Towards Privacy in Personal Data Management

P.S. Efraimidis G. Drosatos F. Nalbadis A. Tasidou
Department of Electrical and Computer Engineering
Democritus University of Thrace
Vas. Sophias 12, 67100 Xanthi, Greece
{pefraimi, gdrosato, fnalmpan, atasidou}@ee.duth.gr

Abstract

We present a personal data management framework called Polis, which abides by the following principle: Every individual has absolute control over her personal data, which reside only at her own side. Preliminary results indicate that beyond the apparent advantages of such an environment for users' privacy, everyday transactions remain both feasible and straightforward.

1 Introduction

As the use of computers and the Internet becomes more popular by the minute, the issue of protecting one's personal data is more essential than ever. The way electronic transactions are conducted nowadays, makes it necessary for the customer to give away her personal data to the service provider and hope that the latter will not use them in a malicious way. In order to protect personal information, several organizations and countries have issued privacy regulations, which should be followed in order for personal information to be protected; the collectively referred to as Fair Information Practices (FIP). Examples of important FIP regulation frameworks are the Data Protection Directive 95/46/EC (henceforth referred to as The Directive) and follow-ups like the Canadian PIPEDA and UK's Data Protection Act (DPA).

In this work, we assert that it is feasible for electronic transactions to work, whilst personal data resides at the individuals' side. To support this claim, we design, build and evaluate the prototype system Polis, which implements the above principle. We show that Polis can satisfy important data protection principles in a natural and efficient way and describe how Polis can be integrated into online transactions to manage personal data. Preliminary results indicate that the Polis approach can lead to a simple, scalable solution that can be beneficial to both individuals and service providers.

Related Work. Laudon in [22] discusses the idea that individuals should own their personal information themselves and decide how this information is used. A point made in [27] is that, although considering personal data the owner's private property is a very appealing idea, it would be rather difficult to practically apply it and legally enforce it. Our approach proposes an idea that has the same practical effect as considering personal data the owner's private property, but withdraws the legal objections involved with this idea. The economic aspects of privacy are examined in [28] where Varian argues: "It is worth observing that the Fair Information Practices principles would automatically be implemented if the property rights in individual information resided solely with those individuals". The argument that personal data would be safer at the user's side is also examined in [24].

Different kinds of frameworks that are related to personal data have recently been proposed or are in progress. In particular, privacy sensitive management of personal data in ubiquitous computing is discussed in [14], storing personal data in an individual's mobile device is examined in [15]. Protecting personal data that is stored within a company is considered in [26, 17]. More related to Polis is a rich framework for privacy protection, proposed in [23]. This framework is more complicated than Polis and is built on the principle that personal data is kept inside a "Discreet Box", located at the service provider's side. Other results in this field, less related to Polis, can be found in [10, 3]. General surveys on privacy enhancing technologies are given in [11, 13]. To our knowledge, Polis is the first general framework for managing personal data only at the owner's side.

Outline. The rest of this work is organized in the following way: The Polis approach is described in Section 2. In Section 3, the possible applications for Polis are discussed. The Polis prototype is presented and evaluated in Section 4. A final discussion is given in Section 5. Additional figures and snapshots can be found in the Appendix.

978-0-7695-3323-0/08 $25.00 © 2008 IEEE
DOI 10.1109/PCI.2008.11

2 The Polis approach

The Polis approach is based on the following principle:

"Polis-users are prohibited from storing any personal data but their own."

Polis is meant to be employed by privacy concerned internet users which fulfill the requirements of having:

- A reliable, always-on access to the Internet, in order for her agent to be always accessible.
- A certificate from an approved Certification Authority.

