Noname manuscript No.
(will be inserted by the editor)

Ciarán Bryce · Marnix A. C. Dekker · Sandro Etalle · Daniel Le Métayer · Frédéric Le Mouël · Marine Minier · Joel Moret-Bailly · Stéphane Ubéda

Ubiquitous Privacy Protection

Position Paper

Abstract The main message of this position paper is that ubiquitous computing technology need not necessarily be an obstacle to privacy protection: if legal and social issues are given proper consideration, technology can also be used to allow individuals to exercise their rights. We believe that the key issue is to devise techniques able to support ambitious privacy protection policies while allowing for the flexibility required in the ambient intelligence context. We illustrate our position using three technical requirements: (1) formal specification of privacy policies, (2) trust management and (3) auditability, which show both the challenges posed by ubiquitous computing and the opportunities to strengthen privacy. For each of these requirements, we present the legal and social motivations, suggest technical challenges and provide hints on possible solutions based on our on-going work.

Keywords trust management · audit · a posteriori verification · privacy policy model · formal specification · user consent

This research has been supported by the projects PRIAM - Privacy issues in ambient intelligence - funded by INRIA, PEARL - Privacy enhanced security architecture for RFID labels - funded by NL-STW/Sentinelsand, PAW - Privacy in an Ambient World - funded by SenterNovem and SERENITY - System Engineering for Security and Dependability - funded by the European Union.

Contact address: Daniel.Le-Metayer@inrialpes.fr
Affiliations:
C. Bryce, D. Le Métayer: INRIA (France)
M.A.C. Dekker: TNO (the Netherlands)
S. Etalle: university of Twente (the Netherlands) and university of Trento (Italy)
F. Le Mouël, M. Minier, S. Ubéda: INSA-Lyon (France)
J. Moret-Bailly: university Jean Monnet, Saint-Etienne (France)

1 Introduction

Privacy is a complex and multi-faceted notion, both from the social and the legal points of view and it has been interpreted in various ways depending on times, cultures and individual perceptions. Privacy is also very much dependent on technology, as evidenced by the motivation put forward by Warren and Brandeis in their seminal paper [18] back in 1890 - namely the growing use of photography - and, more than one century later, by the increased concerns about privacy raised by the perspectives of ubiquitous computing. In each case, technology plays the role of the villain, the recurrent source of new privacy threats. The position that we put forward here is that technology should not necessarily always be on the wrong side, and that the technical characteristics of ubiquitous computing can also be exploited to extend privacy protection. Obviously, privacy cannot be guaranteed by purely technical means, especially in the context of ubiquitous computing: a combination of legal, social and technological ingredients is required, but technology can and should be designed from the outset with privacy requirements in mind. In other words, if it is true that regulation should in some sense be "technology neutral", technology does not have to be neutral w.r.t. regulation or society; technology should serve the social and the public interests as manifested in regulations.

One of the main difficulties with privacy is that it cannot be absolute, especially in a world of "disappearing computers". It has to be balanced with other rights (e.g. free speech), obligations (e.g. data retention), principles (e.g. liberty) and interests (e.g. personalized services, convenience). This observation does not imply that technical protections have to be weak; on the contrary, they have to be strong, but flexible. Technology can and should support legal and social regulations as much as possible, while accounting for the variety of situations and complexities of the "ambient intelligence" world.

In this abstract, we focus on three technical requirements: (1) formal specification of the privacy policy, (2) trust management and (3) auditability. We choose these

three because they illustrate both challenges posed by ubiquitous computing and opportunities to strengthen privacy. In addition, these requirements apply to the three main stages of communication in ambient intelligence scenarios:

1. The privacy policy model for a users device is defined beforehand, and serves as a reference for all interaction decisions that the device takes.
2. Trust comes into play when a users device interacts with other devices: the devices trust management unit assesses potential risks, given the current context, and grants access rights to the other devices in consequence.
3. Auditability is useful after an interaction, as a means to verify that past actions by the device were consistent with the privacy policy and the granted rights. This is also useful as input for future trust management decisions.

We review these three technical requirements in, respectively, Sections 2, 3 and 4. For each requirement, we present the legal and social motivation, suggest technical challenges and then provide hints on possible solutions based on our on-going work. Section 5 draws some conclusions and points out further legal issues.

