Pseudonymity in the light of evidence-based trust

Daniel Cvrček[1] and Václav Matyáš Jr.[2]

[1] University of Cambridge, Computer Laboratory
Daniel.Cvrcek@cl.cam.ac.uk
[2] University College Dublin, Ireland & Masaryk University Brno, Czech Rep.
matyas@informatics.muni.cz

Abstract. This position paper discusses the relation of privacy, namely pseudonymity, to evidence-based trust (or rather reputation). Critical concepts of evidence-based trust/reputation systems are outlined first, followed by an introduction to the four families of the Common Criteria (for security evaluation) Privacy Class: Unobservability, Anonymity, Unlinkability, and Pseudonymity. The paper then discusses the common problem of many papers that narrow the considerations of privacy to anonymity only, and elaborates on the concept of pseudonymity through aspects of evidence storing, attacks and some of their implications, together with other related issues like use of mixes.

1 Introduction

The reasoning introduced in this paper is based on the following general idea of reputation (or trust-based) systems. Systems do not utilize enrollment of users – there is no objective knowledge about their identities [3, 12]. All the system can use is evidence of previous behaviour. Functionally, there are three types of nodes in the system. Requesters/clients are exploiting services and resources offered by servers. Servers may use recommending service of recommenders – nodes that have an interaction history with requesters. Kinateder and Pearson [9] use slightly refined description of recommenders as they define recommender with own experiences and accumulators with mediated evidence.

The system works basically with two sets of evidence (data describing interaction outcomes). First set contains data relevant to a local system. This data is used for real-time risk analysis evaluating security of the local system in various contexts (it may be non-deterministic process in some sense). This set may be also referenced as derived or secondary data.

Primary data is in the second set. As mentioned above, the evidence is delivered (or selected from locally stored data) according to a given request content. Records from this set are used for reputation evaluation to grant/reject access requests. Data in the second set may contain information from third parties representing evidence about behaviour collected by other nodes – recommenders. Note that there may be an intersection between the two evidence sets with implications to privacy issues that we are investigating at the moment.

The approach of reputation systems is rather probabilistic and this feature directly implies properties of security mechanisms that may be defined on top of such systems. The essential problem arises with recommendations that may be artificially created by distributed types of attacks (Sibyl attack [8]) based on huge number of nodes created just to gather enough evidence and achieve maximum reputation.

A node functional model is based on two data flows. Risk analysis is based on locally stored *secondary* data and results of the analysis should adjust rules for access control. This process may be deterministic or probabilistic according to the way the search for behavioural patterns in the evidence data is conducted. Trust analysis is based on primary data about behaviour of a requester and trust values are compared against access control rules.

1.1 Trust and privacy or Trust vs. privacy?

We come then to the following question: *How much trust may someone assign to me without compromising my privacy?*

The term privacy is currently understood as the right and ability to protect one's personal space (including personal data) and to prevent invasions into this personal space. It can be then expected that increased amount of evidence necessary for higher reputation implies higher probability of privacy compromise.

A limited amount of evidence/data related to the subject can lead to limited trust level being achieved yet desirable for the subject of trust reasoning. This situation is desirable in mix-based services and it in contrary *increases trust* in the mix. The trust then reflects predictability of behaviour (size of possible recipient set) here.

What do we mean with privacy – anonymity or pseudonymity (for more definitions see below), in other words, is it worth and even possible to require anonymity from some level of mutual trust or would it be better to concentrate on pseudonymity?

2 What could privacy be?

Privacy is not only about subject (user) identity. Let us consider, for example, three questions relevant to user privacy:

- *Who* is the user, e.g. browsing xxx.com from this room?
- *Which* is the user, e.g. browsing xxx.com from this room? (I.e., we are not concerned with that user's identificator as such but with her properties.) This question could be risen by a script on xxx.com to decide whether it is possible to infect the user's laptop.
- *Is there any* user, e.g. browsing xxx.com from this room? This may be a question of a law enforcement agency before entering the room.

We can also identify three most suitable ways to "measure" privacy:

- To what degree is data personally identifiable?
- How certain can we be in linking pieces of evidence?
- To what degree is an action observable?

