Attaching Multiple Personal Identifiers in X.509 Digital Certificates

Prokopios Drogkaris and Stefanos Gritzalis

Laboratory of Information and Communication Systems Security,
Department of Information and Communication Systems Engineering,
University of the Aegean Samos, GR-83200, Greece
{pdrogk,sgritz}@aegean.gr

Abstract. The appeals for interoperable and decentralized Electronic Identity Management are rapidly increasing, especially since their contribution towards interoperability across the entire "electronic" public sector, effective information sharing and simplified access to electronic services, is unquestioned. This paper presents an efficient and user-centric method for storing multiple users' identifiers in X.509 digital certificates while preserving their confidentiality, allowing for interoperable user identification in environments where users cannot be identified by an all embracing unique identifier.

Keywords: Privacy, Identity Management, Digital Certificate, e-Government.

1 Introduction

The requirements raised by e-Government environments for interoperability, acceptance and coherence have strengthen the demands for underlying Identity Management Systems. As European Union states in [4] "Electronic Identity Management is a cornerstone of the implementation of the full range of e-Government services, for both citizens and businesses, across the Union". As more government, personal and commercial transactions are conducted electronically – especially where documents exist only in digital form – parties need to be sure of a person's or an organization's identity". Several protocols and standards have been proposed towards this direction; however either most of them involve a centralized entity for administrating user identities or attributes or they require the establishment of trust relationships among different Identity Providers (IdP).

In this paper we propose an efficient and user-centric method for storing multiple users' identifiers in X.509 digital certificates while preserving their confidentiality. To the best of our knowledge, similar research work has not been published yet. The rest of the paper is structured as follows: Section 2 presents the Subject Identification Method on which our proposal is based; Section 3 discusses the necessity of being able to store multiple identifiers; Section 4 presents our proposal while Section 5 provides an evaluation on the proposed methodology. Finally, Section 6 concludes the paper providing directions for future work.

J. Camenisch and C. Lambrinoudakis (Eds.): EuroPKI 2010, LNCS 6711, pp. 171–177, 2011.

2 Subject Identification Method

RFC 4683: Internet X.509 Public Key Infrastructure Subject Identification Method (SIM) [1] is a request for Comments (RFC) published by the Internet Engineering Task Force (IETF); its current status is ‘Proposed Standard”. What SIM proposes is to include user’s privacy-sensitive identifier in the *otherName* field in the *subjectAltName* extension of user’s X.509 certificate. User is being assigned an identifier or Sensitive Identification Information (*SII*) which relates to a certain Sensitive Identification Information type (*SIItype*), an object identifier that specifies the type of SII. User selects a password *P* and along with *SIItype* and *SII* forwards them to the Registration Authority (RA) through a secure channel. After RA validates the association between user, SII and SIItype, generates a random value *R* and calculates the *SIM* value as described in formulas (1) & (2), where *H ( )* is a cryptographically secure hash algorithm. *SIM* value is then passed on from the RA to the user and the Certification Authority (CA) through a secure channel. When user sends a certificate request to the CA, she issues user’s X.509 digital certificate including *SIM* value in *otherName* field which is part of *subjectAltName* extension.

$$\mathrm{PEPSI} = \mathrm{H}\,(\,\mathrm{H}\;(\,\mathrm{P} \,||\, \mathrm{R} \,||\, \mathrm{SIItype} \,||\, \mathrm{SII}\,)\,)\,. \tag{1}$$

$$\mathrm{SIM} = \mathrm{R} \,||\, \mathrm{PEPSI}\,. \tag{2}$$

$$\mathrm{PEPSI}' = \mathrm{H}\,(\,\mathrm{H}\;(\,\mathrm{P} \,||\, \mathrm{R} \,||\, \mathrm{SIItype} \,||\, \mathrm{SII}\,)\,)\,. \tag{3}$$

$$\mathrm{SIM}' = \mathrm{R} \,||\, \mathrm{PEPSI}'\,. \tag{4}$$

$$\mathrm{H}\;(\,\mathrm{P} \,||\, \mathrm{R} \,||\, \mathrm{SIItype} \,||\, \mathrm{SII}\,)\,. \tag{5}$$

