IdPs. Entities use the AS to access attributes of other entities' partial identities. Entities also use the AS to set access control policies to their own partial identity attributes.

Figure 4 shows an example of an entity and its partial identities for a given system. The entity has the real identity with an attribute name *Adam John Wilkes*. Using this real identity the entity has obtained a PPI from the PIdP that includes two attributes: name and role. This entity has also obtained N RPIs from N different IdPs. Some of the RPIs are obtained providing its PPI (such as RPI 1) and some other using its real identity (such as RPI N).

The identity management layer provides the following main features from the point of view of security and privacy:

- *Authentication of Partial Identities.* Entities use RP APIs in order to authenticate the partial identities of the other entities taking part in the trust and reputation system. Therefore, entities are allowed to recognize to each other from interaction to interaction and establish trust and reputation relationships.

- *PIUP avoidance.* Only the PIdP is allowed to issue PPIs for a given trust and reputation system. The PIdP avoids that a previously registered entity (using a real identity) is able to obtain a new PPI. There is no chance for an entity in a trust and reputation system to have two different PPIs. Therefore, trust and reputation relationships built through PPIs avoid PIUP.

[6] IMSs support the federation of IdPs that belong to the same and also different remote security domains across the Internet. A PIdP, then, can be implemented as a federation of IdPs instead of only one IdP, minimizing the typical drawbacks of a centralized trusted third party, such as being a single point of failure (SPOF) and a possible efficiency bottleneck. Examples of identity federation standards are the Liberty Alliance Identity Federation Framework `http://projectliberty.org/resource_center/specifications/liberty_alliance_id_ff_1_2_specifications/` and WS-Federation `http://www.ibm.com/developerworks/library/specification/ws-fed/`.

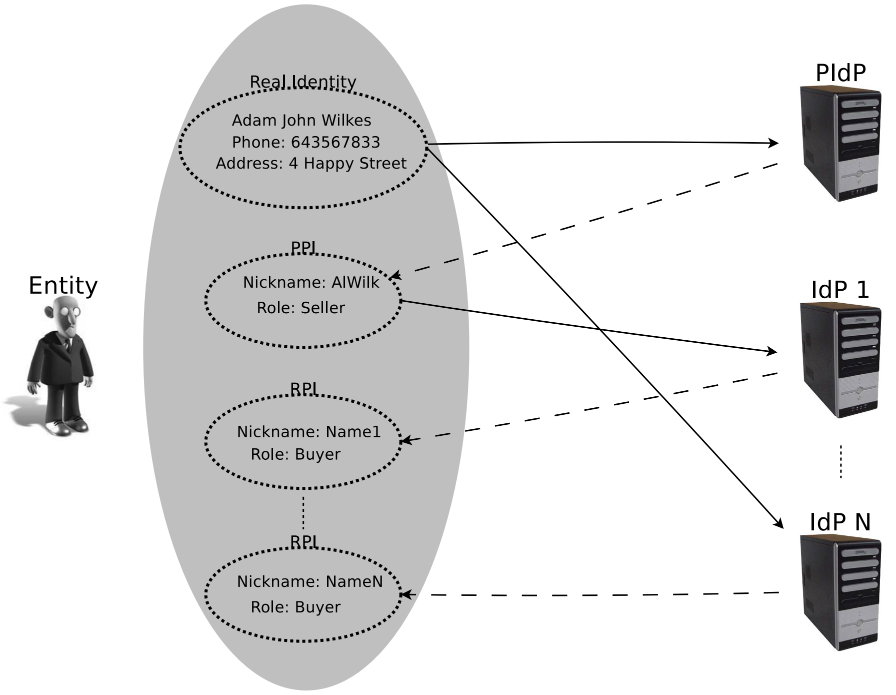

Figure 4: An example of an Entity as seen by the Identity Management Layer

- *Multiple RPIs.* Entities can hold multiple RPIs in a system. There are many situations in which entities could be interested in using multiple RPIs. For instance, multiple RPIs can play a crucial role for preserving privacy. In order to avoid buyer profiling, entities could use a different RPI for each interaction with another entity [19].