2 Requirement for a formal privacy policy model

2.1 Legal and social motivation

Most legal instruments for privacy protection explicitly refer to the unambiguous consent of the data owner as one of the conditions for the collection of personal data (see, e.g., Article 7 of [16]). However, consent of the data owner can be deemed unambiguous only if the information on which such consent is based is also unambiguous. This prerequisite is far from obvious considering the intrinsic complexities of privacy issues and the lack of clarity of certain privacy claims or declarations. The discussions concerning privacy statements in languages such as P3P and their potential inconsistencies show that a formal semantics for such languages is urgently needed.

This need is even more striking in the ubiquitous computing context where the data owner does not want his consent to be requested before each individual data communication: in some sense, it could even be seen as contradictory with the basic assumption of "disappearing computers" and the idea of interactions carried out behind the user's back. This should not imply that the user is unable to control, even indirectly, such data transfers however. Nonetheless, the effective control of his personal data by the user and the legal value of his consent must rely on the existence of a precise privacy policy that has been defined beforehand, understood and accepted without ambiguity. Such policy, which can be seen as a form of "generic (and conditional) consent", may be rather complex given the variety of situations to be considered.

Another significant reason for avoiding ambiguities in the statement of privacy policies is the liability of the data controller (the legal person who determines the purposes and means of the processing of personal data) for damages suffered by the data owner as a result of unlawful processing of personal data (see, e.g., Article 23 of [16]). Indeed, the liability of the data controller may be very difficult to enforce if his commitments, as expressed within the privacy policy, are undermined by inconsistencies or imprecisions.

In addition to the legal requirement, the delivery of a clear and consistent information to the data owner will obviously increase his trust in the system and make him less reluctant to let his personal data be automatically released and processed by devices which are out of his control.

2.2 Technological issues

As already illustrated for decades in the programming languages area [15,13], formal semantics have at least two major benefits: (1) they make it possible to assign a definite meaning to each statement in a language, and (2) they pave the way for the design of a variety of well-founded tools (such as program analyzers, property checkers, program transformers, etc.). Such benefits would be very much welcomed in the context of privacy policy languages as well: examples of useful features include conformance checking (e.g. checking whether the privacy policy of a potential data recipient conforms to the privacy requirements of the data owner), consistency checking (detecting inconsistencies or suspicious configurations in privacy policies), verifying the compliance of an implementation with respect to privacy statements, etc.

This is not to say, however, that the application to privacy policy languages of techniques devised for the formal specification of programming languages is straightforward. Indeed, defining privacy policies in a precise and unambiguous way is not an easy task, especially in the ambient intelligence world where many different contexts, actors, types of data and devices have to be considered. Examples of notions that must be handled include: purpose (of data processing), conditional rights (granted to the data controller), obligations (required from the data controller), revocations (e.g. decision from the data owner to revoke a right), time (e.g. delay for the right to keep data), etc. In addition, as argued in Sections 3 and 4, complex notions such as trust, auditability and liability are also desirable features of a privacy policy language. Last but not least, a mathematical model is not enough to reach the objectives set forth in the previous subsection: individuals should not be expected to understand mathematics, hence formal specifications should

be translated into understandable text in a natural language (or conveyed through any other user-friendly interface).

2.3 Technological solutions

In order to tackle the issues raised in the previous subsection, we have designed a formal system which relies on the very notions of trust and auditability [6,3] discussed in Sections 3 and 4. The trust management part of the system is based on the RT0 language [12]. The proof system for trust management defines the rules for establishing credential formulae such as "A asserts that D is a member of A.friends" where "A" denotes a principal and "friends" a role. The system makes it possible to express different forms of delegations (delegation to a principal, to all principals of a role, partial delegation, etc.).

The semantics of privacy policies can be established based on the state of the trust management system. The privacy policies are sticky in the sense that each data comes with its policy and this policy governs its use as well as its distribution. The proof system for privacy policies defines the conditions under which a principal is authorized to receive a data, to keep it and to send it to another principal. A significant feature of the system is that an agent is authorized to receive data only from trusted principals. Last but not least, as argued in Section 4, the privacy policies are not assumed to be enforced a priori but can be implemented *a posteriori*, through the checking of audit logs.

