ON PRIVACY IN BUSINESS PROCESSES
Observing Delegation of Personal Data by using Digital Watermarking

Sven Wohlgemuth, Isao Echizen, Noboru Sonehara
National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan

Günter Müller
Institute of Computer Science and Social Studies, Albert-Ludwigs University of Freiburg, Germany

Keywords: Privacy, Business processes, Usage control.

Abstract: Privacy in business processes for personalized services is currently a matter of trust. Business processes require the delegation of personal data and users are not able to control their delegation and so their usage. Existing privacy-enhancing technologies consider access control but not usage control of personal data. Current work on usage control considers especially formalization of usage rules, so called obligations, and their enforcement by using mechanisms of Digital Rights Management, secure logging of access requests for an ex post enforcement, and the non-linkable delegation of access rights to personal data. However, either these enforcement mechanisms do not consider a delegation of personal data or they assume trustworthy data consumers or data providers respectively. We investigate on digital watermarking in order to observe the enforcement of obligations for a delegation of personal data without mandatory trusting in service providers.

1 INTRODUCTION

Business processes for personalized services require the collection and delegation of personal data. Depending on the purpose of the data usage, service providers act as a data consumer (DC) or as a data provider (DP). As a data consumer they collect, use, and save personal data. As a data provider they delegate personal data to other service providers. An example for such a change of roles are customer loyalty programs (Customer Relationship Management - CRM). The challenge faced is whether the requirements of the European data protection directive (European Commission, 1995) and the Japanese Act on the Protection of Personal Information (Japanese Government, 2005) are fulfilled so that users are able to enforce respectively control the enforcement of the agreed rules for using personal data.[1]

Anonymity services and identity management systems focus on the collection of personal data and so on the access to it. Identity management systems protect users against an undesired profiling by using pseudonymity. If they are applied on a delegation of personal data, users will lose the control on their personal data (Wohlgemuth and Müller, 2006). Current work on usage control focuses either on the formalization of obligations (Hilty et al., 2005), on enforcement of obligations by using mechanisms of Digital Rights Management (DRM)(Pretschner et al., 2008), on audits concerning the usage of personal data by secure logging (Accorsi, 2007), and on protocols for a non-linkable delegation of rights in order to get access to personal data at service providers (Wohlgemuth and Müller, 2006). However, DRM mechanisms do not consider delegation (Rosenblatt et al., 2001), secure logging does not consider delegation when authorized data consumers have got access to personal data, and non-linkable delegation of rights assumes trustworthy data providers in order to enforce delegated access rights of service providers as data consumers.

Our contribution is a scheme for observing delegations of personal data. This scheme makes use of digital watermarking but without the need of trustworthy service providers, i.e. data providers. Since this scheme abstracts from watermarking algorithms

[1] Concerning a delegation of personal data, the corresponding user has to give his agreement in advance. The data provider has to inform the user before a delegation about the personal data to be delegated, the purpose of the delegation, and about the recipients.

Wohlgemuth S., Echizen I., Sonehara N. and Müller G. (2009).
ON PRIVACY IN BUSINESS PROCESSES - Observing Delegation of Personal Data by using Digital Watermarking.
In *Proceedings of the International Conference on e-Business*, pages 27-32
DOI: 10.5220/0002260600270032

by focussing on the model for embedding and detecting watermarks, it is suitable for any medium format and presents thereby also an approach for enhancing DRM to a controlled delegation of digital content.

Section 2 describes a privacy threat in case of non-compliant data providers in the case study of CRM. Section 3 presents our approach for observing a delegation of personal data. Section 4 investigates on the use of digital watermarking for our scheme and presents the conceptional weaknesses of digital watermarking concerning mandatory trustworthy service providers. Section 5 presents our scheme *DETECTIVE* for using watermarking to observe delegations of personal data. Section 6 reports on related work. Section 7 concludes the result of our work so far.

2 DELEGATION OF PERSONAL DATA IN CRM

An instance of CRM are loyalty programs. Participating users get discounts on services if they use their loyalty card at the service providers of this loyalty program. At the same time, these service providers collects personal data about their users, e.g. customer number, goods or services, payment, time, location, and amount of discount by means of loyalty points, and delegates this data to the provider of the loyalty program. A loyalty provider offers on behalf of the participating service providers personalized services and advertisements to the users and manages their discounts.

