Mandatory Enforcement of Privacy Policies using Trusted Computing Principles

Frank Kargl
University of Twente
Drienerlolaan 5
7522 NB Enschede
The Netherlands
f.kargl@utwente.nl

Florian Schaub and **Stefan Dietzel**
Ulm University
Albert-Einstein-Allee 11
89081 Ulm
Germany
givenname.surname@uni-ulm.de

Abstract

Modern communication systems and information technology create significant new threats to information privacy. In this paper, we discuss the need for proper privacy protection in cooperative intelligent transportation systems (cITS), one instance of such systems. We outline general principles for data protection and their legal basis and argue why pure legal protection is insufficient. Strong privacy-enhancing technologies need to be deployed in cITS to protect user data while it is generated and processed. As data minimization cannot always prevent the need for disclosing relevant personal information, we introduce the new concept of mandatory enforcement of privacy policies. This concept empowers users and data subjects to tightly couple their data with privacy policies and rely on the system to impose such policies onto any data processors. We also describe the PRECIOSA Privacy-enforcing Runtime Architecture that exemplifies our approach. Moreover, we show how an application can utilize this architecture by applying it to a pay as you drive (PAYD) car insurance scenario.

1. Motivation

Privacy and protection of personal data gains more and more importance, especially as our environment becomes more and more covered by sensors that collect data on our personal behavior. While this is considered a significant threat to privacy, the mere advantages that information and communication systems provide in terms of usability and user comfort will surely outweigh privacy concerns for most users. Therefore, preventing deployment of new systems is not a viable strategy for privacy protection. Instead, we must face the fact that users will willingly disclose some personal information to receive certain application benefits in exchange. Thus, a better strategy is to treat personal data as something very valuable that needs to be available for certain purposes that users consent to while, at the same time, non-consenting use must be prevented.

In this paper, we will advocate and describe such an approach and describe a technical architecture to implement it. For this purpose, we first highlight how legal and technical approaches to privacy protection differ and how they need to be combined to efficiently protect user privacy. Next, we motivate why users should be able to specify and control privacy policies that govern access to their personal data. Based on this, we present an approach to enforce policy compliant data processing rooted in trusted computing mechanisms. We further detail our approach by presenting a technical system architecture we have designed for policy enforcement in upcoming cooperative intelligent transport systems (cITS), as part of the European PRECIOSA project.[1] While our approach is not limited to cITS, the privacy problems faced there are well suited to exemplify many major issues. Thus, cITS will serve as a coherent example throughout the paper.

2. Cooperative Intelligent Transport Systems

Nowadays, vehicles contain many electronic assistive systems, like electronic stability control (ESC) or navigation systems, that aim to make the driving experience safer and more enjoyable for the driver. Currently, such systems operate only on a local scale by means of evaluating data of a vehicle's local sensors. However, by exchanging information between vehicles, roadside units, and back-end services, a new wave of applications is enabled that enhances safety and efficiency in a way not possible with local information only. Such systems, which rely on the cooperation between vehicles, back-end service providers, as well as supporting infrastructure, are called cooperative intelligent transportation systems (cITS). Cooperation is based on periodic or event triggered information exchange via either dedicated short range communication (DSRC as implemented in IEEE 802.11p) or cellular networks.

One category of applications that makes use of cITS as an enabling technology is usage based car insurance, also called pay as you drive (PAYD) insurance. In a PAYD scenario, drivers are charged a customized insurance fee based on their driven distance and driving style. Thus, in contrast to classic insurance models, speeding or using roads with higher accident risks can directly influence insurance fees. One way of implementing such an application uses so called floating car data (FCD). Messages consisting of the vehicles' unique identifier, GPS position, current speed, and a time stamp are periodically sent to the insurance provider that later assesses them for billing purposes.

