V. Benjumea, J. Lopez, and J. M. Troya, "Anonymity Analysis in Credentials-based Systems: A Formal Framework", *Computer Standards & Interfaces*, vol. 30, pp. 253-261, 2008.
NICS Lab. Publications: https://www.nics.uma.es/publications

Anonymity Analysis in Credentials-based Systems: A Formal Framework

Vicente Benjumea *, Javier Lopez and Jose M. Troya
Department of Computer Science
University of Malaga, 29071 Malaga, Spain

Abstract

Anonymity has been formalized and some metrics have been defined in the scope of anonymizing communication channels. In this paper, such formalization has been extended to cope with anonymity in those scenarios where users must anonymously prove that they own certain privileges to perform remote transactions. In these types of scenarios, the authorization policy states the privileges required to perform a given remote transaction. The paper presents a framework to analyze the actual degree of anonymity reached in a given transaction and allows its comparison with an ideal anonymity degree as defined by the authorization policy, providing a tool to model, design and analyze anonymous systems in different scenarios.

Key words: anonymity metrics, anonymity degree, adequacy degree, anonymous credential systems

1. Introduction

There is a high number of applications where subjects have to make use of their privileges to perform tasks they are granted for. In most of cases, such procedures are based on the utilization of users' credentials provided and supported by authorization management systems [6]. These systems are able to manage group membership, role, clearance, or any other form of authorization. For instance, when company A needs to set roles among their employees to control the use of the different networked computer systems, or when company B needs to establish distinctions among different types of costumers or providers regarding privileges over resources (either software or hardware), an adequate authorization service becomes essential.

However, there are many situations in which a user wants to make use of her privileges without revealing her identity. Probably, the main reason for that behavior is that the Internet is becoming the largest system ever known to damage individual's privacy due to the increment in the number of remote transactions and the capability of computers to collect and cross-reference a huge amount of information. Therefore, the design of anonymous credentials and anonymity services are receiving a lot of attention [3,9].

We can not hide the importance that the use of authorization solutions have, whatever type of credential is used, in the area of forensic computing. In this sense, there are many times in which detection of an illegal behavior in electronic transactions is possible by tracing the use of the credentials in the operations that the dishonest subject has performed. For this reason it is obvious that providing complete anonymity solutions would make difficult (maybe impossible) to perform certain tasks in forensic computing. However, we still have to consider the privacy rights of the users.

We envision that in the near future it will be necessary to find solutions that provide a balance between the privacy of the user and the capacity to trace, only under certain circumstances, the operations performed if she is suspicious of any illegal activity. There is no doubt that this will be part of a difficult discussion where not only technological issues will have to be considered, but also legal, social and even psychological issues.

This paper does not intend neither to provide any technical system fulfilling the previous requirements nor to elaborate on the above discussion. On the other hand, the goal of this paper is to face the problem of discovering how much anonymity is provided by an anonymous system. The reason for that aim is that, to our understanding, and from the forensic computing point of view, it will be important to know beforehand which is the level of anonymity that

* This work has been partially supported by the projects CRISIS (TIN2006-09242) and ARES (CSD2007-00004), funded by the Spanish Ministry of Education
* Corresponding Author.
Email addresses: benjumea@lcc.uma.es (Vicente Benjumea), jlm@lcc.uma.es (Javier Lopez), troya@lcc.uma.es (Jose M. Troya).

a user can reach in her electronic transactions when using any specific anonymity system. Such type of information will be crucial when finding a trade-off between user's privacy and tracing of transactions.

A formalization of anonymity was presented in [13] focussing on analysis of anonymity in Internet communications; more precisely, in systems that anonymize clients' host IP addresses. In this paper, such formalization of anonymity has been adapted and extended to cope with scenarios where anonymity is a major concern when dealing with privileges and credentials.

Our work proposes a formal framework to measure the degree of anonymity that an authorization system reaches in different scenarios. In these scenarios, the authorization in a given transaction is ruled by a policy that states the privileges that users must own in order to be able to perform it. Additionally, it provides a mechanism to define an upper limit of the anonymity degree that a system can reach when performing a transaction depending on its authorization policy.

From a forensic point of view, the proposed framework provides mechanisms to analyze anonymous remote transactions and to realize *how anonymous* they actually are, specially in relation to how anonymous they could be, and if it is enough to guarantee anonymity. The formal framework provides a test bed for comparing different anonymous systems with respect to the degree of anonymity that they can reach under different scenarios.

The framework also allows to model the requirements imposed to the observers in a transaction, and to what extend the anonymity in a transaction could be damaged if such requirements are not fulfilled. In some anonymous credential systems there exist several entities in possession of privileged information that may collude to disclose sensitive information that allows to reverse the anonymity. These entities can also be modeled as passive observers for a transaction that can disclose such sensitive information when some conditions are met.

The paper is organized as follows. Some other works that are somewhat related are presented in section 2. Then section 3 provides a preliminary background regarding the original paper on which our work is based. Section 4 defines the formal framework for anonymity analysis in several scenarios. Section 5 describes an example that shows how the framework can be applied for analyzing the anonymity and adequacy of different anonymous systems. Finally, section 6 shows how current standards can benefit from this work and section 7 concludes the paper.

2. Related Work

Anonymity in information systems has been largely studied, and it is usually focused from two complementary points of view. On the one hand, the communication channel must be anonymized in order to guarantee the anonymity in the communication [14,11,13]. On the other hand, in those systems where the client must be authorized to carry out a transaction, the client needs to prove that owns enough privileges. Therefore, besides an anonymous communication channel, anonymity in the proofs of privileges becomes essential to protect her privacy [3,9].