In practice, loyalty program providers publish their privacy policy as part of their general terms and conditions.[2] If users want to participate in a loyalty program, they have to accept their general terms and conditions and thereby give loyalty program providers full authority to process their personal data. A loyalty program provider collects personal data of users whenever they use their loyalty card while buying goods and services of the participating service providers. We assume that a user and the loyalty program provider have agreed on a privacy policy and so obligations for delegations of his personal data *d* to *service provider 2* and *d'* to *service provider 3*. Additional delegations of *d* and *d'* are not allowed.

A violation of these obligations has occurred if personal data has been delegated to non-authorized service providers. Figure 1 shows an exemplary flow of personal data according to the model of (Pretschner et al., 2006) and two violations. The first violation stems from the loyalty program provider, since he has delegated *d'* in addition to *d* to *service provider 2*. The second violation stems from *service provider 3*, since he has delegated *d'* to *service provider 2*. Also, the second delegation may happen indirectly by linking user's transactions, e.g. if the user uses the same identifier in several transactions with *service provider 2* and *service provider 3*. The challenge is to detect the permitted information flow and to identify untrustworthy data providers.

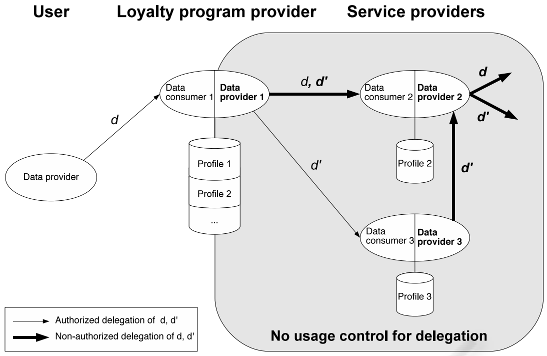

Figure 1: Delegation of personal data in contrary to obligations.

3 OBSERVING DELEGATION OF PERSONAL DATA

Since it is unlikely that service providers will give users some control on their information systems, we treat them as a *black box*. Our approach considers the fact that personal data has been delegated and focuses on the observation of delegations of given personal data between data providers and data consumers. In order to identify these participants, all delegations of given personal data should be *traceable*. Secondly, in order to prevent indirect information flows, users' transactions should be *non-linkable* as far as obligations do not consider the delegation of identifying data. *Traceability* in this context means that an information flow should be uniquely mapped to a data provider, data consumer and the corresponding user. Data providers should not be able to repudiate a delegation but they should be able to prove that they have not delegated given personal data.

We propose to tag an information flow between two parties and to get a proof for data providers and data consumers concerning the delegation and receipt of given personal data. The tag for personal data *d* consists of data provider's identity ID_{DP}, data consumer's identity ID_{DC}, the corresponding user's iden-

[2] cf. the privacy policies of *PAYBACK* at http://www.payback.de and *Miles & More* at http://www.milesandmore.com

tity ID_{User}, and of a $link_{obligations}$ to the agreed obligations: $tag = (ID_{DP}, ID_{DC}, ID_{User}, link_{obligations})$. The obligations are indirectly part of a tag by a link to them ($link_{obligations}$), since they should be modifiable due to a change of the business process or an exchange of the authorized service providers. The tag should be *sticky* to d similar to (Karjoth et al., 2002) so that $d^* = (d, tag)$ is going to be delegated while assuring the integrity of d^*. If d^* is going to be delegated further in compliant to the obligations, the tag has to be updated by replacing ID_{DP} with the identity of the former data consumer and by adding the new data consumer $ID_{DC'}$. A sequence of tags for the same personal data describes an information flow by a delegation chain. In the following we apply symmetric and asymmetric digital watermarking schemes on the scenario in order to find out their suitability for our approach.