[1] PRECIOSA website:
`http://www.preciosa-project.org`

To calculate the resulting insurance fees per driver, the insurance provider only needs to access the collected data in a particular manner, for instance, to average the driving speed per driver or to calculate the total amount of kilometers driven per billing period. Fine-grained data access is only required to prove correctness in case of disputes over charged fees. However, with the data collected, the insurance provider would also be able to deduct exact driving patterns. These patterns can be used for many commercial purposes that are not covered by the terms of usage drivers agreed to. Additionally, insurance providers can use collected data to base insurance calculations on observations which are not part of the contract, e.g., driving through bad neighborhoods. If no additional means are taken to properly protect the fine-grained user data, this is a potential privacy fiasco waiting to happen, which would not be limited to PAYD applications but would apply to FCD applications in general. The potential privacy implications of cITS have been already picked up by mainstream media, for example, in a 2009 article of the Guardian.[2] Moreover, the European ITS action plan that furthers the deployment of cITS has been questioned by the European Data Protection Supervisor (EDPS) for not sufficiently taking privacy into consideration (Hustinx 2009). Similar criticism can be observed in other countries.

3. Legal vs. Technical Privacy Protection

As highlighted by Hustinx (2009), it is of paramount importance for any intelligent transportation system to follow a *Privacy by Design* approach that deeply embeds privacy principles starting at the earliest stages of system design.

Agrawal (2002) presents a list of ten privacy principles, which are also compatible with similar principles stated in European data protection directives (European Parliament and Council 1995; 2002), by the OECD (OECD 1999), and by many national laws like the German data protection law (Bundesrepublik Deutschland 2003):

1. *Purpose Specification.* For personal information, the purposes for which the information has been collected shall be associated with that information.
2. *Consent.* The purposes associated with personal information shall have consent of the donor of the personal information.
3. *Limited Collection.* The personal information collected shall be limited to the minimum necessary for accomplishing the specified purposes.
4. *Limited Use.* The system should only allow such data accesses that are consistent with the purposes for which the information has been collected.
5. *Limited Disclosure.* The personal information stored in the system shall not be communicated outside system boundaries for purposes other than those for which there is consent from the donor of the information.
6. *Limited Retention.* Personal information shall be retained only as long as necessary for the fulfillment of the purposes for which it has been collected.
7. *Accuracy.* Personal information stored in the system shall be accurate and up-to-date.
8. *Safety.* Personal information shall be protected by security safeguards against theft and other misappropriations.
9. *Openness.* A donor shall be able to access all information about the donor stored in the system.
10. *Compliance.* A donor shall be able to verify compliance with the above principles. Similarly, the system shall be able to address a challenge concerning compliance.

Although laws often mandate compliance to those principles, current practice shows that privacy protection laws are regularly broken and consequences for offenders are not significant. Although German privacy laws are generally considered to be among the most stringent world-wide, in 2008, privacy incidents at major companies like Deutsche Telekom[3] or Deutsche Bahn (German Railway)[4] highlighted problems with data protection laws. While having such laws and enforcing them better is a vital ground for privacy protection, we argue that breaching privacy on a large scale is far too easy with today's systems collecting and processing large amounts of personal data. Thus, legal measures cannot be the only line of defense. Instead, systems should have some inherent technical protection to prevent privacy infringements in the first place. This opens the opportunity to assign liability in cases where protection mechanisms are circumvented on purpose.

Currently, most *Privacy-Enhancing Technologies* (PET) either employ *data minimization*, or they use *policies* to govern data usage. The goal of data minimization PETs is to reduce the amount of personal data that is exposed to communication partners. One example here is the use of pseudonymous authentication schemes often proposed for vehicular communication systems (Papadimitratos et al. 2008). Plain vehicle identifiers are removed from authentication credentials, like certificates, and, thus, prevent attackers from learning the identity of vehicles. However, sometimes minimization is not possible, because the communication partner or service provider needs access to personal information to perform its service despite the fact that this information requires privacy protection. Here, policy-based PETs can be used where a user explicitly states the kind of processing he consents to. P3P (Cranor et al. 2006) is an example of such a technology. However, such PETs often only provide the policy language without actually enforcing that policies are kept.