The *degree of anonymity* was informally introduced in [11] to analyze anonymity in communication channels. Later, [13] further formalized these concepts. Some other metrics and measurements for anonymity have been proposed in [4,12,15] too. However, all of them are in the scope of anonymous communications. To the best of our knowledge no formal metrics have been defined in the scope of anonymity in credential systems. Our work further elaborate on [13], re-formalizing concepts and defining new ones to provide formal metrics for anonymity in these scenarios with privileges.

3. Background

A formalization of anonymity was presented in [13] focussing on anonymity analysis in Internet communications. More precisely, it formalized anonymity in those scenarios aimed at anonymizing clients' host IP addresses.

That work defined the degree of anonymity provided for some entity with respect to an observer while using a specified anonymous protocol, based on the probability, as assigned by the observer, that the entity is the initiator of a given transaction. It also defined the *overall degree of anonymity* provided by a protocol as the minimum degree of anonymity for every potential client and observer. Additionally, several intervals of degrees of anonymity were also formalized, from which *minimal anonymity* was defined. These concepts were applied to adjust some system parameters to guarantee that minimum anonymity was provided in the system and to analyze the degree of anonymity reached by several protocols focussed on anonymizing the IP address of the initiator of a transaction.

4. Anonymity analysis

In this section, we reformulate the concepts stated in the paper [13] to better adapt them to a new scenario for anonymous authorization based on privileges and properties that clients fulfill. It starts by formalizing the concepts of *set of observers* and *set of potential clients* for a remote transaction. Then, the *anonymity set*, a concept previously defined in [10], is formalized. We use this concept to reformulate the *anonymity degree*. The same concepts are formalized for a collusion of observers, and for an abstract authorization policy that rules a set of transactions. Finally, it defines the concept of *adequacy degree* that allows to compare the degree of anonymity reached in real transactions with regard to the ideal for that transaction.

4.1. *Some definitions*

The set of observers $\widehat{O}_t$ for a given remote transaction $t \in T$ consists of those entities o in the sample space U that are able to extract some information $I_o(t)$ from such transaction. Among others, usually the client as well as the server belong to that set.

$$o \in \widehat{O}_t \Leftrightarrow I_o(t) \neq \emptyset, \forall\, o \in U$$

Note that $I_o(t)$ may include the transaction client's IP address, some information about some properties and privileges that the transaction client fulfills and enjoys respectively, etc.

Let $C_t \equiv U - \{ s_t \}$ be the set of potential clients for a given remote transaction t, the sample space U and the transaction server s_t.

4.2. *Unique observer*

The *anonymity set* $AS_{t,o}$ for a given remote transaction $t \in T$ and observer $o \in \widehat{O}_t$ consists of those entities from the set of potential clients C_t that are liable to have been the client of the transaction, taking into account the information $I_o(t)$ that the observer is able to extract from it.

In other words, if $\Pr(c, I_o(t))$ is the probability that a given entity $c \in C_t$ is the client for a given transaction t according to what is specified in $I_o(t)$ for a given observer $o \in \widehat{O}_t$, then the anonymity set is composed by those entities that fulfill the following relation:

$$c \in AS_{t,o} \Leftrightarrow \Pr(c, I_o(t)) > 0\ ,\ \forall\, c \in C_t$$

such that:

$$\sum_{i \in AS_{t,o}} \Pr(i, I_o(t)) = 1$$

If $\Pr(C_t, I_o(t))$ is the probability that any member from the set of potential clients C_t fulfills the properties specified in $I_o(t)$ for a given transaction and observer, then the *cardinality* of the anonymity set is:

$$|AS_{t,o}| = |C_t| \Pr(C_t, I_o(t))$$

Note that depending on the kind of information regarding the sample space, two different approaches are possible in order to know the cardinality of the anonymity set. If enough information about individual properties is known to assign individual probabilities ($\Pr(c, I_o(t))$), then the anonymity set can be built with individuals from such information. Otherwise, if the overall probability $\Pr(C_t, I_o(t))$ is known, then the cardinality of the anonymity set can be calculated.

The *anonymity degree* $A_{t,o}$ for a given transaction and observer is the minimum of the inverse of the probability that any entity, belonging to the anonymity set for such transaction and observer, is the client of such transaction.

$$A_{t,o} = min\{\ 1 - \Pr(c, I_o(t))\ \}\ ,\ \forall\, c \in AS_{t,o}$$

In the ideal case where each member of the anonymity set has the same probability of having taken part in a given transaction as a client:

$$\Pr(c, I_o(t)) = \frac{1}{|AS_{t,o}|}\ ,\ \forall\, c \in AS_{t,o}$$

therefore, in the ideal case:

$$A_{t,o} = 1 - \frac{1}{|AS_{t,o}|}$$

In our study, which is oriented towards scenarios where authorization is based on users' privileges, we assume this ideal case of equiprobability for the members of the anonymity set.

As it was previously mentioned, the anonymity set for a given transaction depends on the probability that any entity has been the client of the transaction for a given observer. That probability depends on the information $I_o(t)$ that such observer is able to extract from the transaction. If the information can be split into minor parts, then:

$$I_o(t) = I_o^i(t) \cup I_o^j(t) \Rightarrow AS_{t,o} = AS_{t_i,o} \cap AS_{t_j,o}$$

therefore:

$$|AS_{t,o}| = |C_t| \Pr(C_t, I_o^i(t) \cap I_o^j(t))$$

that in some way points out that the larger is the extracted information, the smaller is the anonymity set for the transaction and, therefore, the smaller is the anonymity degree.