4 DIGITAL WATERMARKING AND DELEGATION

Digital watermarking aims for detecting unauthorized copies of digital content (Cox et al., 2008). Recent approaches consider watermarks for text, e.g. by making use of the structure of XML messages (Zhou et al., 2005). The main characteristic of digital watermarking schemes is the use of a symmetric watermarking key in order to produce *noise* to embed watermarks in (Cox et al., 2008). If one knows this key and the watermarking algorithm, he is able to embed, detect, and remove watermarks. Since neither data providers, data consumers nor users should be able to forge a digital watermark, e.g. in order to obscure a delegation, the watermarking algorithms and the watermarking key are to be kept secret from them. For this reason, a Trusted Third Party (TTP) has to be introduced in order to embed and detect digital watermarks so that solely the TTP knows these secrets.

Figure 2 shows the application of digital watermarking in combination with the use of a TTP. For simplicity, user's profiles are not shown. The user has already disclosed $d*$ to *service provider 1* with the obligation that he is allowed to delegate d' to *service provider 3*. According to the symmetric scheme, *service provider 1* requests the TTP to embed the tag for the authorized delegation and sends ID_{User} from user's authentication, $ID_{DP} = ID_{SP\_1}$ and d^*. The TTP gets $ID_{DC} = ID_{SP\_3}$ and $link_{obligations}$ from *service provider 3*. The TTP returns the embedded tag by d'^* to *service provider 1* who redirects it to *service provider 3*. Thus both service providers get d'^*. If one of them delegates it to *service provider 2*, neither user nor arbiter can decide whether *service provider 1* or *service provider 3* has violated this obligation. Also, service providers have got no proof that they have not violated the obligations. In addition, every participant has to trust the TTP that she will embed and detect digital watermarks according to her policy.

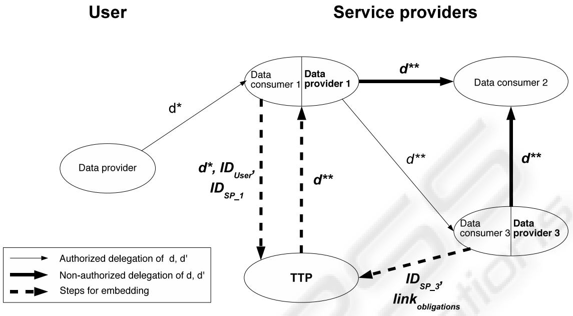

Figure 2: Applying digital watermarking on delegation of personal data *d*.

Asymmetric fingerprinting (Pfitzmann and Schunter, 1996) solves this problem of undecidability. In principle, asymmetric fingerprinting combines a digital watermarking scheme with cryptographic commitments and the digital signature of the data consumer. Data providers embed the watermarks consisting of a random ID_{DC} chosen by data consumers together with a text, here the obligations. The protocol of (Pfitzmann and Schunter, 1996) assures by using commitments that only data consumers get the watermark. The obligations are signed by data consumers and sent to the data provider. This is the proof of the data provider that the given data consumer will get this data with the watermark.

However, asymmetric fingerprinting assumes conflicting interests between providers and consumers. This contradicts with our trust model, i.e. service providers may violate user's obligations, since they have an interest to collude. A solution is again the introduction of a TTP who checks whether service providers as provider and consumer have run the asymmetric fingerprinting protocol as expected by verifying the results of the protocol.

5 *DETECTIVE*: DELEGATION OF PERSONAL DATA WITH DIGITAL WATERMARKS

With *DETECTIVE*, we propose a modification for the asymmetric fingerprinting scheme so that it can be used without a TTP. Concerning the embedding and verifying of digital watermarks, the main difference of *DETECTIVE* to the scheme of (Pfitzmann and Schunter, 1996) is the integration of users' agreement by means of delegated access rights to the requested personal data. Thus our scheme should generate only valid digital watermarks if the delegated access rights of data consumers to the corresponding personal data are used. We modify the protocols of (Pfitzmann and Schunter, 1996) and combine them with our previous work: the *DREISAM* protocols for a non-linkable delegation of rights (Wohlgemuth and Müller, 2006). The outcome of a protocol run of *DREISAM* is an anonymous one-show credential for the data consumer which is linked to his identity which is represented by his cryptographic secret key k_{DC}.(Camenisch and Lysyanskaya, 2001) An anonymous credential consists of user's access right to personal data stored by a data provider and of his obligations for a further delegation of personal data. If a data consumer wants to get access on this personal data, he will show his anonymous credential to the given service provider as the data provider. The messages of a showing protocol's run are written in a transcript which is used for solving disputes and deanonymizing a dishonest participant.(Camenisch and Lysyanskaya, 2001)

