SoK: Anonymous Credentials

Aarav Varshney[1]

[1]**Ashoka University**, aarav.varshney@alumni.ashoka.edu.in

Abstract

Anonymous credentials are a powerful cryptographic tool that enables users to prove some attribute about themselves without revealing their identity. This provides users with privacy while still allowing them to demonstrate their eligibility for specific services or privileges. In this SoK, we provide an overview of anonymous credential systems, including their history, design, and applications.

Keywords: *Anonymous Credentials, Participation Privacy, Pseudonyms*

1. Introduction

Data minimization as an abstract strategy describes the avoidance of unnecessary or unwanted data disclosures. The most fundamental information that can be disclosed about an individual is who he is (an identifier, or which observable events he is related to) [1]. If this information can be kept secret, the individual remains anonymous. Pfitzmann and Hansen, who pioneered the technical privacy research terminology, define anonymity as follows: Anonymity of a subject means that the subject is not identifiable within a set of subjects, the anonymity set [2].

The subject here is any entity defined by facts (names or identifiers), or causing observable events (by sending messages) while if an adversary cannot narrow down the sender of a specific message to less than two possible senders, the actual sender of the message remains anonymous. The two or more possible senders in question form the anonymity set.

Data minimization can be achieved through obfuscating of facts and events (messages, emails, and actions). If the attacker cannot detect confidential materials then they cannot link them to the subjects. This is defined as undectability [2]. We can achieve data minimization by removing any relation between the subject and the confidential materials. This is called unlinkability [2]. We discuss another form of data minimization through the use of psuedonyms. Introduced by Chaum [3], pseudonymity is related to anonymity as both concepts aim at protecting the real identity of a subject. The use of pseudonyms, however, allows to maintain a reference to the subject's real identity.

We structure the SoK by first discussing the history of anonymous credentials, followed by a discussion of the design of anonymous credentials, and finally, we discuss the applications of anonymous credentials.

2. History

A traditional credential, also known as a certificate or attribute certificate, is a collection of personal characteristics, such as name, date of birth, or personal identification number, that is certified by the issuer by signing it and binding it to the owner through cryptographic means,

requiring the use of the user's secret key [1]. Using either traditional or anonymous credentials is preferable to direct requests to the certifying party as this prevents user profiling by the certifying party. However, traditional credentials require the disclosure of all attributes together if the user needs to prove certain properties, making it possible for different uses of the same credential to be linked. In addition, the verifier and issuer can link the various uses of the user's credential to the credential issuance.

Chaum

Anonymous credentials were first introduced by Chaum in his paper *Security Without Identification: Transaction Systems To Make Big Brother Obsolete* [3]. The paper highlighted the struggle between individuals and organizations. Organizations face a challenge in maintaining a pervasive and efficient record-keeping system, which could potentially be linked to national identities and fingerprints to prevent abuse of system resources but that would compromise an individual's privacy.

To address this problem, Chaum proposed the use of **pseudonyms**, which could be used by individuals when interacting with different organizations, preventing them from being linked. Pseudonyms are the false names used to hide users' actual identities and maintains anonymity. This approach is similar to the use of public keys in modern cryptography, where an individual can use a different public key with each organization to receive encrypted messages that only they can decrypt [4].

Chaum also introduced a new approach to conducting three categories of transactions: communication, payment, and credential. For communication, he used the example of *dining cryptographers* [5] to send unconditionally untraceable messages and digital signatures for accountability. For payment, he introduced the idea of *blind signatures* [6], allowing individuals to pay for goods and services without revealing their identity. And finally, for credentials, he proposed using blind signatures to allow individuals to prove their eligibility for certain services or privileges without revealing their complete identity. Similar idea for proving eligibility was adapted by a later system, *Cinderella* [7].

Overall, Chaum's work on anonymous credentials paved the way for the development of new cryptographic techniques that prioritize individual privacy while still allowing for secure transactions.

Jan Camenisch and Anna Lysyanskaya

Camenisch et al. proposed an efficient and secure anonymous credential system that they claim is "considerably superior to previously proposed ones" [8]. In their paper, they provide a formal analysis of the system. The system includes three entities: a **user**, a **verifier**, and an **organization** where the user requests credentials from the organization, and upon successful verification of their previous credentials by the verifier, the organization grants the credential to the user.

Recall that in the anonymous systems so far [3, 9], the organizations know the users only by pseudonyms where different pseudonyms of the same user cannot be linked. Yet, an organization can issue a credential to a pseudonym, and the corresponding user can prove possession of this credential to another organization (who knows her by a different pseudonym), without revealing anything more than the fact that she owns such a credential. Credentials can be for unlimited use (these are called multiple-show credentials) and for one-time use (these are called one-show credentials).