If a given observer is able to correlate two given transactions t_i and t_j as being performed by the same client c, then a new *virtual* transaction t can be defined as result of the join of both transactions:

$$t = t_i \cup t_j \Rightarrow I_o(t) = I_o(t_i) \cup I_o(t_j)$$

and therefore:

$$AS_{t,o} = AS_{t_i,o} \cap AS_{t_j,o}$$

This points out that if it is possible to correlate two o more remote transactions as being carried out by the same client, then the anonymity degree reached in both transactions is equivalent to that one reached by only one virtual transaction, where it would have been simultaneously revealed all the same information equivalent to that revealed in both correlated transactions.

4.3. *Set of observers*

The set of observers O_t for a given transaction is made up of those members in $\widehat{O}_t$ that do not belong to its own anonymity set for that transaction:

$$o \in O_t \Leftrightarrow o \notin AS_{t,o}\ ,\ \forall\, o \in \widehat{O}_t$$

Note that if $o \in AS_{t,o}$ then $\Pr(o, I_o(t)) > 0$. This implies that $\Pr(o, I_o(t)) = 1$, that is, the observer herself is the client of the transaction (except in those ones where the client is not able to identify that it has been performed by herself). In other words, O_t is $\widehat{O}_t$ excluding the client of the transaction if she belongs to such set.

Given a set of observers O_t for a transaction, the *minimal anonymity set* and the *minimal anonymity degree* are defined as:

$$AS_t^- \;=\; min\{AS_{t,o}\}\;,\;\forall\; o \;\in\; O_t$$
$$A_t^- \;=\; min\{A_{t,o}\}\;,\;\forall\; o \;\in\; O_t$$

Given a set of observers O_t for a transaction, the application of the collusion of the observers in O_t to the previous concepts can be defined as those ones in which the joined information $I_O^*(t)$ is used in the definition of the anonymity set and the anonymity degree.

$$I_O^*(t) \;=\; \bigcup_{o\in O_t} I_o(t)$$

Therefore, $\Pr(c, I_O^*(t))$ is the probability that a given entity $c \in C_t$ is the client in the transaction t according to what is specified in $I_O^*(t)$, and $\Pr(C_t, I_O^*(t))$ is the probability that any member from the set of potential clients C_t fulfills the properties specified in $I_O^*(t)$:

$$c \;\in\; AS_t^* \;\Leftrightarrow\; \Pr(c, I_O^*(t)) \;>\; 0\;,\;\forall\, c \;\in\; C_t$$

such that:

$$\sum_{i\in AS_t^*} \Pr(i, I_O^*(t)) \;=\; 1$$

and

$$|AS_t^*| \;=\; |C_t|\,\Pr(C_t, I_O^*(t))$$
$$A_t^* \;=\; min\{\;1 \;-\; \Pr(c, I_O^*(t))\;\}\;,\;\forall\, c \;\in\; AS_t^*$$

In the ideal case where all elements in AS_t^* have the same probability:

$$A_t^* \;=\; 1 \;-\; \frac{1}{|AS_t^*|}$$

It is important to take into account that this collusion of observers must respect the constraints imposed by the protocols used for its proper use, i.e., if *onion routing* requires that at least one intermediary node be honest, then the one that is able to extract the smaller amount of information from the transaction is removed from the collusion of observers. Obviously, it would also be interesting to study the system behavior in the case that some constraints are broken, and how that fact affects the anonymity. Likewise, in some anonymous credential systems, there exist some privileged entities with the capability of disclosing some sensitive information that allows to reverse the anonymity in a transaction. These privileged entities are supposed to be honest and they collude to reverse the anonymity only under some special circumstances. These entities are, a priori, outside the set of observers, however they can also play the role of passive observers, modelling the system having into account such circumstances.

4.4. *Authorization policy*

The previous study is focussed on the analysis of anonymity for a given transaction (already performed or in the process), and it is based on the probability that a given entity is the client in the transaction, which in turn depends on the information that can be extracted from such transaction by some observers. Some part of the information comes from the information revealed by the client (or her own system) in order to prove that owns enough privileges (such as credentials, signatures, IP address, etc.) to perform a given transaction.

When an entity wants to perform a given transaction usually must prove, by some specified means, that owns enough privileges. This privilege requirement is usually specified by means of an *authorization policy*.

It is possible to apply the preceding analysis to a given authorization policy, defining in this way an upper bound regarding the anonymity degree that can be reached in a remote transaction ruled by a given authorization policy. These concepts can also be defined regarding the privileges specified in a given authorization policy.

The *anonymity set* AS_p^* for a property p consists of those entities in the set of potential clients C_t that fulfill such property.

$$c \;\in\; AS_p^* \;\Leftrightarrow\; \text{fulfills}(c, p) = \text{true}\;,\;\forall\, c \;\in\; C_t$$

If $\Pr(C_t, p)$ is the probability that any member in the set of potential clients C_t fulfills the property p, then the *cardinality* of the anonymity set for such property is:

$$|AS_p^*| \;=\; |C_t|\,\Pr(C_t, p)$$

The *anonymity degree* A_p^* for a given property p is the inverse of the probability that any member in the anonymity set is selected randomly for such property.

