Enhancing privacy of users in eID schemes

Shrishak, Kris; Erkin, Zekeriya; Schaar, Remco

Publication date
2016

Document Version
Final published version

Published in
37th WIC Symposium on Information Theory in the Benelux / 6th WIC/IEEE SP Symposium on Information Theory and Signal Processing in the Benelux

Citation (APA)
Shrishak, K., Erkin, Z., & Schaar, R. (2016). Enhancing privacy of users in eID schemes. In *37th WIC Symposium on Information Theory in the Benelux / 6th WIC/IEEE SP Symposium on Information Theory and Signal Processing in the Benelux* (pp. 158-165)

Enhancing privacy of users in eID schemes

Kris Shrishak
Cyber Security Group, Department of Intelligent Systems
Delft University of Technology, The Netherlands
k.s.sridaran@student.tudelft.nl

Zekeriya Erkin
z.erkin@tudelft.nl

Remco Schaar
UL Transaction Security
The Netherlands
remco.schaar@ul.com

Abstract

In todays world transactions are increasingly being performed over the internet but require identification of users as in face-to-face transactions. In order to facilitate eGovernance as well as other eCommerce services Electronic Identification (eID) schemes, which intend to provide unique and reliable identification and authentication of the users, have been introduced. eID schemes commonly involve a Service Provider which provides a service, such as online shopping, to the user and an Identity Provider which verifies the users identity and facilitates the user to identify itself to the Service Provider. Every transaction made over the Internet reveals bits of information about the user which can be accumulated and abused, thus necessitating security and privacy in order to prevent misuse of data and invasion of personal privacy. In this work, five eID schemes which are in use or are proposed in EU countries is surveyed and the strengths and weaknesses of these schemes is investigated. All the schemes have given importance to security while only a few of them are designed with privacy in mind. Identity Providers in federated eID schemes are observed to be a privacy hotspot as they store user information and can uniquely identify the user. The use of homomorphic encryption and block chain in eID schemes is further explored in order to prevent the Identity Provider from becoming a privacy hotspot while fulfilling its role in the scheme.

1 Introduction

Long before the Internet came into existence, Governments have had public authentication schemes by issuing Identity documents to identify a person or verify aspects of a person's personal identity. These documents were trusted not only by the governments but also by businesses that required reliable authentication of users [1]. Today the number of services offered online is increasing rapidly and this growth has forced users to maintain multiple credentials for authentication and identification to service providers, which has caused security and usability issues. Password-based authentication mechanism employed by most service providers has lead to users reusing the same password or writing them on paper. Usage of other authentication mechanisms such as hardware security tokens is also not convenient. As a result, many countries in the European Union (EU) have either developed an electronic identification (eID) scheme or are in the process of developing one. All schemes have put security at the forefront but few have considered the privacy implications of a nation-wide single identification scheme. We consider it essential that designers of eID systems focus on privacy, specifically informational privacy of users which we interpret such that when a user mentions 'my information', the 'my'

> "is not the same as 'my' in 'my car' but rather the same as 'my' in 'my body' or 'my feelings'; it expresses a sense of constitutive belonging, not of external ownership, a sense in which my body, my feelings, and my information are part of me but are not my possessions [2]."

We survey five eID schemes in the EU and provide brief descriptions of two solutions which could be applied in eID schemes to improve privacy. Information privacy in eID schemes is investigated in terms of the following properties [3]:

1. Anonymity: Users may use a service without disclosing their identity.
2. Pseudonimity: Users may utilise a service by using pseudonyms.
3. Data minimization: Only the required information about the user must be shared in order to prevent misuse.
4. Unlinkability: User should be able to use resources and services without others being able to link these activities.
5. Unobservability: Users should be able to use services without being observed by others.
6. Transparency: User data should be obtained only when necessary and after user consent.

2 eID Schemes

eID systems intend to provide reliable identification and authentication of the users. The eID systems discussed in this section were designed for use by both public and private services. The parties involved in these systems differ widely but commonly involve the following:

- ◇ *User* - wants to authenticate her/himself to the Service Provider to access a resource.
- ◇ *Service Provider* (SP) - provides a service, such as online shopping or government tax services, and makes transaction decisions based upon the acceptance of a users authenticated credentials and attributes.
- ◇ *Identity Provider* (IDP) - verifies the user's identity or credentials and facilitates the user to authenticate her/himself to the SP. It improves the overall usability since the user does not need to remember multiple authentication credentials.