The Common Criteria [4] class Privacy deals with aspects of privacy in the Trusted Computing Environment and as proposed is composed of four families. Three of these families have a similar grounding with respect to entities (i.e., users or processes) whose privacy might be in danger. They describe vulnerability to varying threats, which make them distinct from each other. These families are Unobservability, Anonymity, and Unlinkability. The fourth family – Pseudonymity – addresses substantially different kind of threats.

Unobservability : *This family ensures that a user may use a resource or service without others, especially third parties, being able to observe that the resource or service is being used.* The protected asset in this case can be information about other users' communications, about access to and use of a certain resource or service, etc. Several countries, e.g. Germany, consider the assurance of communication unobservability as an essential part of the protection of constitutional rights. Threats of malicious observations (e.g., through Trojan Horses) and traffic analysis (by others than communicating parties) are best-known examples.

Anonymity : *This family ensures that a user may use a resource or service without disclosing the user's identity. The requirements for Anonymity provide protection of the user* (bound to a subject or operation) *identity. Anonymity is not intended to protect the subject identity.* Although it may be surprising to find a service of this nature in a Trusted Computing Environment, possible applications include enquiries of a confidential nature to public databases, etc. A protected asset is usually the identity of the requesting entity, but can also include information on the kind of requested operation (and/or information) and aspects such as time and mode of use. The relevant threats are: disclosure of identity or leakage of information leading to disclosure of identity – often described as "usage profiling".

Unlinkability : *This family ensures that a user may make multiple uses of resources or services without others being able to link these uses together.* The protected assets are of the same as in Anonymity. Relevant threats can also be classed as "usage profiling". Note that it is sometimes questioned whether this family deals with substantially different threats and applications from those of Anonymity.

Pseudonymity : *This family ensures that a user may use a resource or service without disclosing its user* (bound to a subject or operation) *identity, but can still be accountable for that use.* Possible applications are usage and charging for phone services without disclosing identity, "anonymous" use of an electronic payment, etc. In addition to the Anonymity services, Pseudonymity provides methods for authorization without identification (at all or directly to the resource or service provider).

3 Evidence and privacy

The crucial question to be answered here is whether it is possible to protect against identification of an individual. What is the cost difference between anonymity and pseudonymity? Can we define relations between amount of evidence and cost of preserving anonymity/pseudonymity? We are (currently) of the opinion that there are higher risks (misuse of evidence that is not personalised) when anonymity is addressed. For it is clear that even if we can reduce the set of suspects such that the set size is 1, i.e. there would be no (or lowest level of) anonymity according to current definitions and perception of anonymity, e.g. [5, 6, 9, 10], we could still have some level of pseudonymity, i.e. probability of linking a personal identity with the given pseudonym less than 1.

3.1 Evidence store

We come to one of the critical issues here – *Who should store information needed for trust evaluation*:

1. *Server* (service provider) – most convenient for the purpose of local risk analysis and adjustment of rules for access control to local resources. We do not need true client identity, although certain amount of "linkability" is desirable, for that. Basically, only the information computed locally as part of previous (e.g., access control) decisions is needed. However, this holds true only when a service provider is using this information and making the decision on its own, and it is often advisable to use more sources of evidence to increase trust (yet obviously this endangers privacy most).
2. *Client* (subject of information) – ideal from the privacy point of view because I (the subject) can decide what I really want, can and will provide with a service request – i.e., typically the required minimum. On the other hand, client usually does not have enough knowledge to guess volume of necessary evidence. This implies some tricky requirements for correct specification of communication protocols. What seems plausible at the moment is that the service request would consist of a number of request-response interactions. Another approach would be publishing the evaluation criteria for clients' behaviour. For the latter, one can find good reasons both for (transparency like in case of state benefit ratings, etc.), and against (commercial secrets, need to suppress change of behaviour reflecting knowledge of the criteria, security critical information from deployed security mechanisms, etc.).
3. *Distributed storage* – pieces of evidence are spread over a network, e.g. using the Eternity Service [2] (this is a variant of accumulators as defined in [9]). The problem is to ensure that evidence will not remain on server using it to calculate reputation value. One of possible ways may be use of distributed calculation. This however brings question of legitimacy of evidence/results used to get final reputation/trust value.