$$A_p^* \;=\; 1 \;-\; \frac{1}{|AS_p^*|}$$

Likewise, the authorization policy usually states relationships among properties by means of boolean expressions. Therefore, the following equations may be useful in order to estimate the anonymity degree reached by a given authorization policy. If $AS_{p_i}^*$ is the set of entities that fulfill the property p_i, and $AS_{p_j}^*$ is the set of entities that fulfill the property p_j, then:

$$p \;=\; p_i \;\vee\; p_j \;\Rightarrow\; AS_p^* \;=\; AS_{p_i}^* \;\cup\; AS_{p_j}^*$$
$$p \;=\; p_i \;\wedge\; p_j \;\Rightarrow\; AS_p^* \;=\; AS_{p_i}^* \;\cap\; AS_{p_j}^*$$
$$p \;=\; \neg\, p_i \qquad \Rightarrow\; AS_p^* \;=\; C_t \;-\; AS_{p_i}^*$$

and therefore:

$$p \;=\; p_i \;\vee\; p_j \;\Rightarrow\; |AS_p^*| \;=\; |C_t|\,\Pr(C_t, p_i \;\cup\; p_j)$$
$$p \;=\; p_i \;\wedge\; p_j \;\Rightarrow\; |AS_p^*| \;=\; |C_t|\,\Pr(C_t, p_i \;\cap\; p_j)$$
$$p \;=\; \neg\, p_i \qquad \Rightarrow\; |AS_p^*| \;=\; |C_t|\,(1 \;-\; \Pr(C_t, p_i))$$

4.5. *Minimal anonymity and adequacy degree*

The *anonymity degree* for a given transaction as well as for an authorization policy is over the threshold of *minimal anonymity* when its value is greater or equal than 1/2 [13].

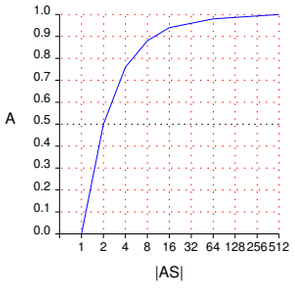

Fig. 1. Anonymity degree

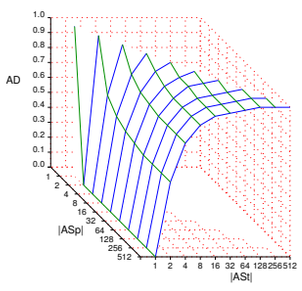

Fig. 2. Adequacy degree

$$minimal\_anonymity(A_t^*) = true \Leftrightarrow A_t^* \geq 1/2$$
$$minimal\_anonymity(A_p^*) = true \Leftrightarrow A_p^* \geq 1/2$$

The anonymity degree reached by a given authorization policy specifies the maximum degree that could be reached in any transaction ruled by that policy. Therefore, a comparison between that ideal value and the value actually reached in a given transaction, after applying the global set of observers, would provide very valuable information regarding the suitability of the technology that has been used to support the transaction with respect to the authorization policy.

The *adequacy degree* $AD_{t,p}^*$ that the anonymity degree for a given transaction has reached with respect to the anonymity degree for a given policy can be defined as ($0 \leq AD_{t,p}^* \leq 1$):

$$AD_{t,p}^* = \begin{cases} 1 & \text{if } A_t^* = A_p^* \\ \dfrac{A_t^*}{A_p^*} & \text{otherwise (note that } A_t^* < A_p^*) \end{cases}$$

Likewise, it is also possible to compare the anonymity degree reached by different authorization policies.

Figures 1 and 2 show how the anonymity degree as well as the adequacy degree rapidly grow to appropriate levels with respect to cardinality values of the corresponding anonymity sets. Figure 1 shows the anonymity degree reached with respect to the cardinality of the anonymity set for either a given transaction or a given policy. Likewise, Fig. 2 shows the adequacy degree reached depending on the cardinality of the anonymity set for a given transaction (AS_t^*) with respect to the cardinality of the anonymity set for a given policy (AS_p^*). Note that for a given transaction and policy, $|AS_t^*| \leq |AS_p^*|$. The level of minimal anonymity is pointed out by the 0.5 mark.

4.6. *Zero knowledge proofs*

When a certain entity uses *zero knowledge proofs* [7] to prove that owns some specific privileges, then the proof *verifier* does not obtain any knowledge from the *prover* except that the entity fulfills the specified property. However, that proof is suitable and completely secure. Regarding the information revealed in that proof, the verifier does not obtain any information other than the fact that the prover fulfills the specified properties. But the proof can not be correlated with anything else and, therefore, the anonymity set for a specific proof equals the ideal defined by the authorization policy.

For example, if the authorization policy for a given transaction specifies that either p_1 or p_2 are required, and the proof provided by the client of a given transaction is $ZKP(p_1 \vee p_2)$, then the anonymity set for the transaction equals the anonymity set defined by the policy, and it is composed of those entities that fulfill either one or both properties (p_1 or p_2). However, if the transaction client decides to prove only one property, as in $ZKP(p_1)$, then the proof is valid to perform the transaction, but the anonymity set for the transaction is composed of those entities that fulfill the property (p_1), which is smaller than the one defined by the policy.

Let's see another scenario. Suppose that there exists another property $p_3 \subseteq p_1$, i.e., each entity that fulfills p_3 also fulfills p_1. Then $ZKP(p_3)$ is also a valid proof to perform the transaction. However, the anonymity set for such transaction is even smaller than the one defined by the set p_1, which in turn is smaller than the one defined by the policy $p_1 \vee p_2$.