The current system while still uses pseudonyms for identification, it also allows for optional revocation of anonymity in case of abuse, either through global or local revocation, revealing either the true identity or the pseudonym behind the credential. It also prevents transfer of credentials so that the user cannot sell or give away their credentials to other users. In such a scenario, a user who allows one of her friend to use one of her credentials once would allow that friend to impersonate her by giving access to all her credentials.

The non-transferability and optional revocation are variation of the basic credential system built by introducing a trusted *CA*. There are two major assumtions, firstly the CA is trusted to perform its duties properly and secondly, they assume that the channel between the user and the organization is secure and that the user is not being attacked by a malicious verifier. However, later credential systems like [10] have addressed these issues.

The security of the system relies on the strong RSA assumption and the decisional Diffie-Hellman assumption modulo a strong prime product. The RSA problem involves finding the plaintext given the public key and ciphertext, while the strong RSA assumption states that the problem remains difficult even if the attacker knows the public key exponent. The decisional Diffie-Hellman problem involves determining whether there exist integers x and y such that $a = g^x$, $b = g^y$, and $c = g^{xy}$, where g is a generator of a finite cyclic group G and $a, b, c \in G$.

In the later paper *Signature Schemes and Anonymous Credentials from Bilinear Maps* [11], Camenisch et al present an anonymous credential system based on LRSW assumption [9]. LRSW Assumption states that let G be a cyclic group with generator g and of order $|G|$. Let g^x and g^y be given. Furthermore, assume that an oracle can be called that answers a query s by a triplet (a, a^{sy}, a^{x+sxy}), where $a = g^z$ is a random group element of G. Let this oracle be called for $s_1, s_2, ...$, Then, the problem is to generate a quadruple $(t, b, b^{ty}, b^{x+txy})$, where t is not in $0, s_1, s_2, ...$, and where b is not equal to the identity element. Essentially, this assumes that forgery is hard even if the adversary has access to the oracle.

This is unique signature scheme since this is comparable to the efficiency of signatures schemes based on strong RSA assumption and no previous signature scheme based on an assumption related to the discrete logarithm assumption in the plain (as opposed to random-oracle) model comes close to the efficiency of schemes based on the strong RSA assumption. A signature scheme is integral for an anonymous credential system since a credential is a signed message that proves the user's eligibility for a service or privilege. To this end, they provide two protocols to firstly, prove knowledge of a signature on a committed message (verification of a credentials) and to obtain a signature on a committed message (issue of a credential).

The protocol to sign a committed message is important because a committed message would contain the identity of the user on which he would want the organization's signature. They do this by committing the message (m) as follows:

$$M = g^m Z^r \tag{1}$$

Here, M is the committed message, $Z = g^z$ where g is a generator of group G and $z, r \in G$. M is now an information-theoretically hidden message. Then using either the signature scheme B (single block message) or signature scheme C (multiple blocks message) a signature is made on M. Their signature scheme C is as follows [11]:

Key generation Run the *Setup* algorithm to generate $(q, G, \mathrm{G}, g, \ \mathrm{g}, e)$. Here, $G = \langle g \rangle$, $\mathrm{G} = \langle\ \mathrm{g} \rangle$ are two groups of prime order q that have a non-degenerate efficiently computable bilinear map e.

Choose $x \leftarrow \mathbb{Z}_q, y \leftarrow \mathbb{Z}_q$, and for $1 \leq i \leq \ell, z_i \leftarrow \mathbb{Z}_q$. Let $X = g^x, Y = g^y$ and, for $1 \leq i \leq \ell, Z_i = g^{z_i}$.

Set $sk = (x, y, z_1, \ldots, z_\ell), pk = (q, G, \mathrm{G}, g, \ \mathrm{g}, e, X, Y, \{Z_i\})$.

Signature On input message $\left(m^{(0)}, m^{(1)}, \ldots, m^{(\ell)}\right)$, secret key $sk = (x, y, \ z_1, \ldots, z_\ell)$, and public key $pk = (q, G, \mathrm{G}, g, \ \mathrm{g}, e, X, Y, \{Z_i\})$ do:

- Choose a random $a \leftarrow G$.
- Let $A_i = a^{z_i}$ for $1 \leq i \leq \ell$.
- Let $b = a^y, B_i = (A_i)^y$.
- Let $c = a^{x+xym^{(0)}} \prod_{i=1}^{\ell} A_i^{xym^{(i)}}$.