2.1 Belgian eID scheme

The Belgian eID scheme is a nation-wide Public Key Infrastructure (PKI) which requires each citizen to present her/himself at the municipality for strong user authentication during the issuing phase. Thus it can be inferred that the Belgian government has taken up the role of IDP. The user is issued a smart card and is required to buy a card reader for online use. The objective of the Belgian eID card has been to fulfil four functions, namely, citizen identification, authentication, digital signature and access control [4].

The eID card has the name, title, nationality, place and date of birth, gender, and a photo of its holder printed on it in addition to a hand written signature of its holder and of the civil servant who issued the card. All this information is also stored on the chip in an Identity file which is signed by the National Register (RRN). The chip also contains an address file which is kept independently as the address of its holder may change within the validity period of the card. The RRN signs the address file together with the identity file to guarantee the link between these two files. The corresponding signature is stored as the address file's signature.

Privacy Analysis

The main concern with the Belgium eID card is privacy. The user's identity is revealed for all transactions and (s)he does not have control over the data that is shared with the SPs while it remains possible for colluding SPs to link user activities as the authentication certificate contains the RRN. Verhaeghe et. al [5] provide privacy and security threats if the card or the middleware is compromised.

2.2 GOV.UK Verify

GOV.UK Verify is the eID scheme of the United Kingdom. It is a federated identification infrastructure where an online central hub mediates user authentications between IDPs and SPs. The role of the hub is to ensure interoperable identification and authentication as well as provide privacy benefits by hiding the IDP from the SP. The eID scheme has been designed considering nine Identity Assurance Principles - user control, transparency, multiplicity, data minimization, data quality, service user access and portability, certification, dispute resolution and exceptional circumstances [6]. Users of the eID scheme are identified by a pseudonym ($\boldsymbol{u}$) at the hub. In addition to IDPs, SPs and hub, the scheme also includes (1) Attribute Providers (ATP), which are responsible for establishing attributes and (2) Matching Service (MS), which helps validate assertions from IDPs and derives a pseudonym $\boldsymbol{v}$ from $\boldsymbol{u}$ for the SPs it serves [7].

Privacy Analysis

In spite of claiming that privacy is one of the design criteria for the eID scheme, GOV.UK Verify has multiple privacy issues. The hub, which has full visibility of the user pseudonym and personal information of citizens can link interactions of the same user across different SPs as well as undetectably impersonate users at any SP without user authentication. The MS, which has the task of matching pseudonyms and the attributes into a local account, can link the user and the SPs that choose the same MS. Colluding SPs using the same MS can link user activities as the same pseudonym is used. Thus it can be inferred that GOV.UK Verify actually degrades the privacy of citizens instead of enhancing it. Finally, the hub can be used for undetectable mass surveillance.

2.3 Dutch eID scheme using polymorphic pseudonyms

The Dutch eID scheme utilises a federated eID infrastructure. Some variations of the Dutch eID scheme have been proposed during the planning stage. In the following, the version with polymorphic pseudonyms [8] is considered as it is the most privacy friendly version among those proposed. In addition to IDPs and SPs, the scheme also includes (1) a Pseudonym Provider (PP), which generates polymorphic pseudonym when the IDP sends a unique identifier($U\text{-}id$) (2) brokers, which mediate user authentication between the various eID parties, essentially IDPs and the many SPs. Other parties include Attribute Providers (ATP) and Authorization Providers, a Key Management Authority (KMA) and a investigation authority. The pseudonymization in this scheme involves three levels of pseudonyms, namely, (1) polymorphic pseudonyms (2) encrypted pseudonyms and (3) pseudonyms.

Privacy Analysis

The Dutch eID scheme with polymorphic pseudonyms has few of the privacy properties mentioned in section 1. It provides pseudonimity to the users such that SPs get a user specific pseudonym from the IDP, but the pseudonym is independent of the IDP. Meanwhile, the pseudonym obtained by different SPs for the same user is not the same, thus preventing colluding SPs from being able to link user activities. Finally, encrypted pseudonyms are randomized such that the broker cannot identify the same user on multiple occasions.

In spite of efforts made to provide privacy to the users, this scheme has countable issues. IDP is a privacy hotspot which knows the attributes of users, which SPs are visited by users and how often. This information might be considered very sensitive in some cases. The IDP does not need to know this information in order to perform its role. An alternative proposed in [8] is to store polymorphic pseudonyms in a smart card and use chip authentication as in the German eID scheme. The Dutch eID scheme is complex and it is possible that in the future SPs might outsource the decryption of encrypted pseudonyms to a third party allowing them to learn users visiting patterns.