[2] Big Brother is watching: surveillance box to track drivers is backed — The Guardian online, 31 March 2009, http://www.guardian.co.uk/uk/2009/mar/31/surveillance-transport-communication-box

[3] Deutsche Telekom collected and scrutinized call data of journalists and members of the supervisory board during 2005 and 2006 — Time.com, 27 May 2008, http://www.time.com/time/business/article/0,8599,1809679,00.html

[4] In 2009, the CEO of Deutsche Bahn resigned after active manipulation of trade union leaders' e-mail traffic became public — Times Online, 31 March 2009, http://www.timesonline.co.uk/tol/news/world/europe/article6004352.ece

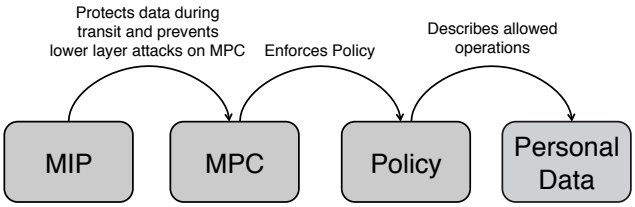

Figure 1: Protection chain created by PeRA Mechanisms that defines the policy enforcement perimeter.

PriPAYD (Troncoso et al. 2007) presents an existing approach for protecting privacy in the above mentioned PAYD insurance applications. PriPAYD follows the data minimization strategy. However, it is highly application specific and also implements only one specific policy. There are currently no general privacy solutions for cITS independent of specific applications.

In the next section, we describe how a generic system for mandatory enforcement of privacy policies can be designed that employs policies while still allowing data minimization principles to be applied. Moreover, it not only allows to declare policies, but also enforces these policies.

4. Mandatory Enforcement of Privacy Policies

We take a data-centric approach on enforcing privacy. According to the data protection principles, a person disclosing personal information gives her consent for a specific purpose only and may also require limited collection and retention of such data. Thus, the person, as the data subject, sets a policy describing the conditions, under which data controllers and processors may us the data. Often, such policies are only stated implicitly. When policies are formalized, e.g., using P3P (Cranor et al. 2006), it is often at the discretion of the data processor to actually respect them (or not), but there is no technical enforcement.

We advocate an approach where policies are stated explicitly and are securely coupled with the data they govern. Furthermore, policies are technically enforced in the IT systems of every data processor. By doing so, we establish a trust domain that extends beyond local system boundaries, in which data subjects can rely on their policies being respected. We call this trust domain the *Policy Enforcement Perimeter* (PEP).

The policy enforcement perimeter is created by a chain of protection mechanisms as depicted in Figure 1. In a first step, personal data is augmented with a policy specifying allowed operations.

Next, *Mandatory Privacy Control* (MPC) enforces privacy policies whenever personal data is accessed. MPC is a reference to the mandatory access control schemes found in access control architectures, like Bell-LaPadula (Bell and Padula 1973). Just like it is not at the user's discretion to change access policies in mandatory *access* control schemes, it is not at the application's discretion to change privacy policies in our mandatory *privacy* control system. To the contrary, it is mandatory that applications have to obey to the user-defined privacy policies whenever personal information is accessed.

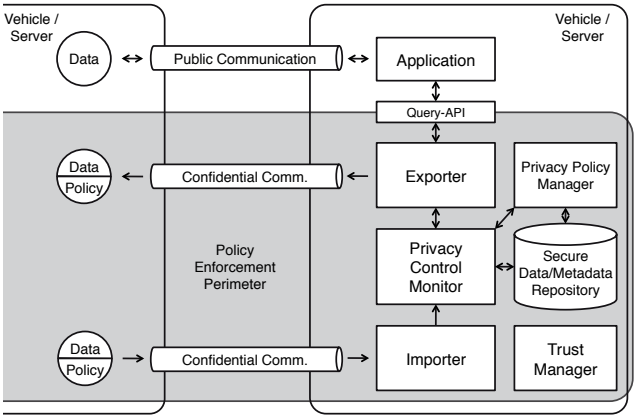

Figure 2: Overview of the privacy enforcing runtime architecture.