- *Hiding of original partial identities.* IdPs (including PIdP) act as independent third parties that must be trusted by the entities taking part in the trust and reputation system. For obtaining new partial identities (PPIs or RPIs), entities must provide a real identity, or a PPI to IdPs. IdPs do not make the original partial identities available. Therefore, the rest of the entities in the trust and reputation system are, a priori[7], not able to link a partial identity used in the system to the corresponding original real identity, PPI, or RPI.

- *Entity control over partial identity attributes.* ASs allow entities to determine the access control rights over each attribute of a partial identity they hold. Entities are able to choose to hide some of the attributes of a partial identity in a system as long as the resulting set of attributes is still a partial identity, i.e., it sufficiently identifies the entity among the set of entities in that system.

- *Entity accountability.* Under special circumstances, such as law enforcement, the real identity of a misbehaving entity can be known. If an entity misbehaves when using its PPI, the PIdP can disclose its real identity if required by a court. If an entity misbehaves when using one of its RPIs, IdPs can disclose the real identity or the PPI that the entity used to obtain a RPI. In case the entity used a PPI to obtain such RPI, then the PIdP can use this PPI to finally disclose the real identity of the entity. Therefore, accountability is assured and entities can be punished if necessary. This leads entities to be liable for their acts and they will take this into consideration before misbehaving.

6.2. Trust and Reputation Model Layer

On the top of the identity management layer, we find the trust and reputation model layer. This layer is the one which implements the actual trust and reputation models being used in the system.

[7] Note that if the attributes between two partial identities of the same entity are similar enough, another entity could infer that these partial identities correspond to the same entity.

Trust and reputation models in this layer are based on the definitions of identity and partial identity and their relationship with trust and reputation detailed in sections 2, 3 and 4. In this sense, partial identities act as a foundation for the establishment of trust and reputation among the entities taking part in the system.

The concept of partial identity is totally independent from the trust and reputation model being used. Therefore, a privacy preserving solution to PIUP is provided without the needing of re-designing the trust and reputation models. However, as explained in sections 3 and 4, partial identities are part of the context in which trust and reputation take place. Therefore, trust and reputation models must be aware of partial identities in order to extract the information they need to compute trust and reputation.

In this sense, partial identities can be used by trust and reputation systems for identifying an entity from interaction to interaction and building trust based on past interactions with her. For instance, Urbano et al. propose the SinAlpha [20] model for trust. This model is based on past experiences (successful or not) which are converted into a measure of trust in [0,1]. 0 means no trust and 1 means completely trust. They recognize entities from interaction to interaction by using the name of each entity. Therefore, the only adaptation needed by this model is to use partial identities as sets of only one attribute: the name of each entity.

Another example of a trust and reputation model which can be built using our two-layer architecture is Fire [21] developed by Huynh et al. This model takes into account not only past experiences but also other sources of information to assess trust and reputation. Concretely, Fire uses the role that an entity is playing in an institutional structure as a mechanism to assign default reputation to the entities. In this sense, the role of the entities can be extracted from their partial identity (whenever entities decide to make it accessible to other entities).

Finally, the trust and reputation model layer also allows heterogeneous trust and reputation systems. In this sense, there is nothing that prevents different entities from using different trust and reputation models in the same trust and reputation system. Entities are not forced to use a concrete particular trust and reputation model in a system. They could choose the trust and reputation model they prefer for a given system. Indeed, this fact opens the possibility of having multiple vendors of trust and reputation models to be used for different entities in the same system.

7. Prototype Implementation and Application Scenario

In this section we describe the implementation of a prototype for our solution to PIUP. We also provide an application scenario and its implementation using this prototype.

7.1. Prototype Implementation

We implemented one PIdP and one IdP both as webapps running on top of the Tomcat[8] web application server. Both are developed using Axis2[9]. PIdP and IdP are implemented as secure web services using the API provided by the Axis2 security module Rampart[10]. Rampart complies with the OASIS WS-Security[11] and WS-Trust[12] standards.