Another question – is (3) suitable candidate for a compromise between (1) and (2) or not? It may be suitable as replacement of (2) in case that agents have constrained memory, presuming that we are able to address issues of access control, confidentiality and availability.

3.2 Issues of evidence store

Initial ideas about attacks partitioned according to holder of evidence data:

1. Evidence held by the service provider – typically about numerous clients, evidence originated at this (one) source. Obviously, situation will be rather different when indirect evidence is taken into account:
   - Privacy – definitely the issue here, main focus would actually be unlinkability when not using any real identities.
   - Integrity is important with respect to server accountability in situations where servers have motivation to change content of data. It is not a big issue when evidence is used just locally but it could be a considerable problem with cooperating servers – recommenders.
   - Another problem is colluding servers – servers may create faked (positive or negative) evidence.
   - Availability attacks – may be interesting to disable recommending functionality, a profound property for distributed environment.
2. Evidence held by the client – it contains data that would probably be integrity-protected (otherwise clients may change evidence) by the service providers, the evidence is most likely only about the particular client. Client would like to increase positive evidence – selling/buying among colluding parties maybe an issue here – cryptography is the only countermeasure. Also, the client would often wish to keep his/her evidence confidential to himself/herself and the services that would request this evidence.

A related question is whether it is advantageous to store complete set of evidence about my previous behaviour and be able/must select the proper subset (ensuring positive response) while requiring a service. Note that selecting "bad" evidence can prompt the provider to lock the service, i.e. making this alias/ID useless to its owner. Note then in turn that this can be useful in case someone would consider leading an obvious related attack.

4 Accidents happen

One has to consider that building up a credit associated with a given pseudonym is usually not an effortless exercise – the value is determined by the cost (money, time, etc.). It is a loss when a pseudonym must be thrown away and a new pseudonym has to be built from scratch. This cost may be tuned by weights assigned to contexts of evidence pieces.

Issue – should this involve also the requirement to have this new pseudonym not related to the old pseudonym, i.e. take care of such (between-set) unlinkability? It is usually the case, otherwise it is pointless to create new identity. A simplistic view is that negative reputation is not relevant since an agent can just change its name and his trustworthiness is on an initial value. However, there is a risk of someone being able to link two identities of an agent, namely those who would profit from knowing the relation between the *NewID* that the *OldID* (pointing to a negative trust value for the *NewID* rather a zero/initial value). Contradiction between negative (trust/reputation) value and positive "context" value could be worth examining in this context. And obviously, the owner of *NewID* and *OldID* may not wish anyone to link two sets of evidence for privacy reasons.

4.1 Accident recovery

The question then is – *how hard it should be to create a new pseudonym with good reputation rating*? Good question – we believe that there are two requirements: be as friendly as possible to users, and sufficiently robust against possible attacks. We can set the cost of threshold "just useful" (for a given purpose) pseudonym with associated evidence from the statistical data about system usage. The problem is how to measure the cost [1] so that we can set it correctly. A vital issue to consider is then the cost difference between creating reputation for x and $x + k$ clients.

When a number of pseudonyms exist for a particular user, (s)he may (when deemed suitable to his/her own intentions) decide and link several pseudonyms to give a particular server enough evidence. However, his/her crucial requirement would be that the linkage depreciates as soon as possible – the data is part of one's history that is potentially valuable for a long period of time. So far we are aware of only one way to provide linkage that cannot be (ab)used in the future – to provide a zero-knowledge type of proof of the link existence. Yet while this solution might be used in some situations, on others (where the link as such must be shown) it would be of no use.

5 More questions instead of conclusions

The issues of anonymity and pseudonymity in relation to trust are investigated from various perspectives and in different environments [3, 6, 7, 12]. While we focus on the issue of evidence-based trust (reputation) with respect to clients (of and within) mix networks, there are clearly more areas being (and ready to be) explored.

Anonymous remailers constitute an existing application directly involved in anonymity issues. In fact, this is an application allowing for reasoning about several interesting aspects of reputation. The Sibyl attack [8] is of limited importance here as remailers are usually loosely federated systems with known set of nodes. Each node may locally compute properties of other nodes in the network and identify those sending considerably lower number of messages – i.e., suspected subset of colluding servers trying to reveal identity of mix clients.