Our scheme consists of three protocols: *init*, *tag*, and *verify*. The *init* protocol generates the cryptographic key pair (pk_{DC}, sk_{DC}) of the data consumer if necessary and broadcasts it via an authentic channel, e.g. a public-key infrastructure. If the service provider acting as a data consumer has not got rights from the user, this service provider starts one run of the *DREISAM* protocol with the results $anonymousCredential(rights, DC, DP, obligations, CA)$ and $transcript(DP, DC, anonymousCredential)$. If the data has not been disclosed, the user discloses it to the requesting service providers and embeds a watermark by running the *tag* protocol.

The *tag* protocol involves two service providers, one as a data provider and the other as a data consumer. The data provider uses a symmetric watermarking scheme in order to embed the identity of the data consumer together with user's identity in his personal data d. In order to distinguish between the service provider as data provider and the service provider as data consumer, the data provider uses the identity of the data consumer k_{DC} as cryptographic commitments, embeds them in the personal data, and computes a commitment of the watermark. This commitment is not the resulting watermark. Otherwise the service provider who is acting as the data provider would also have it. Therefore, the data consumer opens his commitments. The result is the product of data provider's commitment with the opened data consumer's commitments. This is the watermark of the personal data. On the other side, the data provider needs a confirmation that he has got the identity of the data consumer, i.e. his commitments, and the rights which the data consumer has got from the user. This requirement of non-repudiation by the data consumer is fulfilled by the digital signature on the transcript and the commitments concerning k_{DC}. After getting this signature, the data provider will compute the commitment of the embedded data and send it as a commitment to the data consumer. Figure 3 shows the tagging function.

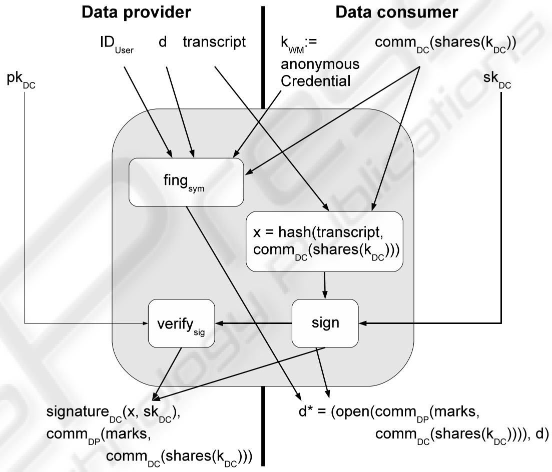
Data provider
Data consumer
ID_User
d
transcript
k_WM:= anonymous Credential
comm_DC(shares(k_DC))
pk_DC
sk_DC
fing_sym
x = hash(transcript, comm_DC(shares(k_DC)))
verify_sig
sign
signature_DC(x, sk_DC), comm_DP(marks, comm_DC(shares(k_DC)))
d* = (open(comm_DP(marks, comm_DC(shares(k_DC)))), d)

Figure 3: Tagging personal data with delegated rights based on (Pfitzmann and Schunter, 1996; Wohlgemuth and Müller, 2006).

The *tag* protocol is as follows:

1. The data consumer chooses the right which the user has delegated to him by means of $anonymousCredential(rights, DC, DP, obligations, CA)$ as the watermarking key k_{WM} for a given watermarking algorithm and sends $transcript(DP, DC, anonymousCredential)$ to the data provider.
2. The data consumer sends his identity k_{DC} by means of commitments $comm_{DC}(shares(k_{DC}))$ to the data provider. The data provider cannot use k_{DC}, e.g. for impersonation the data consumer. The data consumer sends the commitments of $l-1$ shares to the data provider,[3] so that k_{DC} will be embedded at various places in d. The data consumer shows the correctness of his commitments and their relationship to the shares of k_{DC} by using the verifiable secret sharing scheme of (Pedersen, 1992).

[3] $l < n$ shares are at least necessary to reconstruct the key(Pedersen, 1992).