2.1 Polis concepts and architecture

At this point we consider it necessary to introduce a few terms that will be used in the rest of this work:

- In Polis, personal data refers to primitive personal information of individuals like name, birth date, address, etc. Personal data corresponds to what is called *off-line identity* in [1]. Our focus is on privacy-enhanced management of the off-line identity.
- An individual Internet user is a potential customer who can purchase either goods or services. This user can be called *individual*, *customer* or *data subject (according to The Directive)*. We will use the terms *individual* and *customer*, interchangeably.
- An entity that provides the aforementioned goods or services can be called *shop*, *company*, *service provider* or *data controller (The Directive)*. We will use the terms *shop*, *company* and *service providers*.
- Both individuals and companies can become *Polis-users*.

Every Polis-user is represented by a dedicated entity. This entity can be used to instantiate a corresponding Polis-agent, which is the main architectural component of Polis. The agent is used to manage the personal data of the entity and to provide controlled access to it. Service providers use the agent to retrieve personal data from affiliated users. The general architecture of Polis, as well as the constituents of a customer agent and a shop agent are presented in Figure 1.

We would like to emphasize the following characteristics of the Polis architecture:

- From the service provider's point of view, Polis provides a decentralized approach for the storage and management of personal data.
- On the contrary, from the customer's point of view, Polis is a fully centralized system, in the sense that personal data is located and managed locally by the owner's agent.

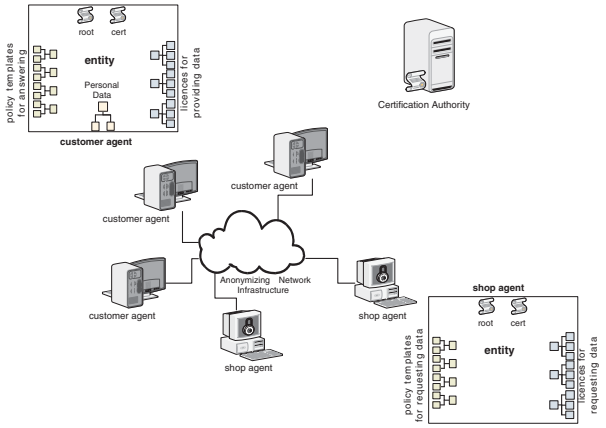

Figure 1. The Polis architecture

2.2 Schemes for personal data and policies

Critical components for a personal data management framework like Polis are the schemes for representing personal data and policies. Some known schemes for personal data are P3P [29] and CPExchange [4]. Approaches for policies related to personal data are also discussed in [16, 17], while work on personal data and policy schemes is in progress in [7].

We currently use schemes that are simple, yet powerful enough, for the current needs of the Polis prototype. Examples of a personal data scheme and a policy, as used in Polis, are shown in the long version of the paper [8]. There are eight general categories of personal data in Polis, organized hierarchically, namely Name, BDate, Cert, Skill, Characteristic, Home-Info, Business-Info and CreditCard. Each of them has one or more subcategories. The terminology used is based on P3P for the user information part, with the addition of the financial information (CreditCard) taken from CPExchange, plus the extra personal information fields (Skill and Characteristic). Each entity stores its personal data in a local XML document.

The components of a policy are the following:

- *Principals*: The Polis-entities.
- *Data*: Every single item of a user's (Polis-entity) personal data.
- *Purposes*: The set of purposes that entitle principals to retrieve data.
- *Usage restrictions*: Additional restrictions exist that limit access rights to a specific number of accesses or a specific time interval, or both.

Other important concepts of Polis are the licence and the contract. A *licence* comprises of the data involved, the valid

purposes that allow data retrieval, as well as the rules to provide either full or restricted access. The use of licences to protect personal data is discussed in [5, 21, 9]. A *contract* concerns two principals and an arbitrary set of licences. An agent can sign any number of contracts with an arbitrary number of entities.