We considered two kinds of entities: legal persons and agents. In this way, we implemented agents as simple Java objects that interact to each other by sending and receiving messages using object method calls inside the same JVM. These agents act on behalf of legal persons. Agents also use Axis2 and Rampart APIs to call the services offered by the PIdP and the IdP. These services (following the WS-Trust standard) include the issuance, renewal, cancellation and validation of partial identities in the form of SAML2[13] security tokens. Entities can, then, use these SAML2 security tokens to prove their partial identities to other agents. Moreover, agents can choose which attributes of the attributes in a partial identity to include in each security token. Thus, they have the control over what attributes are disclosed to what other agents.

An agent calls the services of the PIdP using WS-Security with X.509 certificates to obtain a PPI. These X.509 certificates contains the real identity of the legal person that an agent is acting on behalf of. We considered the set of legal persons in Spain. Thus, the PIdP requires X.509 certificates issued by either

[8] http://tomcat.apache.org/

[9] http://ws.apache.org/axis2/

[10] http://ws.apache.org/rampart/

[11] http://www.oasis-open.org/committees/download.php/16790/wss-v1.1-spec-os-SOAPMessageSecurity.pdf

[12] http://docs.oasis-open.org/ws-sx/ws-trust/v1.4/ws-trust.html

[13] http://saml.xml.org/saml-specifications

the Spanish Electronic Identification[14] (DNIe) or the *Fábrica Nacional de Moneda y Timbre*[15] (FNMT).

The PPIs issued by the PIdP can contain attributes that an agent chooses for itself (self-issued) or can contain attributes from the real identity (managed) of the legal person the agent is acting on behalf of. The important point is that once the PIdP issues a PPI, the PIdP keeps track of what real identity holds what PPI and will always issue the same PPI to the same real identity in a given system. Thus, the PIdP avoids that an agent can have more than one PPI in a given system. PPIs can also contain attributes that the PIdP verified considering the real identities behind them. For instance, an entity can be willing to include an attribute in its PPI stating that it is over 18 years old. Then, the PIdP verifies it against the birth date in the X.509 certificate, and if it is true the PIdP includes the attribute in the PPI issued. Afterwards, the entity is able to prove to other entities that it is over 18 years old without disclosing its birth date.

Entities call the services of the IdP using WS-Security with SAML2 tokens representing its PPI to obtain a RPI. After that, the entity is able to prove that it holds the RPI to other entities. Agents can obtain as many as RPIs they desire. The IdP only keeps track of which PPI is associated to which RPIs for accountability concerns in case of law enforcement.

7.2. Application Scenario

An application Scenario for our proposed solution to PIUP is an agent-mediated e-commerce ([22], [23]) application. Agent-mediated electronic commerce refers to electronic commerce in which agent technologies are applied to provide personalized, continuously running, semi-autonomous behavior. In agent-mediated electronic commerce applications security, privacy, trust, and reputation play a crucial role [24].

We describe an electronic market where seller agents and buyer agents trade online services. In this sense, buyer agents must be able to choose among seller agents which sell the same services. One of the important dimensions that a buyer will take into account in her decision is the trust she has in each seller agent. This trust can be based on successful previous interactions with the same seller agent. A buyer agent can trust in a seller agent regarding past interactions by measuring: whether or not the seller agent provisioned the service, the overall quality of the service (QoS) bought, if there were hidden costs, etc. A buyer agent can also trust in a seller agent regarding some attributes of the seller agent's partial identity in the electronic market: registration date, corporate title, skills, etc.

Another important dimension that a buyer agent will take into account in her decision buying a service is the reputation of the seller agent. In this case, it is not what an agent thinks of a given seller agent but what it is generally said about the seller agent in the electronic market.

For the sake of simplicity, we assume that seller agents do not provide a service until they are paid. Therefore, the reputation of buyer agents and the trust other buyer and seller agents have in them are not treated. We also assume that payments are carried out using some kind of anonymous payment mechanism. Hence, the real identity of an entity is not needed when paying for a service. For instance, the untraceable electronic cash presented by Chaum et al. [25] may be used.