Output $\sigma = (a, \{A_i\}, b, \{B_i\}, c)$.

Verification On input $pk = (q, G, \mathrm{G}, g, \ \mathrm{g}, e, X, Y, \{Z_i\})$, message $\left(m^{(0)}, \ldots, \ m^{(\ell)}\right)$, and purported signature $\sigma = (a, \{A_i\}, b, \{B_i\}, c)$, check the following:

1. $\{A_i\}$ were formed correctly: $e\left(a, Z_i\right) = e\left(g, A_i\right)$.
2. b and $\{B_i\}$ were formed correctly: $e(a, Y) = e(g, b)$ and $e\left(A_i, Y\right) = e\left(g, B_i\right)$
3. c was formed correctly: $e(X, a) \cdot e(X, b)^{m^{(0)}} \cdot \prod_{i=1}^{\ell} e\left(X, B_i\right)^{m^{(i)}} = e(g, c)$.

Note that while engaged in the protocol for obtainaing a signature, the user has to show a zero-knowledge proof of knowledge of the opening of the commitment.

$$PK\left\{\left(\mu^{(0)}, \ldots, \mu^{(\ell)}\right) : M = g^{\mu^{(0)}} \prod_{i=1}^{\ell} Z_i^{\mu^{(i)}}\right\} \tag{2}$$

Here, $\mu^{(0)}, \ldots, \mu^{(\ell)}$ are the quantities that the user has to prove knowledge of.

Once this signature is made, we use the protocol to prove knowledge of a signature on a committed message which requires another zero-knowledge proof.

3. Design

The issue of providing anonymity to users is a complex problem, and although there are several systems claiming to offer anonymity, it is challenging to determine the level of anonymity provided by these systems. Serjantov et al. proposed a probability-based model to define the extent of anonymity provided by a system [12]. This model seeks to quantify the level of anonymity based on the probability of identifying the user who generated a particular message.

In a similar vein, David Chaum, while studying the security of Dining Cryptographers' networks, introduced the concept of an anonymity set [5]. The anonymity set is defined as the group of all possible participants in a network who could have sent a specific message, while all communication in the network is visible to an attacker. According to Chaum, the larger the anonymity set, the more difficult it becomes for an attacker to identify the sender of a message. Thus, the size of the anonymity set is an excellent indicator of the degree of anonymity provided by the network.

However, attacks in their literature review showed the probability of any subject sending or receiving a message is uniform and hence the size of the set is a good indicator of anonymity. However, in real-world scenarios, the probability of a subject sending or receiving a message is not uniform since an attacker can secure additional information and reduce the set size based on this information. For instance, in Figure 1, we see that if the attacker has the information that A sent a message to R then the attacker can immediately deduce the fact that S recieved a message from E. This is depsite the fact that the anonymity set for senders of S is {A, B, C, D, E}.

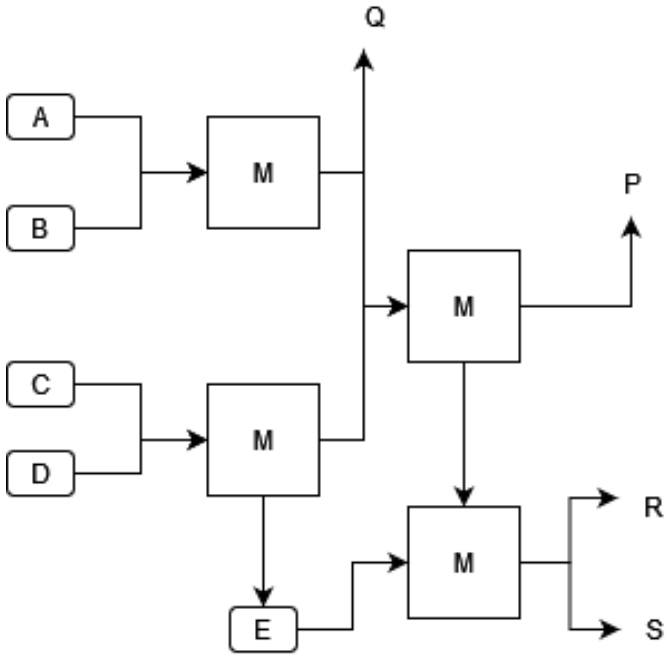

Figure 1: An attack in mix networks. Here, M are the mixes and A, B, C, D, E are the senders.