2.3 Incentives and Objections

The fact that a Polis-user's personal data must be retrieved from the owner's side every time it is needed, automatically fulfills many critical requirements found in privacy regulations. A detailed discussion of possible incentives and objections for Polis can be found in the long version of the paper [8]

2.4 Polis in common transactions

Let us describe a few common transactions that can take place in the Polis-world, in order to display its functionality. **Online shopping** can work effortlessly in Polis while protecting the user's privacy. When a user has to fill in an order form with her personal data, she instead provides the contact details for her agent. The agents of the shop and the customer establish an agreement. A successful agreement grants access to the customer's private data only for the data items and the amount of time needed to complete the order. In Figure 2 the procedure of a Polis-transaction with an e-shop is described.

Registrations can similarly take place in Polis without the individual having to give away any personal data. Instead, only the username and password fields have to be filled in, while the agent handles the rest of the information. According to the agreed privacy policy, the website will have access to the necessary information for specific purposes only.

In addition to the above examples, Polis can potentially be employed in other applications that involve personal data, like identity management systems [25] and e-government platforms. The need for privacy protection in e-business applications is stressed in [18]. The ease of employing Polis lies in the fact that it can work as middleware, which takes care of the personal data exchange between parties in higher level applications.

2.5 Enforcement and Detection

An important aspect of every (electronic) contract is the ability to verify and enforce that the parties will not violate its terms. Polis can handle detectable privacy breaches, i.e., breaches for which data released to the shop finds its way back to the individual who submitted that information [12]. In this case a Polis compliant shop must be able to present evidence that those data were rightfully obtained for the specific purpose, at the specific time, using data licences [5]. A more challenging task would be to detect Polis-shops that leak customer's personal information. A relevant problem is discussed in [12].

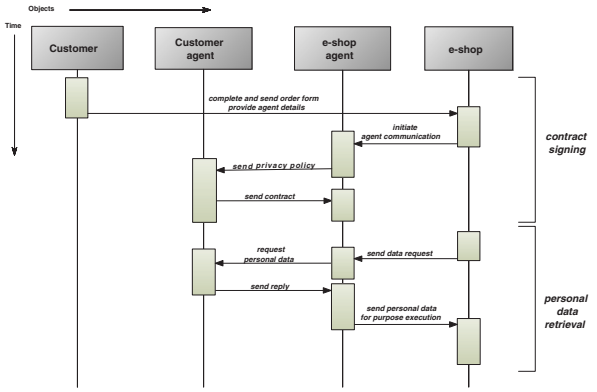

Figure 2. Polis in transactions with an e-shop

Due to the very nature of personal data, it seems that once a service provider possesses some data, there is no technically feasible way for absolute abuse prevention. Consequently, apart from technical measures, we will have to rely on market, legal and social dynamics for handling personal data properly ([12, 2] and [14, Section 5.8.5]).

As far as violations from the user's side are concerned, if the terms of an agreement are violated and the individual refuses to fulfil her contract defined obligation of providing personal information, then the service provider can use her customer-signed licence to prove entitlement to access the data.

3 Prospective Applications for Polis

An infrastructure like Polis can be a realistic step in the direction of effectively controlling personal information. Apart from the direct gains of using Polis in every day electronic transactions, there are some interesting possibilities for higher level applications that could utilize it.

Microtrades and Information Markets. The financial aspects of privacy are studied in several works like [22, 28, 19, 1]. Polis could be utilized to facilitate personal data exchange in personal-level microtrades between Polis-agents. Polis-users can give permission to information gathering companies to access (some of) their personal data, for an agreed price. Each time a company needs to regain access to them, the agreed amount of money should be paid. Furthermore, Polis could provide the ground for more advanced financial applications for personal data. The market for personal data described by Laudon in [22] is an example of such applications. In particular, Laudon proposes the

so called National Information Market (NIM), where personal information can be traded in a National Information Exchange. The adoption of a framework like Polis would simplify the evolution of NIM-like infrastructures.

Privacy-enhanced ubiquitous computing. Online data of an individual can be conveyed through her Polis-agent. In this case, Polis could work as an open architecture for ubiquitous computing applications. For example, dynamic location information could be retrieved from the individual's Polis-agent, like the rest of her personal data.