Every time that the user wishes to request an electronic service from a service provider (SP), she must transmit SP, through a secure channel, her *SII, SIItype, P* and her certificate. SP can then compute *PEPSI′* and *SIM′*, through formulas (3) and (4), using user’s submitted values, compare *SIM′* to *SIM* value from user’s digital certificate and verify user’s *SII.* In cases where the SP already knows user’s SII and the user wants to prove that she is the owner of the specific identifier, she must send to the SP only P and her digital certificate. Again, SP can compute *PEPSI′* and *SIM′*, through formulas (3) and (4) compare *SIM′* to *SIM* value and verify user’s *SII.* Finally, when the user wants to present proof to SP that she is the subject of a SII, without disclosing the identifier itself to SP, she submits the value of formula (5) to SP, which is an intermediate value of formula (1), along with her digital certificate. The SP can then acquire *R* from *SIM* value in user’s certificate, compute the hash of formula (5), compare it to the SIM value and verify user’s knowledge of *P* and *SII.*

3 Multiple Identifiers

The aforementioned identification method could be deployed by several European countries which have adopted the utilization of a national unique identifier for each citizen [7]. In fact this identification scheme seems much more suitable for electronic services, since every user can be easily identified, irrespective of the requested service, and can also ease the exchange of information (interoperability) among different public departments. However, not all environments can uniquely identify their users through an all embracing identifier; instead they utilize multiple identifiers, one for each sub-environment where each participant – user can be uniquely identified through a sectorial identifier. Even if the transition towards the introduction of an all embracing identifier is obvious, it is not that straightforward from a technical and also from a legal point of view. Each public service should have to update their records with the new user identifier, a process not only costly but also of high risk. Moreover, in several occasions, an all embracing identifier is not viable due to constitutional or legal constrains. In Greece for example, as described in [9], the dignity and the right to protect personal data [10], according to the Constitution, sets a normative obstacle to the intentions of the Greek Government to deploy such a solution.

4 Proposed Method

In this paper we propose an efficient user-centric method that utilizes a Public Key Infrastructure, for storing multiple users' identifiers in X.509 digital certificates while, at the same time, preserving their confidentiality. This method sets the basis for interoperable Electronic Identity Management (eIDM) by enabling seamless, portable and user-centric digital identity exchange and a decentralized validation. Through the support of multiple identifiers, our proposal is not only applicable to environments which identify users through multiple identifiers but also in cross environment identification. The basis of our proposal is Subject Identification Method (SIM) as described in Section 2, which is applicable to environments where users can be identified through an all embracing unique identifier. The differentiation of our proposal, compared with SIM methodology, lies in the introduction of distinct passwords, PEPSI and SIM values for each identifier along with a master password P_{master}.

4.1 SIM Computation

Let's assume that user A has been assigned *n* identifiers (ID_1, ID_2 … ID_n) and each one corresponds to an explicit SIItype. Each one of these identifiers must be related to a unique password P_{IDx}, where $x=\{1,2,3,\ldots,n-1,n\}$. However, the selection of *n* different passwords (one for each identifier) would be highly impractical from password management point of view. Consequently, the user selects a master password P_{master}, which will be utilized for the generation of the required n unique P_{IDx} passwords, based on formula (6), where *H( )* is a cryptographically secure hash algorithm and SIItype is the identifier type. The concatenation of P_{master} and SIItype ensures that each password will be unique, since two user identifiers cannot correspond to the same type and the utilization of a secure cryptographic hash function *H( )* ensures that master password will not be disclosed to unauthorized parties, enabling them to compute identifier passwords P_{IDx}.

$$P_{IDx} = H\ (P_{master} \,||\, SIItype\ )\ . \qquad x=\{1,2,3,\ldots,n\text{-}1,n\} \tag{6}$$

$$PEPSI_{IDx} = H\ (\ H\ \ (P_{IDx} \,||\, R_{IDx} \,||\, SIItype \,||\, IDx\ )\ )\ . \qquad x=\{1,2,3,\ldots,n\text{-}1,n\} \tag{7}$$

$$SIM_{IDx} : (\ R_{IDx} \,||\, PEPSI_{IDx})\ . \qquad x=\{1,2,3,\ldots,n\text{-}1,n\} \tag{8}$$

$$SIM_{total} : (\ SIM_{IDx} \,||\, SIM_{IDx+1} \,||\, SIM_{IDx+2} \,||\, \ldots \,||\, SIM_{IDn-1} \,||\, SIM_{IDn}\ )\ . \qquad x=\{1,2,3,\ldots,n\text{-}1,n\} \tag{9}$$