A bit of algebra about these kind of proofs:

$$ZKP(p) = ZKP(p_1 \wedge p_2) \Rightarrow AS_p^* = AS_{p_1}^* \cap AS_{p_2}^*$$
$$ZKP(p) = ZKP(p_1 \vee p_2) \Rightarrow AS_p^* = AS_{p_1}^* \cup AS_{p_2}^*$$
$$ZKP(p) = ZKP(\neg p_1) \Rightarrow AS_p^* = C_t - AS_{p_1}^*$$

It is possible that the observers are able to extract some more information in addition to that one revealed by the client privilege proof (which can be a zero knowledge proof), such as the Internet address of the client host, etc., which in turn defines an anonymity set. The anonymity set for the whole transaction is the intersection of each anonymity set defined by the information extracted from the transaction.

5. Application Example

In this section, the presented metric is applied to different scenarios with the aim of analyzing the anonymity degree as well as the adequacy degree that are reached in each of them.

Figure 3 shows a scheme with the actors in the scenario. When a client carries out a transaction with a given server (s), the messages are sent through several routing nodes (r_1, r_2, r_3) that can be either active or passive depending on the technology used. Regarding the transference of messages, the following technologies have been considered: the messages transferred between the client and the server are sent either unencrypted (in clear form) or encrypted in such a way that only the recipient entity is able to read them. Additionally, the message transference can be either direct between the sender and receiver, or through a system that hides the IP address of the sender host. In the first case, the intermediary nodes are simply passive observers able to

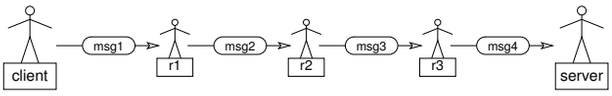

Fig. 3. Actors in the scenario

Table 1
Information Extracted from a Communication

	Clear Message		Ciphered Message	
	Direct	Onion-Routing	Direct	Onion-Routing
r_1	ip∪msg	ip	ip	ip
r_2	ip∪msg	∅	ip	∅
r_3	ip∪msg	msg	ip	∅
s	ip∪msg	msg	ip∪msg	msg

Prop	Prob	Clients
p_1	1/2	A, B, C, D, G, H, I, J
p_2	1/2	A, B, C, D, E, F, K, L
p_3	1/2	A, B, E, F, G, H, M, N
p_4	1/8	O, P
g_3	1/8	M, N
h_1	1/2	A, C, E, G, I, K, M, O
h_2	1/2	B, D, F, H, J, L, N, P

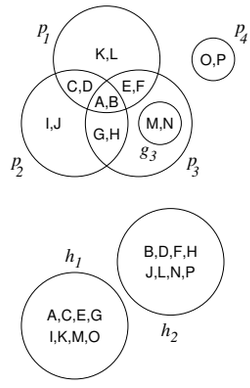

Fig. 4. Clients, Properties and IP Addresses

read the interchanged messages, and the sender IP address is exposed to every node that is involved in the communication. In the second case, the intermediary nodes are active ones whose behavior is defined by the technology used, in this latter case onion routing [14] has been considered.

Table 1 shows a summary of the information that the observers (each intermediary node as well as the server) are able to extract for a delivered message in the aforementioned scenarios. In the case of onion routing, the client owns the *application proxy* and the *onion proxy*, hence the message is already protected before sending it to the first router. If the sender does not encrypt the message for the recipient, then the last router sends the message in clear to the recipient. It is supposed that the constraints specified for each technology are fulfilled. Thus, in onion routing at least one intermediary node must be honest.

It is supposed equiprobability in the anonymity set, though in real world scenarios, anonymizers of IP-addresses are subject to several attacks that unbalance such probability having into account traffic analysis, however these analyses are outside of the scope of this paper, since the proposed metric focus on anonymity at authorization level.

In a transaction system based on privileges, usually an authorization policy specifies which requirements a client must hold in order to carry out a given transaction. In this context, the potential client must prove that owns enough privileges to perform that transaction. The figure 4 shows the clients, the properties that they fulfill, and the hosts from where they carry out the transactions for the following scenarios.

Table 2 shows the transactions that compose the example scenario, where for each transaction it shows the technology used for communications, the authorization policy and the information $I_s(t)$ received by the server when the client proved her privileges to be authorized to carry out such transaction. The table also shows the anonymity sets for each observer: the intermediary nodes $(AS_{t,r_1}, AS_{t,r_2}, AS_{t,r_3})$ and the transaction server $(AS_{t,s})$. Additionally, it also shows the anonymity set for a global observer (AS_t^*), the global anonymity degree (A_t^*), the anonymity set of the corresponding authorization policy (AS_p^*), the anonymity degree for the policy (A_p^*), and the adequacy degree for the technology used $(AD_{t,p}^*)$. In this scenario, the amount of information extracted by the server of the transaction coincides with the one extracted by the collusion of observers (according to the constraints imposed by each technology used). Therefore, the global collusion as well as the minimal anonymity set equals to the one provided by the transaction server.

$$AS_{t,s} \equiv AS_t^- \equiv AS_t^*$$
$$A_{t,s} \equiv A_t^- \equiv A_t^*$$

It follows a more detailed analysis of the transactions in table 2 with respect to the proposed metric:

– In transactions 1 to 3, the authorization policy specifies that everyone is authorized to carry out the transaction, therefore the client did not prove any privilege at all and remained anonymous. In those cases where client's IP address $(IP(h_1))$ is exposed, the anonymity set is defined by the set of entities that share the exposed IP address. Then, it can be seen that a transaction that hides the IP address reaches a higher degree of anonymity (t_3 vs. t_2). As the policy does not impose any restriction, then the anonymity set for the policy is the sample space (U). Additionally, encrypting the message improves the client privacy with respect to the intermediary nodes (t_2 vs. t_1).
– In transactions 4 to 6, the authorization policy specifies that only a given identified client (A) is allowed to perform the transaction. In this case, the identity of the client was authenticated $(ID(A))$, and therefore the degree of anonymity reached is null, since the anonymity set is only composed by a single entity. As the policy states that the client must be identified, then the used technology reaches the higher adequacy degree. However, encrypting the message and hiding the IP address improves the client privacy with respect to the intermediary nodes.
– In transaction 7, the authorization policy specifies that any client that fulfills the property p_3 is authorized to carry out the transaction. In order to prove her privileges, the identity of the client was authenticated and proved that such entity fulfills the required property $(ID(A) \wedge P(A, p_3))$, and thus the degree of anonymity reached by the client in the transaction is null, since

Table 2
Transactions

Trans	Ciph.	IP	Policy	$I_s(t)$	AS_{t,r_1}	AS_{t,r_2}	AS_{t,r_3}	$AS_{t,s}$	AS^*_t	A^*_t	AS^*_p	A^*_p	$AD^*_{t,p}$
t_1	N	Exposed	*free*	$IP(h_1)$	h_1	h_1	h_1	h_1	h_1	0.88	U	0.94	0.93
t_2	Y	Exposed	*free*	$IP(h_1)$	U	U	U	h_1	h_1	0.88	U	0.94	0.93
t_3	Y	Hidden	*free*	$\emptyset$	U	U	U	U	U	0.94	U	0.94	1
t_4	N	Exposed	$ID(c) = A$	$IP(h_1) \wedge ID(A)$	$\{A\}$	$\{A\}$	$\{A\}$	$\{A\}$	$\{A\}$	0	$\{A\}$	0	1
t_5	N	Hidden	$ID(c) = A$	$ID(A)$	U	U	$\{A\}$	$\{A\}$	$\{A\}$	0	$\{A\}$	0	1
t_6	Y	Hidden	$ID(c) = A$	$ID(A)$	U	U	U	$\{A\}$	$\{A\}$	0	$\{A\}$	0	1
t_7	Y	Hidden	p_3	$ID(A) \wedge P(A, p_3)$	U	U	U	$\{A\}$	$\{A\}$	0	p_3	0.88	0
t_8	Y	Exposed	p_3	$IP(h_1) \wedge ZKP(p_3)$	U	U	U	$h_1 \cap p_3$	$h_1 \cap p_3$	0.75	p_3	0.88	0.85
t_9	Y	Hidden	p_3	$ZKP(p_3)$	U	U	U	p_3	p_3	0.88	p_3	0.88	1
t_{10}	Y	Hidden	p_3	$ZKP(g_3)$	U	U	U	g_3	g_3	0.5	p_3	0.88	0.57
t_{11}	Y	Hidden	$p_1 \wedge p_2$	$ZKP(p_1) \wedge ZKP(p_2)$	U	U	U	$p_1 \cap p_2$	$p_1 \cap p_2$	0.75	$p_1 \cap p_2$	0.75	1
t_{12}	Y	Hidden	$p_1 \vee p_2 \vee p_3$	$ZKP(p_3)$	U	U	U	p_3	p_3	0.88	$p_1 \cup p_2 \cup p_3$	0.93	0.95
t_{13}	Y	Hidden	$p_1 \vee p_2 \vee p_3$	$ZKP(p_1 \vee p_2 \vee p_3)$	U	U	U	$p_1 \cup p_2 \cup p_3$	$p_1 \cup p_2 \cup p_3$	0.93	$p_1 \cup p_2 \cup p_3$	0.93	1
t_{14}	Y	Hidden	p_3	$LAP(x, p_3)$	U	U	U	$x : p_3$	$x : p_3$	0.88^*	p_3	0.88	1^*
t_{15}	Y	Hidden	$p_1 \wedge p_2$	$LAP(x, p_1) \wedge LAP(x, p_2)$	U	U	U	$x : p_1 \cap p_2$	$x : p_1 \cap p_2$	0.75^*	$p_1 \cap p_2$	0.75	1^*
$t_{14'}$	Y	Hidden	p_3	$I_s(t_{14}) \wedge I_s(t_{15})$	U	U	U	$x : p_1 \cap p_2 \cap p_3$	$x : p_1 \cap p_2 \cap p_3$	0.5	p_3	0.88	0.57
$t_{15'}$	Y	Hidden	$p_1 \wedge p_2$	$I_s(t_{14}) \wedge I_s(t_{15})$	U	U	U	$x : p_1 \cap p_2 \cap p_3$	$x : p_1 \cap p_2 \cap p_3$	0.5	$p_1 \cap p_2$	0.75	0.67

the anonymity set is only composed by a single entity. However, the policy defines the anonymity set for the transaction as those entities that fulfill p_3. In this case, the adequacy degree is null, as opposed to the previous transaction.