2.4 German eID scheme

The German eID scheme uses a direct authentication eID infrastructure. The design goals of this scheme were data minimization, data security and transparency. In order to use this scheme, the user needs an eID card, a reader and the Ausweisapp software while the SP needs an authorization certificate, an eID-Server that handles authentication by communicating with the card. The eID card contains a chip which stores the information printed on the card as well as the fingerprints of the holder, if the holder wishes. The document number and the fingerprints can be read offline only by authorities who have machines certified by the Federal Office for Information Security (BSI).

The German eID scheme makes use of cryptographic protocols, to perform mutual identification, which are also used in EU passports [9]. Password Authenticated Connection Establishment (PACE) protocol provides secure communication and explicit password-based authentication of the eID card and the terminal while the Extended Access Control (EAC) protocol provides secure key establishment between a chip card and a terminal, using a PKI. It serves the purpose of limiting access to the sensitive data stored on the chip card. EAC comprises of terminal authentication and chip authentication. Finally Restricted Identification (RI) protocol generates a sector-specific identifier for each card, enabling the pseudonymous identification of the card-holder.

Privacy Analysis

The German eID scheme provides pseudonimity to the users and reduces sharing of excessive data with SPs. For instance, to verify if the age of the user is above 18, only a yes/no is sent instead of the age. By using secure channels, user data is not observable in transit. However the security of eID authentication relies on the tamper-resistance of the smart card chips. If an attacker manages to extract the chip authentication key from any eID card, then this attacker would be able to forge arbitrary identities. Also, user data that is transmitted after selective disclosure by the user does not contain any signature in order to verify if it is indeed the data that was originally issued by BSI and sent by a legitimate eID cardholder [1]. Only the context of the EAC protocols run and the secure channel thus established assure the eID-Server of the authenticity of the eID data. Finally, The eID-Server is recommended to be implemented by the SP but can also be implemented by third parties, thus creating a privacy risk as it serves more than one SP and can learn the visiting patterns of users as the attributes that are revealed by the users are identifying. Until 2012, eID servers implemented by only two third parties were being used by the SPs [10].

2.5 IRMA (I Reveal My Attributes)

IRMA (I Reveal My Attributes) is a attribute-based credentials (ABC) eID scheme. The credential issuer or the IDP issues credentials to the user and vouches for the validity of the attributes contained in the credential. After issuing the credentials, the IDP cannot recognise the credential as it is signed using blind signatures. This eliminates the possibility of the issuer tracking the card owner. The user can use the credentials to prove the possession of an attribute to the SP. Two important technologies that make use of an ABC approach are Microsoft's U-Prove [11] and IBM's Identity Mixer (Idemix) [12]. Idemix is built upon the concept of Camenisch-Lysyankaya signature scheme and its protocols [13]. IRMA is a partial implementation of Idemix that demonstrates the applicability of ABCs on smart cards [14]. The implementation includes privacy enhancing features of ABCs such as selective disclosure of attributes using zero-knowledge protocols.

Privacy Analysis

Users can perform transactions anonymously for transactions that do not need an identity but can be completed if the credentials are satisfactory. IRMA provides issuer

unlinkability as issuance involves creating a blind signature which conceals the resulting credential from the IDP. The user is also guaranteed that when a credential is verified multiple times by a SP, these sessions cannot be linked. By selective disclosure, the user can choose to reveal only a selection of the attributes and has greater control.

IRMA is certainly the most privacy friendly scheme discussed in this document. There still remains few issues that need to be addressed. If a card is lost or stolen, the lack of revocation procedure in the current implementation allows the possibility of misuse of credentials till it expires. A revocation scheme [15] for IRMA has been recently proposed but it introduces a new problem. The procedure uses a revocation value which is encoded by the credential such that the credential can be identified when it is revoked. Thus weakening the unlinkability argument.

Table 1: Comparison of eID schemes

Privacy Properties	eID Schemes				
	Belgian	UK	Dutch	German	IRMA
Anonymity / Pseudonimity		✓	✓	✓	✓
Data minimization				✓	✓
Unlinkability				✓	✓
Unobservability			✓	✓	✓
Transparency			✓	✓	✓

3 Privacy Enhancing Solutions

In section 2, a number of privacy issues in some of the existing/proposed eID schemes were identified. In this section, we briefly discuss two possible solutions to address the issue of privacy hotspots in federated eID systems.