Privacy-Enhanced personalization. An important feature of services provided by web sites is personalization. Even though most users are interested in personalization, their privacy concerns are a serious obstacle to the wider use of personalization [20]. Both approaches, either having customer data and personalization code only at the shop side, or only at the customer side, have serious drawbacks. Polis can support an intermediary solution to overcome many of the difficulties involved. The personalization rules can be executed on the shop-side while the customer data reside at the customer-side.

4 The Polis prototype

Polis is work in progress. The main objective of the current prototype has been to demonstrate that electronic transactions are feasible while personal data remain only at the owner's side. Preliminary experiments confirm the above claim. Another technical objective of the development of the Polis prototype was to make its deployment simple and friendly to contemporary information management practices. We believe that we have fulfilled the above design goal adequately. In order to deploy Polis:

- Customers install the Polis-agent, store their personal information and prepare the necessary policy templates.
- Companies install the Polis-agent, prepare policy templates and integrate the agent with the company's back-office. *Polis-customers can co-exist with normal customers at a company side.*

Furthermore, we believe that a fully developed Polis platform can satisfy the general properties that a privacy technology must have in order to be considered useful [11].

We prepared an elementary Polis environment with a set of Polis-agents installed on the local network of our laboratory. A set of web pages, web forms and dynamic web pages were used to perform the experiments. The customer database contained 27 customers in total, 11 conventional customers and 16 Polis-customers (4 of which using Tor hidden services [6] for their agents). Customer registration with and without Tor anonymization and personal data retrieval operations have been performed. Due to limited space this material can be found in the long version of the paper [8].

5 Discussion

In this work, we describe Polis, a conceptual data management framework which embodies a promising fundamental privacy principle. Polis aims at making storage of personal data unnecessary for contemporary online transactions to work efficiently. This way, users will be able to monitor and limit the distribution of their personal data, according to their needs and preferences. Furthermore, the safety of stored personal data is enhanced and personal data accuracy is ensured.

In conclusion, this work demonstrates the fact that it is possible to deploy a security prototype like Polis, in order to achieve significant privacy protection, in the current electronic world. We cannot expect Polis to become a panacea for all kinds of privacy problems. However, we believe that Polis has more advantages than disadvantages compared to current practices for personal data management. Finally, it is very encouraging that given only a few assumptions, the transition to a personal data respecting way of conducting online transactions, can be natural and smooth.

References

[1] A. Acquisti. Privacy and security of personal information: Technological solutions and economic incentives. In J. Camp and R. Lewis, editors, *The Economics of Information Security*, pages 165–178. Kluwer, 2004.

[2] G. Aggarwal, M. Bawa, P. Ganesan, H. Garcia-Molina, K. Kenthapadi, R. Motwani, U. Srivastava, D. Thomas, and Ying Xu 0002. Two can keep a secret: A distributed architecture for secure database services. In *CIDR*, pages 186–199, 2005.

[3] E. Bangerter, J. Camenisch, and A. Lysyanskaya. A cryptographic framework for the controlled release of certified data. In B. Christianson, B. Crispo, J.A. Malcolm, and M. Roe, editors, *Security Protocols Workshop*, volume 3957 of *LNCS*, pages 20–42. Springer, 2004.

[4] K. Bohrer and B. Holland, editors. *Customer Profile Exchange (CPExchange) Specification*. IDEAlliance, 2000.

[5] Shi-Cho Cha and Yuh-Jzer Joung. From p3p to data licenses. In *Privacy Enhancing Technologies*, pages 205–222, 2003.

[6] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The second-generation onion router. 2004.

[7] DISCREET. Discreet service provision in smart environments. FP6-2004-IST-4 contract no. 27679. http://www.ist-discreet.org/.

[8] P.S. Efraimidis, G. Drosatos, F. Nalbadis, and A. Tasidou. Towards privacy in personal data management. Technical report, Democritus University of Thrace, Greece, June 2008. http://polis.ee.duth.gr.