In this scenario the PIUP is a great concern. Seller agents should not be able to get rid of their trust and reputation assessments. This could cause important money loss. For instance, a seller agent can be cheating buyer agents by getting paid for a service which will never be delivered. This obviously decreases the trust and reputation that buyer agents have in this seller agent. Hence, this seller agent decides to quit the electronic market and re-entry into it with a new fresh identity, restarting her trust and reputation assessments from scratch. Another example would be a seller agent which sell the same service under different partial identities. In this sense, the probability that a buyer agent chooses one of their partial identities as the provider of the service increases.

We implemented one seller and three buyers. Each buyer uses its own trust and reputation machinery to model the trustworthiness of the sellers based on previous interactions and personal attributes of the sellers. The PPIs issued by the PIdP take values for two attributes: name and role. Both sellers and buyers register into the system using the PPI that the PIdP issued for them — so that the system does not know the real identity of the legal person that agents are acting on behalf of. In this way, buyers are able to identify providers from previous interactions and build their own trust and reputation models being

[14] http://www.dnielectronico.es/

[15] http://www.fnmt.es/

sure that the seller will not be able to hold any other PPI.

The seller follows a normal distribution with a mean of 0 and standard deviation of 1 to model whether it carries out the service requested in the way consumers expect it. In this sense, when a buyer requests a service to the seller, if the value returned is in the interval [-1,1], the buyer considers that the seller performed as expected. If the value returned is out of this interval the buyers consider that the seller did not perform as expected. When the seller performs as expected, buyers rate them with 1. When the seller does not perform as expected, buyers rate them with 0. These ratings are inputs of the trust and reputation model each buyer has.

Each buyer runs a different trust and reputation model that is fed using past interactions with sellers and attributes from sellers' partial identities. We implemented three models (each one for each buyer), one simply using a mean of all the previous performances to compute a trust value, one using the SinAlpha trust model that considers previous interactions, and finally, one using the Fire trust and reputation model which uses, among other information, previous interactions and the role of the entities to be trusted.

The application scenario benefits from the following features that the identity management layer provides (as stated in section 6.1):

- *Authentication of Partial Identities.* Buyers and sellers are able to authenticate their partial identities (both PPIs and RPIs). Therefore, they are allowed to recognize to each other from interaction to interaction and establish trust and reputation relationships.

- *PIUP avoidance.* There is no chance for a buyer or a seller to have two different PPIs. Therefore, trust and reputation relationships built through PPIs avoid PIUP.

- *Multiple RPIs.* Buyers can hold multiple RPIs and use a different one for each interaction with the seller. Therefore, they are able to avoid that the seller performs buyer profiling.

- *Hiding of original partial identities.* Both the PIdP and the IdP do not make the partial identities needed to obtain a PPI or a RPI available. Therefore, the rest of the agents are a priori not able to link a partial identity used to the corresponding original real identity or PPI.

- *Entity accountability.* If an agent misbehaves when using its PPI, the PIdP can disclose its real identity if required by a court. If an entity misbehaves when using one of its RPIs, IdPs can disclose the PPI that the entity used to obtain a RPI. Then the PIdP can use this PPI to finally disclose the real identity of the entity.

8. Related Work

Rehák and Pěchouček [26] relate trust and identity by modeling trust context and identity representation. They mainly focus on scenarios with scarce resources such as sensor networks, in which an underlying identity infrastructure cannot be assumed. Jennings and Finkelstein [27] propose a unified identity for social software in business processes. They propose building this unified identity by mining data from different social silos. Once this unified identity is built, it can be used as a foundation for trust and reputation. These two approaches obviate privacy concerns related to identity attributes.

Friedman and Resnick [18] propose a mechanism for preventing name changes in a social arena. They assume an intermediary, trusted by all of the entities in the specific social arena without revealing one's real identity. However, they do not consider that real identities should be revealed in special situations such as law enforcement.