The first version of the system [6,3] has been kept minimal, for theoretical purposes, but several extensions are currently under investigation, including the introduction of obligations and the management of *reputation based trust* (in addition to the current *rule based trust management* - see Subsection 3.2).

3 Requirement for trust management

3.1 Legal and social motivation

Humans use trust daily to promote interaction and accept risk in situations where they have only partial information. Actually, the notions of privacy and trust have various connections:

- At the "operational level", trust is a social parameter for risk management; it helps in the decision making process, for example to decide whether or not personal data should be communicated. Until recently, trust was established through direct, face-to-face, interactions. Nowadays, in the global connected society, alternative methods for establishing trust are needed.
- At the "philosophical level", it has been argued that one of the first values encapsulated by privacy was to contribute to the well-being and trustworthiness of people, especially in the context of their social relationships.

Thus, trust helps maintain privacy, which in turn strengthens trust. However, establishing trust may also affect privacy since a minimal amount of information has to be provided to initiate trust (bootstrapping process). This illustrates again the fact that privacy should not be without limits: the extreme scenario where no one agrees to release personal information would completely undermine trust and thus endanger one of the very foundations of privacy itself.

From the legal point of view, the notion of trust underlies several notions such as "burden of proof", "minimal care" or "good faith". For example, the trust in an electronic signature satisfying legal requirements (e.g. use of a certified signature device and established certificate provider) is considered high enough to place the burden of proof on the party questioning the validity of the proof. In the same vein, individuals receiving personal data from a completely unknown third party could be accused of lack of care (just as a consumer buying counterfeited products in the street). It has also been argued that privacy protection has been afforded in USA courts on the basis of "breach of implied contract or of a trust or confidence" [18]. Obviously, such qualification remains under the judges's appraisal, but any piece of evidence substantiating this notion of trust could be of value in the decision process.

3.2 Technological issues

Regarding the protection privacy problem from a trust management point of view appears as a reasonable solution in a distributed framework such as ubiquitous computing where no control from a central authority can be assumed. Trust management can be a way to help individuals taking decisions on whether or not to consent to the processing of their private data. Two notions of trust are usually distinguished (see [14] for comprehensive presentation): *confiding* and *reliability trust*. Confiding is a merely passive attitude, corresponding to the case where certain negative events are expected not to take place. On the other hand, reliability trust can be seen, following Gambetta [8], as "the subjective probability by which an individual expects that another individual performs a given action on which his own welfare depends". Confiding and reliability trust can be implemented by two different trust management frameworks: confiding is addressed by *rule based trust management* while reliability trust is addressed by *reputation based trust management* (as already put into practice by commercial sites such as eBay).

Rule based trust management falls in the area of access control. The adaptation to decentralized collaborative systems of solutions available for centralized systems is very challenging, especially when considering ambient intelligence environments. On the other hand, in order to be socially acceptable, reputation based trust management systems must provide effects close to the human perception of trust. They must also be able to capture the dynamics of trust, taking into account not only the performed interactions but also the social context of those interactions.

In addition, any trust management framework for ambient intelligence requires a minimal exchange of personal data. As shown in the next subsection, several solutions are possible to limit the risks resulting from such disclosure: limitation of the quantity of information, of its quality, or of the associated use rights.

3.3 Technological solutions

Our trust management framework is based on the notion of communication history [7] : after a successful interaction, each node builds a proof of interaction signed by both parties. Each node then keeps in its history this common certificate to cryptographically prove that the interaction has taken place (and, implicitly, to the satisfaction of both parties). These certificates are used to build trust with other devices based upon their numbers of interactions with common parties (parties which have had successful interactions with the two nodes). More information can also be added to the certificates to further qualify the interactions (including, e.g., information about the level of satisfaction of the parties). The certificate could also contain a contract between the two entities including beacons, time-stamping information, localization information, and various constraints on use rights (e.g. time constraints or localization constraints).