[9] M. Fahrmair, W. Sitou, and B. Spanfelner. Security and privacy rights management for mobile and ubiquitous computing. In *Workshop on UbiComp Privacy*, 2005.

[10] Ian Goldberg. *A Pseudonymous Communications Infrastructure for the Internet*. PhD thesis, Univ. of California at Berkeley, 2000.

[11] Ian Goldberg. Privacy-enancing technologies for the internet iii: Ten years later. In A. Acquisti, S. Gritzalis, C. Lambrinoudakis, and S. di Vimercati, editors, *Chapter 1 of Digital Privacy: Theory, Technologies, and Practices*. December 2007.

[12] P. Golle, F. McSherry, and I. Mironov. Data collection with self-enforcing privacy. In *CCS '06: 13th ACM conference on Computer and communications security*, pages 69–78, New York, NY, USA, 2006. ACM.

[13] S. Gritzalis. Enhancing web privacy and anonymity in the digital era. *Information Management and Computer Security*, 12(3):255–287, 2004.

[14] J.I. Hong. *An Architecture for Privacy-Sensitive Ubiquitous Computing*. PhD thesis, University of California at Berkeley, Computer Science Division, Berkeley, 2005.

[15] P. Jäppinen. *ME - Mobile Electronic Personality*. PhD thesis, Lappeenranta University of Technology, Finland, 2004.

[16] G. Karjoth and M. Schunter. A privacy policy model for enterprises. In 15th IEEE Computer Security Foundations Workshop, 2002.

[17] G. Karjoth, M. Schunter, and M. Waidner. The platform for enterprise privacy practices - privacy enabled management of customer data. In 2nd Workshop on Privacy Enhancing Technologies (PET), 2002.

[18] S.K. Katsikas, J. Lopez, and G. Pernul. Trust, privacy and security in e-business: Requirements and solutions. In *Panhellenic Conference on Informatics*, pages 548–558, 2005.

[19] J. Kleinberg, C. Papadimitriou, and P. Raghavan. On the value of private information. *TARK: Theoretical Aspects of Reasoning about Knowledge*, 8, 2001.

[20] A. Kobsa. Privacy-enhanced personalization. *Commun. ACM*, 50(8):24–33, 2007.

[21] L. Korba and S. Kenny. Towards meeting the privacy challenge: Adapting drm. In *Digital Rights Management (LNCS 2696/2003)*, pages 118–136. Springer Berlin / Heidelberg, 2003.

[22] K.C. Laudon. Markets and privacy. *Commun. ACM*, 39(9):92–104, 1996.

[23] G.V. Lioudakis, E.A. Koutsoloukas, N.L. Dellas, N. Tselikas, S. Kapellaki, G.N. Prezerakos, D.I. Kaklamani, and I.S. Venieris. A middleware architecture for privacy protection. *Comput. Networks*, 51(16):4679–4696, 2007.

[24] D. Mulligan and A. Schwartz. Your place or mine?: privacy concerns and solutions for server and client-side storage of personal information. In *CFP '00: Proceedings of the tenth conference on Computers, freedom and privacy*, pages 81–84, New York, NY, USA, 2000. ACM Press.

[25] PRIME. Privacy and identity management for europe. EC Contract No. IST-2002-507591. https://www.prime-project.eu/.

[26] F. Salim, N.P. Sheppard, and R. Safavi-Naini. Enforcing p3p policies using a digital rights management system. In N. Borisov and P. Golle, editors, *Privacy Enhancing Technologies*, volume 4776 of *LNCS*, pages 200–217. Springer, 2007.

[27] P. Samuelson. Privacy as intellectual property? *Stanford Law Review*, 52:1125, 2000.

[28] Hal Varian. Economic aspects of personal privacy. U.S. Dept. of Commerce, Privacy and Self-Regulation in the Information Age, 1996.

[29] W3C. The platform for privacy preferences 1.0 (p3p1.0) specification, 2002. http://www.w3.org/TR/P3P.