3.1 Homomorphic Encryption for Privacy

Homomorphic encryption is a form of encryption which allows processing of ciphertexts and generate an encrypted result which, when decrypted, matches the result of operations performed on the plaintexts. Rivest, Adleman, and Dertouzos proposed the idea of homomorphic cryptosystems in their 1978 paper [16] and soon partially homomorphic cryptosystems such as unpadded RSA [17], ElGamal [18] and Paillier [19] were proposed. But it was only in 2009 that the first fully homomorphic encryption (FHE) scheme was proposed by Gentry [20]. The earliest FHE scheme which was based on ideal lattices was not suitable for practical implementation as it was computationally expensive and the ciphertext sizes were also very large [21]. Variants of the scheme based on different hardness assumptions such as Learning With errors (LWE) [22] or Ring-LWE [23] and integer-based or approximate Greatest Common Divisor (GCD) problem [24], also turned out to be impractical as the noise contained in the ciphertexts could not be managed after certain number of homomorphic operations and required an expensive bootstrapping step to refresh the ciphertext. But real world applications do not need to be able to perform handle all circuits. Thus a leveled homomorphic encryption scheme which can handle a circuit of low depth is sufficient. Many optimizations such as modulus switching, tensoring and re-linearization [25] have been proposed to make these schemes more practical.

The property of homomorphic encryption which allows computation on encrypted data can be utilised in federated eID schemes. In the Dutch eID scheme using polymorphic pseudonyms, the multiplicative homomorphic property of ElGamal cryptosystem is utilised to transform pseudonyms. A similar approach could be used for attributes as well. But multiplicative homomorphism may not be sufficient if we want to prevent

IDPs from having direct access to user data. Leveled homomorphic encryption schemes may find application in this scenario.

3.2 Block chain for Decentralization

Block chain was introduced by Satoshi Nakamoto as a timestamp server as part of the Bitcoin protocol [26]. A block chain is a public ledger shared by all nodes participating in a system based on the Bitcoin protocol [27]. A full copy of a block chain contains every transaction ever executed. Every block contains a hash of the previous block such that a chain of blocks is created from the first block of the chain, also known as genesis block, to the current block. This way the blocks are arranged in chronological order. It is also computationally infeasible to modify a block as every block that follows must also be regenerated.

Block chain allows to eliminate the necessity of a central party. But it introduces additional issues. A public ledger provides everyone in the network access to the data on the block chain. This means that instead of one central party having access to all the data, now all parties in the network have access. So block chain as it is cannot be used to store private data and hence does not address the problem of privacy but merely shifts the problem. Even though Bitcoin protocol allows the usage of multiple pseudonyms or public keys, it is possible to link the activity of users [28]. Enigma [29], a decentralized computation platform which allows storage and computation of private data, uses off-chains in addition to the block chain to store private data. A similar approach has been used in [30] but it requires a minimally trusted manager.

As a primitive idea we propose that block chain could be used in federated eID systems by using off-chain consisting of IDPs. User data can be split up among the IDPs thus preventing any one IDP from having complete information about its users. Another possibility is to use a private block chain in which the IDPs take up the role of miners. These ideas require further research and refinement before they can be considered for nation-scale eID systems.

4 Conclusion

We have surveyed five eID schemes in the EU and analysed their privacy properties. Belgian, GOV.UK and Dutch schemes use the approach of identifying entities by a unique number at the IDP. This approach is convenient for bookkeeping but it is not privacy friendly. Unique identifiers can be used to trace the user and her/his activities. IDPs have been identified as a privacy hotspot due to the large trove of user data that they store and process which allows them to link user activities. Finally, we proposed two possible solutions - homomorphic encryption and block chain - to allow the IDP to perform its role of authenticating users to SPs without becoming privacy hotspots.

References

[1] A. Poller, U. Waldmann, S. Vowe, and S. Türpe, "Electronic identity cards for user authentication - promise and practice," *IEEE Security & Privacy*, vol. 10, no. 1, pp. 46–54, 2012.

[2] L. Floridi, "The ontological interpretation of informational privacy," *Ethics and Information Technology*, vol. 7, no. 4, pp. 185–200, 2005.

[3] ISO15408-2:2005, "Information technology - security techniques - evaluation criteria for it security - part 2: Security functional requirements," tech. rep., International Standard Organization, 2005.

[4] D. D. Cock, C. Wolf, and B. Preneel, "The belgian electronic identity card (overview)," in *Sicherheit*, vol. 77 of *LNI*, pp. 298–301, GI, 2006.

[5] P. Verhaeghe, J. Lapon, B. D. Decker, V. Naessens, and K. Verslype, "Security and privacy improvements for the belgian eid technology," in *SEC*, vol. 297 of *IFIP Advances in Information and Communication Technology*, pp. 237–247, Springer, 2009.

[6] "Identity assurance principles," tech. rep. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/361496/PCAG_IDA_Principles_3.1__4_.pdf.