Anonymity also plays a crucial role for preserving privacy [28]. Anonymity is characterized by the fact that an agent can interact with other parties in a form that these other parties do not know the identity of the agent [24]. For instance, Korba et al. [29] presents an anonymous agent communication mechanism based on the Tor network [30]. However, complete anonymity poses a great disadvantage, it does not allow trust and reputation assessments.

Warnier and Brazier [19] also present an agent communication mechanism that offers some degree of anonymity by means of what they call *handlers.* Handlers act as partial identities of only one attribute (a pseudonym) that agents can use to send messages to other agents. An agent can preserve its privacy by using a different handler for each interaction. They also consider that an agent can build up a reputation and being trusted by other parties by reusing the same handler across different interactions. However, they do not provide any protection against PIUP. Moreover, in their proposal the system on which the agents run knows the association between the users of the system and the agents that act on their behalf (usually known as semi-anonymity [19]). In our proposal, neither the agents nor the system know the real identity of the entities. The PIdP only discloses this information in case of law enforcement.

9. Conclusions

In this paper, we propose formalized definitions of partial identities and their relationship to trust and reputation. Partial identities are a key concept for identifying entities. Moreover, they play a crucial role in trust and reputation, modeling part of the context where trust and reputation take place. In this sense, both trust and reputation are established through partial identities.

We also define the *partial identity unlinkability* problem (PIUP) based on partial identities. PIUP can be more or less harmful depending on the final domain of the application using trust and reputation models. In domains where users can be seriously harmed (e.g. in an e-marketplace by losing money) PIUP needs, at least, to be considered.

We finally propose a privacy preserving solution to PIUP which takes into account privacy concerns. It allows the building of trust and reputation through partial identities while preventing entities from getting rid of trust and reputation assessments in a given system. The real identities of the entities in a system are not disclosed except under special circumstances such as law enforcement.

We implemented a prototype to validate our solution to PIUP. However, further research is needed in order to integrate our proposal into an agent platform. Such an integration will result in a complete architecture for deploying agent-based trust and reputation systems without PIUP and respecting privacy concerns. Thus, future work includes the design of this architecture, its implementation and a performance evaluation to assess the efficiency of our solution.

References

[1] L. Rasmusson, S. Jansson, Simulated social control for secure internet commerce, in: NSPW '96: Proceedings of the 1996 workshop on New security paradigms, ACM, New York, NY, USA, 1996, pp. 18–25. doi:http://doi.acm.org/10.1145/304851.304857.

[2] A. Jøsang, R. Ismail, C. Boyd, A survey of trust and reputation systems for online service provision, Decis. Support Syst. 43 (2) (2007) 618–644.

[3] E. Carrara, G. Hogben, Reputation-based systems: a security analysis, ENISA Position Paper (2007).

[4] R. Kerr, R. Cohen, Smart cheaters do prosper: defeating trust and reputation systems, in: Proc. of The 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009), 2009, pp. 993–1000.

[5] A. Jøsang, J. Golbeck, Challenges for Robust of Trust and Reputation Systems, in: Proceedings of the 5th International Workshop on Security and Trust Management (STM 2009), 2009.

[6] S. Fischer-Hübner, H. Hedbom, Benefits of privacy-enhancing identity management, Asia-Pacific Business Review 10 (4) (2008) 36–52.

[7] S. Clauβ, D. Kesdogan, T. Kölsch, Privacy enhancing identity management: protection against re-identification and profiling, in: DIM '05: Proceedings of the 2005 workshop on Digital identity management, ACM, New York, NY, USA, 2005, pp. 84–93.

[8] A. Pfitzmann, M. Hansen, Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management – a consolidated proposal for terminology, v0.31 (Feb. 2008).
URL `http://dud.inf.tu-dresden.de/Anon_Terminology.shtml`

[9] K. Rannenberg, D. Royer, A. Deuker (Eds.), The Future of Identity in the Information Society: Challenges and Opportunities, Springer Publishing Company, Incorporated, 2009.

[10] S. Chopra, L. White, Artificial agents - personhood in law and philosophy, in: Proc. of The 13th European Conference on Artificial Intelligence (ECAI 2004), 2004, pp. 635–639.