After the user has computed her P_{IDx} passwords, she informs the Registration Authority (RA), through a secure communication channel, about her identifiers, their SIItype and their passwords. The user may choose to submit less identifiers than the ones that she has been assigned; for simplicity we assume that she submits all her identifiers. The RA verifies that the submitted identifiers are correctly associated with user and SII type and generates a random value R_{IDx}, one for each submitted identifier. This random value must be created by a Random Number Generator (RNG) that meet the requirements defined in FIPS 140-2 [12] and it's length must mandatory be the same as the output of the secure hash algorithm utilized by RA. Having generated *n* random values (R_{ID1}, R_{ID2}, ... P_{IDn}), RA must then compute *n* PEPSI values, one for each identifier, based on formula (7), where *H( )* is again a cryptographically secure hash algorithm. RA must then compute SIM_{IDx} value for each identifier, based on formula (8), similarly to SIM method described in Section 2, SIM_{ID} value consists of $PEPSI_{ID}$ value concatenated with Random value. After RA computes *n* SIM_{ID} values, one for each identifier, concatenates them to each other and creates SIM_{total} value, as shown in formula (9). Since the number of identifiers that could be utilized to identify a user in some environment is predetermined, so is the number of SIM_{ID} that could exist. Consequently, a prearranged sequence of blocks allows easy retrieval of each SIM_{ID} and consequently of each $PEPSI_{ID}$ and R_{ID}. In cases where the user does not wish to submit an identifier, the corresponding SIM_{ID} cannot be replaced by another and must be left blank, preserving the aforementioned sequence. Finally, the RA informs the user of her SIM_{total} value and passes it on to the Certification Authority (CA). When the user sends a certificate request to the CA, she issues user's X.509 digital certificate including SIM_{total} value in *otherName* field.

4.2 Identification from Service Providers

Each time a user wishes to request an electronic service from a Service Provider (SP), she should inform the SP of her identifier and also present proof that she has been assigned the specific identifier. However, there are also cases where the SP has some prior knowledge about user's identifier, either from the user herself or from another electronic service, or more infrequently, cases where only proof that the user has been assigned a specific identifier type is required. The proposed methodology is able of supporting all these cases, based on the amount of information the user submits to the SP during an electronic service request.

The first case, which is the most common, is when the SP does not have any information on user who requests the provision of an electronic service. Thus, during the request, the user must also submit to the SP her identifier, the corresponding SIItype, the identifier password (P_{ID}) and her digital certificate. SP can then compute $PEPSI'_{ID}$ value based on formula (10) and knowing the predefined sequence in which SIM_{ID} are stored in SIM_{total}, compare it with SIM'_{ID}. If there is a match then the user has submitted the correct identifier and if she is eligible, can start using the requested electronic service.

$$PEPSI'_{ID} = H\,(\,H\;\,(P_{ID} \,||\, R_{ID} \,||\, SIItype \,||\, ID\,)\,)\;. \tag{10}$$

$$SIM'_{ID} : (\,R_{ID} \,||\, PEPSI'_{ID})\;. \tag{11}$$

$$H\;\,(P_{ID} \,||\, R_{ID} \,||\, SIItype \,||\, ID\,)\;. \tag{12}$$

The second case is when SP already knows user's identifier (ID) and only needs proof that the specific user has been assigned the specific identifier. In that case, the user submits her identifier password (P_{ID}) and her digital certificate. Similarly to the first case, the SP can compute PEPSI′ value based on formula (10) and knowing the predefined sequence in which SIM_{ID} are stored in SIM_{total}, compare it with SIM′. The third case is when the user wants to prove to the SP that she is the subject of an identifier, without disclosing the identifier itself to the SP. The user computes the value of formula (12) and submits it to the SP along with her certificate. In order to perform this computation, the user must acquire R_{ID} value from SIM_{ID}, which is stored on her digital certificate. Similarly to her, the SP also acquires R_{ID} from user's certificate, computes the hash function of the value that the user submitted earlier and concatenates it with R_{ID}. If the outcome of this computation matches the value of SIM_{ID} stored in user's certificate, then the user's knowledge of identifier (ID) and SIItype is verified.

5 Evaluation

Since users' digital certificates can be publicly available, the preservation of identifiers confidentiality is compulsory and is achieved through the repetitive usage of secure cryptographic hash functions on identifier (ID), identifier password (P_{ID}) and Random Value (R_{ID}). Since several security flaws have been identified in SHA-1 [16], we propose the utilization of SHA-2 or a more secure hash algorithm. Moreover, another vital aspect of SIM_{total} confidentiality is identifiers' passwords (P_{ID}) strength as users tend to chose easy memorable passwords that relate to personal information. Consequently, users must be encouraged to select strong passwords that will at least comply with FIPS 112 and FIPS 180-1 [2][3]. Finally, Random Value (R_{ID}) should be created through a Random Number Generator (RNG) that meets the requirements defined in FIPS 140-2 [12].