3. The data provider embeds $comm_{DC}(shares(k_{DC}))$ in the representation d and stores these places as marks. Afterwards, he computes a commitment $comm_{DP}(marks,\ comm_{DC}(shares(k_{DC})))$. The result is the tagged personal data as the commitment of the data provider.

4. The data consumer computes the hash value of $transcript(DP, DC, anonymousCredential)$ and his identity ($comm_{DC}(shares(k_{DC}))$). Afterwards, he signs the hash value with his signature key sk_{DC} and sends it with the digital signature to the data provider.

5. The data provider verifies the digital signature. If it is valid, he saves it with the commitments of the data consumer. Then he sends his commitment of the watermark to the data consumer.

6. The data consumer opens data provider's commitments. This results is the valid watermark of d. Since only the data consumer is able to open his commitments, the data provider has not got this watermark.

The *enforcement* of running the *tag* protocol depends on the correctness of the verification. Dishonest service providers who either does not embed a watermark or modify respectively remove a watermark must be identified by running the *verify* protocol. Either the user or an arbiter starts the *verify* protocol. The assumption is that one of those has found personal data and want to check whether the data consumer where this data has been found has the authorization to use it. At the beginning of the *verify* protocol, they re-construct the delegation chain of personal data under investigation by delegated access rights. If the user has delegated the corresponding rights to this service provider, then there is no violation of the obligation and the verification is finished. Otherwise the obligation has violated and the aim of the *verify* protocol is to identify the dishonest service provider who has non-authorized acted as a data provider.

The user or the arbiter extract all watermarks of the found personal data by using the anonymous credentials of the delegation chain as the watermarking keys. By mapping these watermarks to the delegation chain via the anonymous credentials, the verifier knows the last data consumers of the chain. In the worst case, this is the watermark which has been created between the user and the first service provider by the *init* protocol. If one of these watermarks is identical to the found one, the user request the the digital signature of the data consumer from the corresponding data provider and checks it. In the next step, the user requests the data consumer to open his commitments. The verifier compares these opened commitments with those found in the watermark. If the digital signature and the commitments of the data consumer are correct, then the data consumer has violated the obligation. Otherwise, the data provider has violated the obligations.

6 RELATED WORK

Current work on enforcement of a policy for information flows concentrate on formal methods (Mantel, 2001) or on encryption (Casassa Mont and Pearson, 2005). However formal methods consider an information flow via covert channels or an indirect path from a data provider to a data consumer. In addition, a corresponding verification of a system implies that this system doesn't change afterwards. Otherwise, it has to be verified again.

Obligations for a delegation of personal data are realized by sticky policies (Karjoth et al., 2002). An implementation of sticky policies for delegation of personal data is the *adaptive privacy management system (Adaptive PMS)* by (Casassa Mont and Pearson, 2005). Sticky policies are linked to certain personal data at the time of their collection by an encryption scheme. A data consumer will get the decryption key from a TTP, if he is authorized by the sticky policy. After the decryption of the personal data, these data consumers are able to delegate the decrypted personal data further.

7 CONCLUSIONS

We have shown that using digital watermarking and asymmetric fingerprinting schemes for observing a delegation of personal data results in the introduction of a TTP. By modifying the asymmetric fingerprinting scheme by adding data consumer's rights to get this personal data, we have proposed a scheme where users will be able to control the delegation of personal data according to their delegated rights. Since the suitability of watermarking schemes depends on the existence of watermarking algorithms, our scheme may also be used for personal data which is represented by images, e.g. x-ray images in case of electronic health records (EHR). The model of EHR services is the same as the model of CRM. Especially the providers of EHR should not be seen as trustworthy in contrary to doctors or nursery services. Current technical infrastructures and services for EHR, e.g. for the German electronic health card (gematik - Gesellschaft für Telematikanwendungen der Gesundheitskarte mbH, 2008), uses encryption of patient data in order to protect them against non-authorized use by the EHR provider. Consequently, the EHR provider cannot use this data and so cannot offer additional services.