[7] L. T. A. N. Brandão, N. Christin, G. Danezis, and anonymous, "Toward mending two nation-scale brokered identification systems," *PoPETs*, vol. 2015, no. 2, pp. 135–155, 2015.

[8] "Polymorphic pseudonymization." Available at https://www.idensys.nl/fileadmin/bestanden/idensys/documenten/basisdocumentatie/documentatieset/PP_Scheme_091.pdf.

[9] BSI, "Advanced security mechanisms for machine readable travel documents and eIDAS token part 2 protocols for electronic identification, authentication and trust services (eIDAS)," tr-03110-2, Bundesamt fur Sicherheit in der Informationstechnik, 2015.

[10] R. Bjones, I. Krontiris, P. Paillier, and K. Rannenberg, "Integrating anonymous credentials with eids for privacy-respecting online authentication," in *APF*, vol. 8319 of *Lecture Notes in Computer Science*, pp. 111–124, Springer, 2012.

[11] S. Brands, *Rethinking Public Key Infrastructures and Digital Certificates: Building in Privacy*. MIT Press, 2000.

[12] IBM Research Zurich Security team, "Specification of the identity mixer cryptographic library," tech. rep., IBM Research, April 2010.

[13] J. Camenisch and A. Lysyanskaya, "An efficient system for non-transferable anonymous credentials with optional anonymity revocation," in *EUROCRYPT*, vol. 2045 of *Lecture Notes in Computer Science*, pp. 93–118, Springer, 2001.

[14] P. Vullers and G. Alpár, "Efficient selective disclosure on smart cards using idemix," in *IDMAN*, vol. 396 of *IFIP Advances in Information and Communication Technology*, pp. 53–67, Springer, 2013.

[15] W. Lueks, G. Alpár, J. Hoepman, and P. Vullers, "Fast revocation of attribute-based credentials for both users and verifiers," in *SEC*, vol. 455 of *IFIP Advances in Information and Communication Technology*, pp. 463–478, Springer, 2015.

[16] R. L. Rivest, L. Adleman, and M. L. Dertouzos, "On data banks and privacy homomorphisms," in *Foundations of Secure Computation* (R. A. DeMillo, D. P. Dobkin, A. K. Jones, and R. J. Lipton, eds.), pp. 165–179, Academic Press.

[17] R. L. Rivest, A. Shamir, and L. M. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Commun. ACM*, vol. 21, no. 2, pp. 120–126, 1978.

[18] T. E. Gamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," in *CRYPTO*, vol. 196 of *Lecture Notes in Computer Science*, pp. 10–18, Springer, 1984.

[19] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *EUROCRYPT*, vol. 1592 of *Lecture Notes in Computer Science*, pp. 223–238, Springer, 1999.

[20] C. Gentry, *A fully homomorphic encryption scheme*. PhD thesis, Stanford University, 2009.

[21] C. Gentry and S. Halevi, "Implementing gentry's fully-homomorphic encryption scheme," in *EUROCRYPT*, vol. 6632 of *Lecture Notes in Computer Science*, pp. 129–148, Springer, 2011.

[22] Z. Brakerski and V. Vaikuntanathan, "Efficient fully homomorphic encryption from (standard) LWE," in *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011*, pp. 97–106.

[23] Z. Brakerski and V. Vaikuntanathan, "Fully homomorphic encryption from ring-lwe and security for key dependent messages," in *CRYPTO*, vol. 6841 of *Lecture Notes in Computer Science*, pp. 505–524, Springer, 2011.

[24] M. van Dijk, C. Gentry, S. Halevi, and V. Vaikuntanathan, "Fully homomorphic encryption over the integers," in *EUROCRYPT*, vol. 6110 of *Lecture Notes in Computer Science*, pp. 24–43, Springer, 2010.

[25] Z. Brakerski, C. Gentry, and V. Vaikuntanathan, "(leveled) fully homomorphic encryption without bootstrapping," 2012.

[26] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008. https://bitcoin.org/bitcoin.pdf.

[27] "Blockchain." https://en.bitcoin.it/wiki/Block_chain.

[28] E. Androulaki, G. Karame, M. Roeschlin, T. Scherer, and S. Capkun, "Evaluating user privacy in bitcoin," in *Financial Cryptography*, vol. 7859 of *Lecture Notes in Computer Science*, pp. 34–51, Springer, 2013.

[29] G. Zyskind, O. Nathan, and A. Pentland, "Enigma: Decentralized computation platform with guaranteed privacy," *CoRR*, vol. abs/1506.03471, 2015.

[30] A. E. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts," *IACR Cryptology ePrint Archive*, vol. 2015, p. 675, 2015.