Another goal of reputation reasoning can be to measure the behaviour of mix clients. As described in [11], an adversary may launch a statistical attack based on predictable behaviour of mix clients when their sets of recipients are rather small compared to the total number of mix clients. The reputation of clients may be used to identify messages with small predictable set of recipients and bounce them back to senders or generate dummy messages to *decrease* reputation of senders and thus increase set of probable recipients. Note that reputation is a gauge for behaviour predictability in this case.

One related issue with respect to use of mixes that has already been discussed and should be taken into account (e.g. [5]) relates to the use of a mechanism for detection of attacks based on number of colluding nodes – a probabilistic countermeasure of some realistic applicability. The idea is that you know whom you are asking for reference of someone's trustworthiness, but you (and neither a recipient of the request) have no idea about the path of your request – the randomness in the routing could be used to detect cliques of colluding subnets. Yet note that this is in turn essentially unlinkability attack from the privacy point of view.

Last but not least, enormously challenging area for future research concerns plausible ways to "depreciate" links between evidence (data), at least for the server (deletion/depreciation) only. Would fresh temporary pseudonyms (unlinkable to primary identities) used just for one-time links be the right way to approach this issue?

References

[1] A. Acquisti, R. Dingledine, and P. Syverson. On the Economics of Anonymity. In J. Camp and R. Wright, editors, Financial Cryptography, 7th International Conference, FC 2003, SpringerVerlag, LNCS 2742, pp. 84–102, 2003.

[2] R. J. Anderson. The Eternity Service. In Proceedings of Pragocrypt'96, pp. 242–252, Prague, Czech Republic, 1996.

[3] V. Cahill, et al.: Using Trust for Secure Collaboration in Uncertain Environments. In IEEE Pervasive Computing Magazine, pp. 52–61, July-September 2003.

[4] The Common Criteria Project Sponsoring Organisations: Common Criteria for Information Technology Security Evaluation - part 2, Version 2.1, August 1999.

[5] R. Dingledine and M. J. Freedman and D. Hopwood and D. Molnar: A Reputation System to Increase MIX-Net Reliability, In Proceedings of the 4th International Workshop on Information Hiding, Springer-Verlag LNCS 2137, pp. 126–141, 2001.

[6] R. Dingledine, N. Mathewson, and P. Syverson: Reputation in Privacy Enhancing Technologies, In Proceedings of the 12th annual conference on Computers, freedom and privacy, pp. 1–6, ACM Press, San Francisco, California, 2002.

[7] R. Dingledine, N. Mathewson, and P. Syverson: Reputation in P2P Anonymity Systems. Workshop on Economics of Peer-To-Peer Systems, June 2003, Berkeley, California, 2003.

[8] J. Douceur: The Sybil Attack. In 1st International Workshop on Peer-to-Peer Systems (IPTPS'02), Springer-Verlag LNCS 2429, pp. 251–260, 2002.

[9] M. Kinateder, S. Pearson: A Privacy-Enhanced Peer-to-Peer Reputation System. In K. Bauknecht, A. Min Tjoa, G. Quirchmayr (ed.), Proceedings of the 4th International Conference on Electronic Commerce and Web Technologies, EC-Web 2003, Prague, Czech Republic, September 2003.

[10] A. Rezgui, A. Bouguettaya, and Z. Malik. A Reputation-based Approach to Preserving Privacy in Web Services. In Proceedings of the 4th International Workshop of Technologies for E-Services, TES 2003, LNCS 2819, pp. 91–103, Berlin, Germany, September 2003.

[11] A. Serjantov and G. Danezis: Towards an Information Theoretic Metric for Anonymity. In 2nd International Workshop on Privacy Enhancing Technologies (PET 2002), LNCS 2482, pp. 41–53, San Franciskp, USA, April 2002.

[12] B. Shand, N. Dimmock and J. Bacon: Trust for Ubiquitous, Transparent Collaboration. In Proceedings of the First International Conference on Pervasive Computing and Communications, PerCom'03, pp. 153–160, Texas, USA, 2003.