In our framework, the quantity of personal data communicated during the trust establishment process is limited by the use of an iterative protocol allowing the user to disclose his data gradually according to dynamically computed trust levels. This iterative process can also take into account the current context. The level of quality of the information can be adapted through the use of pseudonyms. Each node is allocated a unique identity and a set of derived pseudonyms that can be used to build certificates without revealing the identity of the node. In addition, the nodes can report bad behaviors to the authority, which is able to establish the identity of the incriminated node, blacklist it with all its pseudonyms and forward this information to all other nodes. Needless to say, the authority is not assumed to be permanently accessible.

4 Requirement for auditability

4.1 Legal and social motivation

Classical mechanisms for the protection of confidential data are *preventive*, in the sense that unauthorized actions are prevented from occurring. Preventive measures can be very successful and should be used whenever possible. However, both for technical and legal reasons, they should be complemented with *a posteriori* measures such as auditing, especially in the context of ubiquitous computing. First, relying exclusively on preventive (or *a priori*) controls would not be realistic for certain applications because the system would necessarily be:

- Either inconvenient, because it would not tolerate any exceptions and would thus deprive the user of many potential interactions or repeatedly require his consent before authorizing them; or,
- Ineffective, either because the user, tired with repeated consent queries, would systematically give his consent or because, in order to alleviate this problem, the system would have to implement a much weaker privacy policy.

The crux of the problem is that a strict and precise preventive control mechanism requires a complete decision procedure to establish whether a certain action is allowed or not, and such a complete decision procedure cannot be effectively implemented since it depends on vague concepts such as the "purpose" of an action. In addition, preventive approaches cannot deal with conflicting requirements that often arise when policies and regulations are emanated by different authorities; nor can they cope with unforeseen interactions, which are very common in ambient intelligence.

From a social point of view, knowing that a company or a third party is auditable obviously increases the trust of the individuals and their willingness to enter into interactions. In addition, ensuring that individuals can be held accountable for their actions can have a strong deterrence effect and we believe that, in certain situations, deterring illegitimate behaviors can be almost as effective as preventing them, with the great advantage of being much more flexible.

From the legal perspective, international instruments such as the European Directive 95/46/EC [16] explicitly refer to the accountability and liability of the data controller. Generally speaking, the preventive (incentive or deterrence) effects of regulation may depend a lot on the actual possibilities of legal action, and these possibilities depend in turn on the availability of evidence that can be used in legal proceedings.

4.2 Technological issues

Auditing can be a valid substitute to preventive access control in that it can implement a form of *a posteriori*

access control: a system in which policy infringements are not prevented, but are systematically logged and detected. An advantage of a posteriori access control is that it makes it possible to define more expressive and powerful privacy policies, including, e.g. conditions on the future use of personal data by the recipient. It is thus another example of the application of the technology to strengthen the rights of the individuals.

Systematic use of auditing to achieve compliance control may seem futuristic, but we believe it is bound to happen. To realize this, a number of technical issues have to be addressed. First, *completeness*: auditing can never be a fully automated process because the decision of whether an action constitutes an infringement is often based on factors (such as the social context) that cannot be evaluated by an automatic system. Therefore, auditing systems will have to isolate dubious cases that may require human intervention. Also, a necessary balance has to be found with respect to the amount of information recorded because logging too much data could in some cases be considered in itself as a breach of privacy. Secondly, *effectiveness*: in the presence of dubious cases, the problem of false positives may be a limiting factor in the effectiveness of automatic auditing. A third challenge is bridging the gap between policies and events. While privacy policies are expressed in terms of high-level concepts (e.g., disclosure, purpose, etc), logging is usually carried out by collecting low-level events (IP packets, file accesses etc). The difficulty here is determining which policies may be infringed by an event and whether the event actually constitutes an infringement.

Concerning the expressiveness of the auditing (i.e. the amount of information what should be collected and logged), it should allow checking of compliance to policies that include conditional rights and obligations, revocations, time constraints, etc. Finally, logging and auditing are techniques that require a substantial amount of resources both in terms of storage and computing power; the limited resources of typical ubiquitous computing devices is thus an additional difficulty to implement an a posteriori control system.

Next to the technical challenges, there are the legal and social ones. In particular, so long as audit logs have no validity as evidence in court, their effectiveness as deterrents is bound to be limited. One of the key requirements that audit logs will have to satisfy to achieve this aim is security (especially integrity and authenticity).