Another vital aspect regarding the viability of the proposed method is the measurement of the computation overhead introduced in the overall environment due to the increased size of the user's certificate. Certificate's size depends on the number of indentifies (n), the size of hash function digest and the length of Random Value (R_{ID}). Currently we are conducting performance simulations using various different values in the aforementioned parameters, in order to estimate the relation between performance and identifier's confidentiality.

6 Conclusions

In this paper we have proposed a user-centric method for storing multiple users' identifiers in X.509 digital certificates. Through our proposal, users' identification can be performed based solely on digital certificates, thus diminishing the necessity for a centralized Identity Provider that will be responsible for managing user identifiers, attributes and roles. Even if the proposed methodology has been designed having in mind the specific needs and requirements of e-Government environments, it can also be applied in cross environment identification, through the deployment of a catholic Public Key Infrastructure. User identifiers' regarding bank institutions or medical environments could also be embodied in digital certificates, allowing for higher levels of interoperability amongst different environments, while preserving their confidentiality. Our next steps, relate to the design of a framework that will support the adaptation of the proposed methodology from organizations outside the initial environment.

References

1. Park, J., Lee, J., Lee, H., Park, S., Polk, T.: Internet X.509 Public Key Infrastructure Subject Identification Method (SIM), National Institute of Standards and Technology (2006)
2. Federal Information Processing Standards, Publication (FIPS PUB) 112, Password Usage (1985)
3. Federal Information Processing Standards Publication (FIPS PUB) 180-1, Secure Hash Standard (1995)
4. Europe's Information Society Thematic Portal. A question of identity, `http://www.ec.europa.eu`
5. Mont, C., Bramhall, P., Pato, J.: On Adaptive Identity Management: The next generation of Identity Management Technologies, HP Labs Technical Report, HPL-2003-149 (2003)
6. Lips M.: Identity Management in Information age Government exploring Concepts, Definitions, Aproaches and Solutions (2008)
7. Hayat, A., Leitold, H., Rechberger, C., Rossler, T.: Survey on EU's Electronic-ID Solutions, Vienna (2004)
8. Drogkaris, P., Lambrinoudakis, C., Gritzalis, S.: Introducing Federated Identities to One-Stop-Shop e-Government Environments: The Greek Case. In: Cunningham, P., Cunningham, D. (eds.) 19th Conference on eChallenges 2009, Istanbul, Turkey. eChallenges e-2009 Conference Proceedings, pp. 115–121 (October 2009)
9. Drogkaris, P., Geneiatakis, D., Gritzalis, S., Lambrinoudakis, C., Mitrou, L.: Towards an Enhanced Authentication Framework for eGovernment Services: The Greek case. In: Ferro, E., Scholl, J., Wimmer, M. (eds.) EGOV 2008, 7th International Conference on Electronic Government, Torino, Italy, pp. 189–196. Trauner Verlag Schriftenreihe Informatik (September 2008)

10. Greek Constitution Articles 2 § 1 (human dignity) and 9 A (right to protection of personal data)
11. Hayat A., Leitold H., Rechberger C., Rössler T.: Survey on EU's Electronic-ID Solutions', Vienna (2004)
12. Federal Information Processing Standards Publication (FIPS PUB) 140-2, Security Requirements for Cryptographic Modules (2001)
13. Menezes, A., Oorschot, P., Vanstone, S.: Handbook of Applied Cryptography. CRC Press, Boca Raton (1996)
14. Lenstra, A., Verheul, E.: Selecting cryptographic key sizes. Journal of Cryptology 14(4), 255–293 (2001)
15. Federal Information Processing Standards Publication 180-2, Secure hash standard (2002)
16. Schneier on Security, Cryptanalysis of SHA-1, `http://www.schneier.com`
17. McKenzie, R., Crompton, M., Wallis, C.: Use Cases for Identity Management in E-Government. IEEE Security and Privacy 6(2), 51–57 (2008)
18. Greenwood, D., Dempster, A., Laird, M., Rubin, D.: The context for Identity Management Architectures and Trust Models. In: OECD Workshop on Digital Identity Management (2007)
19. Directive 97/66/EC of the European Parliament and of the Council concerning the processing of personal data and the protection of privacy in the telecommunications sector. Official Journal L L 024, 1–8 (1997)
20. Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures. Official Journal L 013, 00120020 (2000)
21. Directive 01/45/EC of the European Parliament and the Council of Ministers on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data. Official Journal L 008, 122 (2001)
22. Hansen, M., Pfitzmann, A., Steinbrecher, S.: Identity management throughout one's whole life. Information Security Technical Report 13(2), 83–94 (2008)