However, assuming that the data processor actually operates the IT systems, it would be rather trivial to circumvent MPC. Consequently, we introduce the concept of *MPC Integrity Protection* (MIP). Based on trusted computing concepts, the goal of MIP is to establish trust in remote systems and to ensure that only proper recipients with properly functioning MPC components in place are actually able to access the data during transmit or storage. If the MPC mechanism is tampered with, the MIP will detect the integrity violation and will prevent further data access.

Policies, Mandatory Privacy Control, and MPC Integrity Protection form a chain of control instances that establish the Policy Enforcement Perimeter, in which non-policy-compliant data access is prevented.

5. PRECIOSA Privacy-enforcing Runtime Architecture

The concept of mandatory policy enforcement as presented in the previous section is currently evaluated and implemented in the *Privacy-enforcing Runtime Architecture* (PeRA) of the PRECIOSA project. The PeRA is part of a larger framework for protecting privacy in cITS applications, for example, in the PAYD insurance scenario presented before. We will present the architecture while discussing the PAYD example in parallel.

Figure 2 gives an overview of the PeRA architecture. It shows the policy enforcement perimeter, which spans multiple physical systems, each running a PeRA instance. Re-

mote trust is established by means of trusted computing components encapsulated in the trust manager. The trust manager also protects the integrity of all components in the policy enforcement perimeter and prevents tampering.

Data is stored in a secure repository where access and policy compliance is checked by the privacy control monitor. The privacy policy manager assists by managing and analyzing policies. Applications can access policy protected data only by means of a specific query language. This query language directly relates to operations specified in policies. Transfer of policy protected data is achieved by confidential (i.e., encrypted and authenticated) communication channels, whereas public information can be sent and received by an application directly.

The first step towards ensuring privacy is the definition of a policy language, which allows users to specify exactly under which circumstances their data can be used. As policy languages are not the focus of this paper, we will just give an abstract example, which we subsequently use to explain our architecture. The following policy describes a possible usage restriction for a PAYD application:

BEGIN POLICY $payd\_policy$:

Controls:	$\text{payd\_dataset}(UID, pos, speed, timestamp)$
Processor:	FlexiInsurance Inc.
Purpose:	billing
Retention:	6 months
FIELD $speed$:	ALLOW AVG(per month)
FIELD pos:	ALLOW SUM(distance($p_1, \ldots, p_n$) per month)
FIELD UID:	ALLOW ATOMIC

END POLICY.

The policy is linked to a specific data set, and specifies the data processor that is allowed to process the data as well as a purpose. The retention period declares how long the data may be kept before it must be deleted. Finally, there are three restrictions on allowed operations for certain data fields, e.g., positions may only be used to calculate distances travelled per month.

Next we look at the policy enforcement perimeter and the trusted computing components, before describing in more detail how the policy control monitor mandates policy compliance.

Policy Enforcement Perimeter

After coupling the PAYD data set with the privacy policy, it needs to be sent to the insurance provider. As the data set contains sensitive information, the data transfer from vehicle to back-end service should use a confidential communication channel. A first step towards confidential communication would be to use asymmetric encryption. The PAYD data sets could be encrypted with the public key of the insurance provider to ensure that no adversary can eavesdrop on plaintext data. However, asymmetric encryption alone is not sufficient to mandate the insurance provider's adherence to the privacy policy, because the receiver then could decrypt the data and use it in arbitrary ways. Therefore, we employ a common system architecture at each cITS node that mediates access to stored data and metadata and mandates policy compliance. Figure 3 shows how asymmetric encryption is enhanced by the trust manager, which employs trusted computing principles to establish trust in remote systems. During system set-up, all components of the privacy preserving architecture are verified and it is checked that they have not been tampered with. Then, the trust manager measures and sets so-called platform configuration register (PCR) values that uniquely identify the currently running platform configuration. An asymmetric key pair $(P_T|S_T)$ is generated that is locked to the measured PCR values. The secret key is securely stored and managed by the trust manager, e.g., in a hardware security module (HSM).