4.3 Technological solutions

In order to address some of the issues raised in the previous subsection, we have defined a logic for accountability [2,3], the formal basics of *a posteriori compliance control* [6], and a logical framework for assessing log systems [5]. These systems address in particular the formal questions underlying the structure and meaning of auditing as a way to enforce compliance control. As mentioned before, this technology can help strengthening privacy and the trust of the individuals also because it can be driven by the individual himself and used for instance to discover *a posteriori* how some private information has actually been used by a third party.

Still unexplored are the practical issues related to the design and deployment of the logging and accountability systems. Particularly challenging seems the realization of automatic auditing authorities with low "false positive" rates. Another crucial issue that is currently under study is the scalability of the process with solutions based on the relaxation of the conditions under which actions have to be kept in the audit log.

As set forth in the previous subsection, a key requirement for auditing is that logged data be stored securely. We are investigating this issue using the notion of *trusted hardware*. An example of a general purpose hardware is the *Trusted Platform Module* (TPM) [17]. Currently, the TPM is used to store digests of the code loaded on a platform; the TPM data thus allows the remote party to verify that the platform is not running compromised code. There are three aspects to the usefulness of the TPM for privacy enforcement. First, the TPM can be used to securely store hashes of any data on a device, and attest the hash to other devices. The TPM can thus be used to store secure hashes of audit logs. Second, TPMs store keys that never leave the trustworthy confines of the TPM; these keys can be used for secure identity management and authentication. Third, the attestation feature can be used to validate any privacy enforcement software running on a device. Thus, a device A might only decide to interact with B if B is able to attest that it is running the correct version of a privacy policy engine. This is also important for detecting viruses, since privacy violations can arise from virus-infected software as easily as through the actions of malicious users.

Fundamentally, auditing for privacy entails recording the actions of entities in relation to the exchange of information. Sometimes information can be identified through the storage device that contains it. In this case, RFID technology can be exploited to trace the movement of the storage device. For paper printouts of personal information, RFID can also be used. In February 2007 Hitachi unveiled a RFID tag measuring 0.05x0.05mm, thus thin enough to be embedded in a sheet of paper. Thus, tags can even be used to track paper documents from the moment that they leave the printer[1]. This is another example of where an ambient computing technology, criticized for the threat it poses to privacy, can also be exploited as a privacy enforcing technology.

[1] Obviously, such scenario may still suffer from the "analogue hole" if tagged documents can be copied on untagged sheets of paper, but the purpose of the solutions outlined here is to lift up the level of difficulty to breach privacy policies, rather than to provide definite (or even extremely strong) security guarantees, so that breaching actors cannot easily use good faith arguments for their defence.

5 Conclusion

For the sake of conciseness, we have focused on three technologies that we believe are of significant importance for reconciling privacy and ubiquitous computing. In each case, we have suggested how the technology can be used to extend privacy protection while allowing for the flexibility required by the ambient intelligence context. Obviously, this is not to minimize the value of many other techniques (such as anonymity, cryptography, authentication or proximity based techniques) that should be part of any privacy friendly ubiquitous environment. Another fundamental issue not covered here is the privacy architecture that integrates these techniques ([10,11]). This notion of privacy architecture is especially important in the PRIAM project[2] which is following a top-down approach: starting from the legal requirements, we successively derive requirements for a privacy policy model and its implementation.

Although the main focus in this abstract is on technological issues, we do not want to suggest that technology can solve all problems. Indeed, further legal protection is desirable to ensure that the technical solutions put forward here are viable. For example, the legal value of transactions conducted by electronic means (such as a partially automated user consent, decisions based on automatic trust evaluation or on automated audits) has to be established. This also holds for "privacy enhancing tools" [1], as recently acknowledged by the European Commission [4].