In the next step we will specify the cryptographic protocols of our scheme and apply a proof-of-concept implementation on the use case of EHR. The proof-of-concept implementation of *DETECTIVE* will be evaluated against attacks on the protocol layer, i.e. on the tagging and detecting protocol. Attacks will be derived from the German IT Baseline Protection Catalogue and from the Japanese Act on the Protection of Personal Information. Secondly, the evaluation will consider the economic requirements of services based on electronic health records and investigate on its feasibility for personal data as, e.g., x-ray images in the EHR scenario.

ACKNOWLEDGEMENTS

This work has been funded by the Japan Society for the Promotion of Science (JSPS). The cooperative work between the University of Freiburg and the National Institute of Informatics has been done within their Memorandum of Understanding.

REFERENCES

Accorsi, R. (2007). Automated Privacy Audits to Complement the Notion of Control for Identity Management. In *Policies and Research in Identity Management*. IFIP.

Camenisch, J. and Lysyanskaya, A. (2001). Efficient non-transferable anonymous multi-show credential system with optional anonymity revocation. In *EUROCRYPT 2001*, volume 2045 of *Lecture Notes in Computer Science*, pages 91–118. Springer.

Casassa Mont, M. and Pearson, S. (2005). An Adaptive Privacy Management System for Data Repositories. In Kazikas, S., Lopez, J., and Pernul, G., editors, *TrustBus 2005*, volume 3592 of *Lectures Notes in Computer Science*, pages 236–245, Heidelberg. Springer.

Cox, I. J., Miller, M. L., Bloom, J. A., Fridrich, J., and Kalker, T. (2008). *Digital Watermarking and Steganography*. Morgan Kaufmann.

European Commission (1995). Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. *Official Journal of the European Communities*, L 281(395L0046):31–50.

gematik - Gesellschaft für Telematikanwendungen der Gesundheitskarte mbH (2008). Übergreifendes Datenschutzkonzept der Gesundheitstelematik Version 0.9.0. http://www.gematik.de/upload/gematik_DS_Datenschutzkonzept_V0.9.0_3803.pdf.

Hilty, M., Basin, D., and Pretschner, A. (2005). On Obligations. In de Capitani di Vimercati, S., Syverson, P., and Gollmann, D., editors, *10th European Symposium on Research in Computer Security (ESORICS 2005)*, volume 3679 of *Lecture Notes in Computer Science*, pages 98–117. Springer.

Japanese Government (2005). Act on the Protection of Personal Information. http://www5.cao.go.jp/seikatsu/kojin/foreign/act.pdf,.

Karjoth, G., Schunter, M., and Waidner, M. (2002). Privacy-enabled Services for Enterprises. In *International Workshop on Trust and Privacy in Digital Business (Trustbus 2002)*, pages 483–487.

Mantel, H. (2001). Information Flow Control – Bridging a Gap. In *FME 2001*, volume 2021 of *Lecture Notes in Computer Science*, pages 153–172. Springer.

Pedersen, T. P. (1992). Non-Interactive and Information-Theoretic Secure Verifiable Secret Sharing. In *CRYPTO 1991*, volume 576 of *Lecture Notes in Computer Science*, pages 129–140. Springer.

Pfitzmann, B. and Schunter, M. (1996). Asymmetric Fingerprinting. In *Eurocrypt 1996*, volume 1070 of *Lecture Notes in Computer Science*, pages 84–95. Springer.

Pretschner, A., Hilty, M., and Basin, D. (2006). Distributed Usage Control. *Communications of the ACM*, 49(9):39–44.

Pretschner, A., Hilty, M., Schütz, F., Schaefer, C., and Walter, T. (2008). Usage Control Enforcement: Present and Future. *IEEE Security and Privacy*, 6(4):44–53.

Rosenblatt, B., Trippe, B., and Mooney, S. (2001). *Digital Rights Management: Business and Technology*. John Wiley & Sons.

Wohlgemuth, S. and Müller, G. (2006). Privacy with Delegation of Rights by Identity Management. In *International Conference on Emerging Trends in Information and Communication Security (ETRICS 2006)*, volume 3995 of *Lectures Notes in Computer Science*, pages 175–190. Springer.

Zhou, X., Pang, H., Tan, K., and Mangla, D. (2005). WmXML: A System for Watermarking XML Data. In *Proceedings of the 31st international conference on Very Large Data Bases*, pages 1138–1321.