Whenever a vehicle wants to send a PAYD data set, it first retrieves the public key P_T from the insurance provider. It then seals the PAYD data set D together with the policy Pol by encrypting it with the insurance provider's public key, obtaining $C = Enc_{P_T}(D|Pol)$. Then, the result C is sent to the insurance provider. Because the secret key S_T required for decryption is managed by the trust manager, the trust manager first receives the encrypted data. The trust manager checks whether all system components are still in the desired, i.e., validated, state. Only if this verification is successful, the encrypted data is decrypted by the trust manager and then placed in the secure data / metadata repository. Whenever data is queried by an application, the trust manager first checks whether the system configuration is still in a valid state before allowing access to the data repository.

This process ensures that data is only exchanged inside a trusted perimeter controlled by the trust managers of participating entities. Inside this perimeter, it is guaranteed that data is always coupled with its according metadata, which contains the policies specified by the data subjects, e.g., the insurance provider's customers.

Mandatory Privacy Control and Policy Control Monitor

Once data has been successfully stored in the secure data / metadata repository, all further access to the data is controlled by the privacy control monitor. Whenever an application wants to access the stored data, it submits a query through the query API together with the application's pur-

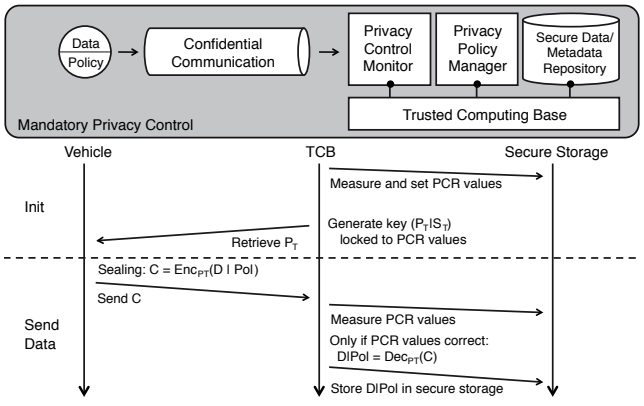

Figure 3: Interaction of a vehicle and a server to exchange confidential data.

pose and role. For our PAYD example, consider the following query request:

BEGIN REQUEST $payd\_request$:
Role: FlexiInsurance Inc.
Purpose: billing
Query: SELECT AVG($speed$) from $payd\_data$ WHERE $timestamp$ IN(09-01-01, 09-01-31) AND UID = 10
END REQUEST.

The query is analyzed and processed by the PCM. The PCM fetches all data affected by the query while the privacy policy manager retrieves all policies associated with the requested data. At this point, the role and purpose of the request are matched with the policies. In our example, a privacy policy is found that corresponds to the usage by FlexiInsurance Inc. for the purpose of billing. Then, the accessed data fields as well as the applied operations are matched with the policy defined. The example query uses the fields $speed$, $timestamp$ and UID:

- **Speed and timestamp:** The average function is applied on the requested speed and the time granularity is as coarse as mandated by the policy, therefore the access is allowed.
- **UID:** Only information about one single individual is accessed, therefore, the access is allowed.

Thus, the whole query adheres to the policy and the result set can be returned to the insurance provider. Note that the resulting data is again coupled with a merged policy based on policies of the single data items that contributed to the result. However, since the information leaves the privacy enforcement perimeter after passing the exporter, further adherence to the policy cannot be guaranteed anymore and is purely best effort on discretion of the application.

Queries by the insurance provider's application that are in conflict with the defined privacy policies would be rejected by the PCM. For example, an insurance provider **cannot**:

- *Access the maximum speed driven*, because the policy only allows access to the speed values using the average function.
- *Track vehicles*, because the policy only allows to calculate the total distance driven per month.
- *Build aggregated data over all customers*, because the policy mandates access only on unique UIDs. Note that this requires an intrusion detection component to detect subsequent probe queries.

Therefore the user can technically rely on the adherence to a given policy throughout the system.

6. Conclusion

In this paper, we have highlighted the need for proper privacy protection in cooperative intelligent transportation systems. We have outlined general principles for data protection and their legal basis. Furthermore, we argued why we consider pure legal protection an insufficient solution. We rather advocate a joint approach that combines legal protection with strong privacy-enhancing technologies, which should be deployed in cITS to protect data while it is generated and processed.