Following this review we come to a conlcusion that we need a better metric to determine the quality of anonymity offered to the users. Serjantov et al. introduce entropy to describe quality of anonymity [12]. Entropy is a measure of the uncertainty of a random variable [13]. In this scneario, the attacker has a probability distribuition, P, over the set of all participants with regards to them sending or receiving a message. They now define the effective size of the probability distribuition to be equal to the entropy of the distribuition:

$$S = -\sum_{u \in \Psi} p_u \log_2 p_u \tag{3}$$

Here, p_u is $P(u, r)$, Ψ is the set of all participants and $r \in \{\text{sender, recepient}\}$. This can also be interpreted as the number of bits of additional information required to identify a user of a particular message.

4. Applications

Idemix [14] is an implementation of an anonymous credential system based on protocols described in [8]. The system offers an easy-to-use access control mechanism that can be integrated with user applications. The library needs to be imported into the application, and the application extends the library with its own access control policies.

For instance, the authors illustrate the system's utility in the context of a New York Times (NYT) news service subscription. In this example, a user first acquires a credential from a bank (ARGENTIX) for $10, which is then shown to the New York Times news subscription service (KIOSK) to obtain another credential. The final credential is presented to NYT to get the subscription. Throughout these transactions, pseudonyms are used, and the process is unlinkable, thereby ensuring full anonymity for the user.

However, the system still requires an external trusted third party to issue root pseudonyms and credentials based on some verification of the user's real identity. To prevent credential sharing, the authors propose two types of safeguards: *all-or-nothing*, where the user can share all or none of their credentials, and *PKI-assured non-transferability*, which links a master secret to a valuable secret key from outside the system (e.g., he secret key that gives access to the user's bank account). The latter option is rarely available.

Optional anonymity revocation also requires a third party to reveal the pseudonym used for the transaction (local revocation) or reveal the user's real identity (global revocation). Both the user and the organization need to agree on the conditions that would lead to revocation (e.g., illegal transactions).

The system's implementation uses zk-proofs to validate user credentials, resulting in significant computational overhead on both the user and organization sides. The authors' benchmark results show that registering a pseudonym, issuing a credential, and showing a credential can take up to 25 seconds.

References

[1] S. Fischer-Hbner and S. Berthold, "Chapter 53 - privacy-enhancing technologies," in *Computer and Information Security Handbook (Third Edition)*, J. R. Vacca, Ed. Boston: Morgan Kaufmann, 2017, pp. 759–778. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128038437000533

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

[3] D. Chaum, "Security without identification: Transaction systems to make big brother obsolete," *Commun. ACM*, vol. 28, no. 10, p. 1030–1044, oct 1985. [Online]. Available: https://doi.org/10.1145/4372.4373

[4] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Commun. ACM*, vol. 21, no. 2, p. 120–126, feb 1978. [Online]. Available: https://doi.org/10.1145/359340.359342

[5] D. Chaum, "The dining cryptographers problem: Unconditional sender and recipient untraceability," *J. Cryptology*, vol. 1, pp. 65–75, 1988.

[6] ——, "Blind signatures for untraceable payments," in *Advances in cryptology–CRYPTO'82*. Springer, 1983, pp. 199–203.

[7] A. Delignat-Lavaud, C. Fournet, M. Kohlweiss, and B. Parno, "Cinderella: Turning shabby x.509 certificates into elegant anonymous credentials with the magic of verifiable computation," in *2016 IEEE Symposium on Security and Privacy (SP)*, 2016, pp. 235–254.

[8] J. Camenisch and A. Lysyanskaya, "An efficient system for non-transferable anonymous credentials with optional anonymity revocation," *IACR Cryptol. ePrint Arch.*, vol. 2001, p. 19, 2001.

[9] A. Lysyanskaya, R. L. Rivest, A. Sahai, and S. Wolf, "Pseudonym systems," in *Selected Areas in Cryptography*, H. Heys and C. Adams, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 184–199.

[10] L. Wang, G. Asharov, R. Pass, T. Ristenpart, and A. Shelat, "Blind certificate authorities," in *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, May 2019.

[11] J. Camenisch and A. Lysyanskaya, "Signature schemes and anonymous credentials from bilinear maps," in *Advances in Cryptology – CRYPTO 2004*, M. Franklin, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 56–72.

[12] A. Serjantov and G. Danezis, "Towards an information theoretic metric for anonymity," in *Privacy Enhancing Technologies*, R. Dingledine and P. Syverson, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 41–53.

[13] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[14] J. Camenisch and E. Van Herreweghen, "Design and implementation of the idemix anonymous credential system," in *Proceedings of the 9th ACM Conference on Computer and Communications Security*, ser. CCS '02. New York, NY, USA: Association for Computing Machinery, 2002, p. 21–30. [Online]. Available: https://doi.org/10.1145/586110.586114