To conclude, the main message that we want to convey through this position paper is that technology need not necessarily be an obstacle for privacy protection: if legal and social issues are considered from the outset [9], then technology can also be used to allow individuals to exercise their rights. Just to take a few final examples: the privacy policy system sketched in Section 4 can be used to implement a data tracking tool, allowing an individual to check who has received his personal data and, if he so desires, to ask him to correct or erase it; tools can also be designed for negotiating privacy policies and adapting the amount of personal information delivered (including reverting from identity to pseudonym or to anonymity) or for verifying (or assisting in the verification of) the conformance of a published policy statement with the actual implementation (e.g. within a certification or an audit process).

References

1. L. A. Bygrave. Privacy-enhancing technologies - caught between a rock and a hard place. *Privacy Law and Policy Reporter*, 9:135–137, 2002.
2. J. G. Cederquist, R. J. Corin, M. A. C. Dekker, S. Etalle, and J. I. den Hartog. An audit logic for accountability. In A. Sahai and W. H. Winsborough, editors, *Proc. of the Int. Workshop on Policies for Distributed Systems and Networks (POLICY)*, pages 34–43. IEEE Computer Society Press, June 2005.
3. J. G. Cederquist, R. J. Corin, M. A. C. Dekker, S. Etalle, J. I. den Hartog, and G. Lenzini. Audit-based compliance control. *International Journal of Information Security*, 6(2-3):133–151, 2007.
4. European Commission. Promoting data protection by privacy enhancing technologies. *EC Communication, IP/07/598*, 2007.
5. S. Etalle, F. Massacci, and A. Yautsiukhin. The meaning of logs. In T. A. Min and L. Costas, editors, *Proc. 4th Int. Conference on Trust, Privacy & Security in Digital Business (TrustBus)*, page to appear. Springer, 2007.
6. S. Etalle and W. H. Winsborough. A posteriori commpliance control. In B. Thuraisingham, editor, *Proc. 12th ACM Symp. on Access Control Models and Technologies (SACMAT)*, page to appear. ACM Press, 2007.
7. S. Galice, M. Minier, J. Mullins, and S. Ubéda. Cryptographic protocol to establish trusted history of interactions. In Levente Buttyán, Virgil D. Gligor, and Dirk Westhoff, editors, *ESAS*, volume 4357 of *Lecture Notes in Computer Science*, pages 136–149. Springer, 2006.
8. D. Gambetta. Can we trust trust? In Diego Gambetta, editor, *Trust: Making and Breaking Cooperative Relatioins*, chapter 13, pages 213–237. Published Online, 2000.
9. M. Langheinrich. Privacy by design - principles of privacy aware ubiquitous systems. In G.D. Abowd, B. Brumitt, and S. Shafer, editors, *Ubicomp 2001 Proceedings*, volume 2201 of *Lecture Notes in Computer Science*, pages 273–291. Springer, 2001.
10. M. Langheinrich. A privacy awareness system for ubiquitous computing environments. In *Ubicomp 2002 Proceedings*, volume 2498 of *Lecture Notes in Computer Science*, pages 237–245. Springer, 2002.
11. M. Langheinrich. Personal privacy in ubiquitous computing. tools and system support. *Dissertation Document ETH 16100*, 2005.
12. N. Li, J. Mitchell, and W. Winsborough. Design of a role-based trust-management framework. In M. Abadi and S. M. Bellovin, editors, *Proc. of the Symp. on Research in Security and Privacy (S&P)*, pages 114–130. IEEE Computer Society Press, 2002.
13. D. Le Métayer and D. Schmidt. Structural operational semantics as a basis for static program analysis. *ACM Computing surveys*, 28(2), 1996.
14. G. Sartor. Privacy, reputation, and trust: Some implications for data protection. EUI-LAW Working Papers 4, European University Institute (EUI), Department of Law, March 2006. available at http://ideas.repec.org/p/erp/euilaw/p0040.html.
15. D. Schmidt. Denotational semantics : a methodology for language development. 1986.
16. The European Parliament and the Council of the European Union. UE DIRECTIVE 95/46/EC 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, 1995.
17. Trusted Computing Group. TPM main specification. Main Specification Version 1.2 rev. 85, Trusted Computing Group, February 2005.
18. S. Warren and L. Brandeis. The right to privacy. *Harvard Law Review*, pages 193–220, 1890.

[2] PRIAM is a multidisciplinary project involving the university of law of Saint-Etienne, INRIA and the university of Twente.