– In transactions 8 and 9, the authorization policy specifies that any client that fulfills the property p_3 is authorized to carry out the transaction. In order to prove her privileges, the client anonymously proved, by means of a *zero knowledge proof* ($ZKP(p_3)$), that fulfills the specified property. The degree of anonymity reached by the client is defined by the set of entities that fulfill p_3. In transaction t_8, the client also exposes the IP address ($IP(h_1)$), therefore the anonymity degree is decreased since in this transaction only those entities in host h_1 that fulfill p_3 compose the anonymity set. When compared with transaction t_7, it can be appreciate how these technologies provide better support for client privacy.
– In transaction 10 the client anonymously proves that fulfills the property g_3, that is a subset of p_3, and therefore such proof is valid to carry out the transaction. However, the anonymity degree reached in the transaction is less than the ideal specified in the policy. The suitability of this method depends on the difference between the superset and the subset.
– In transaction 11 the policy requires that a client must fulfill two properties, p_1 and p_2, for being authorized. Thus, the client anonymously proved, by means of two zero knowledge proofs, that fulfills both properties. Note how as the number of required properties increases, the anonymity set diminishes.
– In transactions 12 and 13 the authorization policy specifies that any client that fulfills any of the three specified properties is allowed to carry out the transaction. The client of the transaction t_{12} anonymously proved that fulfills p_3, but that fact reduces the anonymity set of the transaction with respect to the anonymity set defined by the policy. However, the client of the transaction t_{13} anonymously proved, in zero knowledge, that owns enough privileges to carry out the transaction, but did not expose indeed which privilege was actually satisfied ($ZKP(p_1 \vee p_2 \vee p_3)$), and therefore the anonymity set for transaction t_{13} is composed of those entities that fulfill any of the three properties.
– In transaction 14 the client anonymously proved, by means of a *linkable anonymous proof* ($LAP(x, p_3)$), that fulfills p_3. In this case, x represents a client pseudonym or similar that allows to link such proof with other proofs performed by the same client. The anonymity degree of the transaction is defined by the set of entities that fulfill such property, which coincides with the one defined by the policy.
– In transaction 15 the client proved, by means of two linkable anonymous proofs, that fulfills the two required properties, where x represents linking nexus.
– Transactions $14'$ and $15'$ are a redefinition of the previous ones, after realizing that both have been carried out by the same anonymous user (x), which therefore fulfills those three properties, an important fact that reduces the anonymity degree reached in these transactions. Moreover, future transactions that could be linked with these ones are also affected and contribute on this degradation of anonymity.

On the other part, this formal framework also allows to model anonymous systems where some privileged entities have some sensitive information that allows to reverse the anonymity of certain clients under some circumstances. Let us suppose that B_1 is a privileged entity that holds the binding of *Alice* with a public pseudonym $\hat{x}$, and that B_2 is also another privileged entity that holds the binding of the

public pseudonym $\hat{x}$ with the private pseudonym x. Then a collusion of both entities allow to disclose the identity of the client for the transaction t_{14} if they are able to access the information $I_s(t_{14})$ gathered from that transaction [2]. Then, under these circumstances the set of observers becomes $O^* = \{S, B_1, B_2\}$ and $I^*_O(t_{14}) = \{I_s(t_{14}) \cup [x \leftrightarrow \hat{x}] \cup [\hat{x} \leftrightarrow A]\}$. Therefore, the anonymity set $AS^*_{t_{14}}$ is composed only by A, and the anonymity degree $A^*_{t_{14}}$ becomes null, which means that the anonymity in the transaction has been reversed. Some other schemes also allow to reverse unlinkable anonymous transactions.

6. Contributions and Impacts on Standards and Interfaces

The ITU-T X.509 standard framework [8] defines the format for authentication and authorization credentials, known as public key and attribute certificates, respectively. These standard certificates convey authentication and authorization information that allow interoperability among heterogeneous systems. Moreover, the X.509 framework also defines some mechanisms to guarantee the validity and revocation status of these certificates, as well as some mechanisms to establish and validate certification paths. However, and although this technology provides clear advantages, it can also be seen as a real threat to the privacy of individuals who make use of these credentials. In fact, anonymity in authorization systems based on privileges is an emerging topic aimed at fulfilling current laws in individuals privacy protection regarding the processing and movement of personal data, such as the European Union directive 95/46/EC [5], among others.

Recently, we have proposed [2,1] approaches to incorporate anonymity into the X.509 framework but, most importantly, without modifying the standard data structures proposed by ITU-T. The direct consequence is that the framework is able to deal with new anonymity scenarios. Therefore, in this way, the standard is substantially improved in a direct way, and its contribution goes beyond that one originally intended.

In spite of that advantage, the work would be incomplete if some tools and metrics are not provided in order to support modeling and analysis given that anonymous transactions are a cornerstone in privacy enhancing technologies and the target for numerous threats. That would better complement the improvements to the ITU-T X.509 standard. Precisely, in the present work we provide a tool to help in the design of robust privacy aware systems, as well as a model for forensic analysis when privacy becomes compromised. This allows identifying anonymity weaknesses and potential breaches and, when the anonymity has to be reversed under some circumstances, identify possible breach points and design flaws.

7. Conclusions and Future Work

The formalization of anonymity in [13] has been adapted and extended to define a formal framework for analysis of anonymity in authorization systems based on users' privileges. This formal framework can be used to analyze the degree of anonymity that a remote transaction reaches, and the upper limit that could be reached by an ideal system. These measures provide a test bed for comparison of different systems with respect to the anonymity that they provide in different scenarios. Additionally, the framework makes possible to perform an analysis that allows to tune the system to reach higher degree of anonymity. It also allows to analyze the system based on the requirements that the observers must fulfill.

From a forensic point of view, the framework presented allows to analyze how anonymous a given transaction actually was, specially when compared to how anonymous the transaction could have been. It also provides a way to analyze how the technology used suits, regarding anonymity, to a given scenario. Moreover, it provides a tool that allows to model and analyze anonymous systems, which is useful when forensics becomes necessary in this kind of systems, allowing the analysis of privacy compromises, anonymity breaches and weaknesses, design flaws, etc. Also, an example of how the framework can be used to analyze anonymity and to compare several systems under different scenarios is presented.