[11] T. Balke, T. Eymann, The conclusion of contracts by software agents in the eyes of the law, in: Proc. of The 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008), 2008, pp. 771–778.

[12] D. Gambetta (Ed.), Trust: Making and Breaking Cooperative Relations, Basil Blackwell, 1990.

[13] J. Sabater, C. Sierra, Review on computational trust and reputation models, Artificial Intelligence Review 24 (1) (2005) 33–60.

[14] C. Sierra, J. Debenham, An information-based model for trust, in: Proc. of the fourth int. joint conf. on Autonomous agents and multiagent systems (AAMAS 2005), ACM, New York, NY, USA, 2005, pp. 497–504.

[15] C. Castelfranchi, R. Falcone, Principles of trust for mas: Cognitive anatomy, social importance, and quantification, in: ICMAS '98: Proceedings of the 3rd International Conference on Multi Agent Systems, Washington, DC, USA, 1998, p. 72.

[16] J. Sabater-Mir, M. Paolucci, R. Conte, Repage: REPutation and imAGE among limited autonomous partners, JASSS - Journal of Artificial Societies and Social Simulation 9 (2).

[17] P. Resnick, R. Zeckhauser, Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system, in: M. R. Baye (Ed.), The Economics of the Internet and E-Commerce, Vol. 11 of Advances in Applied Microeconomics, Elsevier Science, 2002, pp. 127–157.

[18] E. J. Friedman, P. Resnick, The social cost of cheap pseudonyms, Journal of Economics and Management Strategy 10 (1998) 173–199.

[19] M. Warnier, F. Brazier, Anonymity services for multi-agent systems, Web Intelligence and Agent Systems 8 (2) (2010) 219–232.

[20] J. Urbano, A. P. Rocha, E. Oliveira, Computing confidence values: Does trust dynamics matter?, in: EPIA '09: Proceedings of the 14th Portuguese Conference on Artificial Intelligence, Springer-Verlag, 2009, pp. 520–531.

[21] T. D. Huynh, N. R. Jennings, N. R. Shadbolt, An integrated trust and reputation model for open multi-agent systems, Autonomous Agents and Multi-Agent Systems 13 (2) (2006) 119–154.

[22] C. Sierra, Agent-mediated electronic commerce, Autonomous Agents and Multi-Agent Systems 9 (3) (2004) 285–301.

[23] M. He, N. R. Jennings, H.-F. Leung, On agent-mediated electronic commerce, IEEE Transactions on Knowledge and Data Engineering 15 (4) (2003) 985–1003.

[24] M. Fasli, On agent technology for e-commerce: trust, security and legal issues, Knowledge Eng. Review 22 (1) (2007) 3–35.

[25] D. Chaum, A. Fiat, M. Naor, Untraceable electronic cash, in: CRYPTO '88: Proceedings on Advances in cryptology, Springer-Verlag New York, Inc., New York, NY, USA, 1990, pp. 319–327.

[26] M. Rehák, M. Pěchouček, Trust modeling with context representation and generalized identities, in: CIA '07: Proceedings of the 11th international workshop on Cooperative Information Agents XI, Berlin, Heidelberg, 2007, pp. 298–312.

[27] B. Jennings, A. Finkelstein, Digital identity and reputation in the context of a bounded social ecosystem., in: Business Process Management Workshops, Springer, 2008, pp. 687–697.

[28] F. Brazier, A. Oskamp, C. Prins, M. Schellekens, N. Wijngaards, Anonymity and software agents: An interdisciplinary challenge, Artificial Intelligence and Law 12 (2004) 137–157.

[29] L. Korba, R. Song, G. Yee, Anonymous communications for mobile agents, in: Proceedings of the 4th International Workshop on Mobile Agents for Telecommunication Applications, MATA '02, 2002, pp. 171–181.

[30] R. Dingledine, N. Mathewson, P. Syverson, Tor: the second-generation onion router, in: Proceedings of the 13th conference on USENIX Security Symposium - Volume 13, 2004, pp. 21–21.