As data minimization cannot always prevent the need for disclosing relevant personal data, we have introduced the concept of mandatory enforcement of privacy policies. This empowers users and data subjects to tightly couple their data with policies and rely on the system to impose such policies onto anyone processing this data.

To achieve mandatory privacy protection, we described the PRECIOSA Privacy-enforcing Runtime Architecture that exemplifies our approach and also showed how a PAYD insurance application can make use of this architecture. Of course the architecture and its underlying principles can be used for many other applications and also in many contexts outside cITS.

While the need for trusted computing components and mandatory privacy control mechanisms will create additional effort for data processors, we argue that deploying such an architecture will enable users to establish significant trust into the system, which may result in greater willingness to provide personal information. Such personal information is required to make many systems and applications possible in the first place. Not providing strong PET, risking privacy infringements, and losing faith of users could prove to be more expensive in the long run.

While we have outlined the basic concepts of our architecture, many challenges still lie ahead. Currently, the partners of the PRECIOSA project are designing a policy language as the basis for the presented mechanisms. Aspects of ease of use and expressiveness have to be taken into account, as well as the automatic evaluation of queries and their compliance with policies.

Further, we are working on the concept of so called controlled applications. The goal of controlled applications is to allow for more complex operations directly on the data items inside the policy enforcement perimeter. For example, elaborate billing calculations based on complex patterns in the PAYD use case, which are not expressible by SQL-like statements. By controlling the information flow to and from controlled applications with a sandbox-like approach, privacy policy enforcement can be retained. However, it is an open issue how operations that a controlled application performs on accessed data can be evaluated to check for policy compliance.

7. Acknowledgments

Parts of this work have been funded by the EU project PRECIOSA (IST-224201). We want to thank all our project partners and especially our colleagues from the group of Prof. Johann-Christoph Freytag, Humboldt University, for fruitful discussions and their contributions to the presented architecture.

References

Agrawal, R.; Kiernan, J.; Srikant, R.; and Xu, Y. 2002. Hippocratic databases. In *28th VLB Conference*.

Bell, D. E., and Padula, L. J. L. 1973. Secure computer sys-

tems: Mathematical foundations. Technical report, MITRE Corporation.

Bundesrepublik Deutschland. 2003. Bundesdatenschutzgesetz (BDSG). Version as published on 14. January 2003 (BGBl. I S. 66), last changed in Article 1 on 14. August 2009 (BGBl. I S. 2814).

Cranor, L.; Dobbs, B.; Egelman, S.; Hogben, G.; Humphrey, J.; Langheinrich, M.; Marchiori, M.; Presler-Marshall, M.; Reagle, J. M.; Schunter, M.; Stampley, D. A.; and Wenning, R. 2006. Platform for privacy preferences 1.1 (P3P1.1) specification. World Wide Web Consortium.

European Parliament and Council. 1995. Directive 95/46/ec (protection of individuals with regard to the processing of personal data and on the free movement of such data). Official Journal L 281 , 23/11/1995 P. 0031 - 0050.

European Parliament and Council. 2002. Directive 2002/58/ec (directive on privacy and electronic communications). Official Journal L 201 , 31/07/2002 P. 0037 - 0047.

Hustinx, P. 2009. EDPS opinion on intelligent transport systems. *Opinions of the EU Data Protection Supervisor*.

OECD. 1999. Oecd guidelines on the protection of privacy and transborder flows of personal data. `http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html`.

Papadimitratos, P.; Buttyan, L.; Holczer, T.; Schoch, E.; Freudiger, J.; Raya, M.; Ma, Z.; Kargl, F.; Kung, A.; and Hubaux, J.-P. 2008. Secure vehicular communications: Design and architecture. *IEEE Communications Magazine* 46(11):100–109.

Troncoso, C.; Danezis, G.; Kosta, E.; and Preneel, B. 2007. Pripayd: privacy friendly pay-as-you-drive insurance. In *Proceedings of the 2007 ACM workshop on Privacy in electronic society*, 99–107. Alexandria, Virginia, USA: ACM.