Moreover, though the proposed metric allows different probability distributions in the anonymity set, we have used the uniform distribution to compare different anonymous authorization systems, regardless of the communication layer. However, it is interesting to use this formal framework for authorization systems in conjunction with a formal metric for anonymity analysis of the underlying communication system. This can provide some non-uniform probability distributions for the elements in the anonymity set, and further explore the capabilities that some authorization systems, in conjunction with a communication system, exhibit regarding anonymity. It would be also interesting to apply the framework presented in this paper to the analysis of widely known anonymous credential systems under several possible scenarios.

References

[1] V. Benjumea, S. G. Choi, J. Lopez, M. Yung, Anonymity 2.0: X.509 extensions supporting privacy-friendly authentication, in: CANS'07: Cryptography and Networks Security, Lecture Notes in Computer Science, Springer-Verlag, 2007.

[2] V. Benjumea, J. Lopez, J. A. Montenegro, J. M. Troya, A first approach to provide anonymity in attribute certificates, in: PKC 2004, 7th International Workshop on Practice and Theory in Public Key Cryptography, vol. 2947 of Lecture Notes in Computer Science, Springer-Verlag, 2004.

[3] J. Camenisch, A. Lysyanskaya, Efficient non-transferable anonymous multi-show credential system with optional anonymity revocation, in: EUROCRYPT 2001: Advances in

Cryptology, vol. 2045 of Lecture Notes in Computer Science, Springer-Verlag, 2001.

[4] C. Diaz, S. Seys, J. Claessens, B. Preneel, Towards measuring anonymity, in: R. Dingledine, P. Syverson (eds.), PET'02: Proc. of the Workshop on Privacy Enhancing Technologies, vol. 2482 of Lecture Notes in Computer Science, Springer-Verlag, 2002.

[5] Directive 95/46/EC of the European Parliament and of the Council, Official Journal L 281, `http://eur-lex.europa.eu/` (Oct. 1995).

[6] D. Ferraiolo, R. Sandhu, S. Gavrila, D. Kuhn, R. Chandramouli, Proposed NIST standard for Role-Based Access Control, ACM Transaction on Information and System Security 4 (3) (2001) 224–274.

[7] O. Goldreich, S. Micali, A. Wigderson, Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems, Journal of the ACM 38 (1) (1991) 691–729.

[8] ITU-T Recommendation X.509. Information Technology - Open systems interconnection - The Directory: Public-key and attribute certificate frameworks (Mar. 2000).

[9] P. Persiano, I. Visconti, An efficient and usable multi-show non-transferable anonymous credential system, in: A. Juels (ed.), FC 2004: Financial Cryptography: 8th Intl. Conf., vol. 3110 of Lecture Notes in Computer Science, Springer-Verlag, 2004.

[10] A. Pfitzmann, M. Köhntopp, Anonymity, unobservability, and pseudonymity a proposal for terminology, in: H. Federrath (ed.), Int'l Workshop Design Issues in Anonymity and Observability, vol. 2009 of Lecture Notes in Computer Science, Springer-Verlag, 2000.

[11] M. Reiter, A. Rubin, Crowds: Anonymity for Web Transactions, ACM Transactions on Information and System Security 1 (1) (1998) 66–92.

[12] A. Serjantov, G. Danezis, Towards an information theoretic metric for anonymity, in: R. Dingledine, P. Syverson (eds.), PET'02: Proc. of the Workshop on Privacy Enhancing Technologies, vol. 2482 of Lecture Notes in Computer Science, Springer-Verlag, 2002.

[13] C. Shields, B. Levine, A Protocol for Anonymous Communication over the Internet, in: Proc. 7th ACM Conference on Computer and Communication Security, 2000.

[14] P. F. Syverson, D. M. Goldschlag, M. G. Reed, Anonymous connections and onion routing, in: IEEE Symposium on Security and Privacy, Oakland, California, 1997.

[15] G. Tóth, Z. Hornák, Measuring anonymity in a non-adaptive, real-time system, in: D. Martin, A. Sejantov (eds.), PET'04: Proc. of the Workshop on Privacy Enhancing Technologies, vol. 3424 of Lecture Notes in Computer Science, Springer-Verlag, 2004.

Vicente Benjumea received his BSc and MSc from the University of Malaga, in Spain. He is currently working on his PhD in Computer Science also at the University of Malaga, focusing his research on anonymous authentication and authorization. He is involved in several national and international research projects.

Javier Lopez Javier Lopez received his M.S. and Ph.D. in Computer Science in 1992 and 2000, from the University of Malaga, respectively. From 1991 to 1994, he worked as a System Analyst and in 1994 he joined the Computer Science Department at the University of Malaga as an Assistant Professor, where he actually is an Associate Professor. He is the Co-Editor in Chief of Springer's International Journal of Information Security (IJIS), and member of the editorial board of other security-related journals. Additionally, he is Spanish representative of the IFIP TC-11 WG (Security and Protection in Information Systems).

Jose M. Troya Jose M. Troya received the MSc (1975) and PhD (1980) degrees in physics from the Universidad Complutense de Madrid. From 1980 to 1988, he was an associate professor at that university and, in 1988, became a full professor at the University of Malaga, where he leads the Software Engineering Group. His research interests include parallel programming, distributed systems, and software architectures. He is very involved in several national and international research projects, has written articles in the most relevant computer conferences and journals, supervised numerous PhD thesis, and organized several workshops and international conferences.