

Privacy by design : towards a systematic approach

Daniel Le Métayer

TURBINE Workshop, January 2011

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE



Context: LICIT

LICIT: Legal Issues in Communication and Information Technologies

Objective: Contribute to the development of methods and tools for a better integration of technical and legal instruments



- **Multidisciplinarity**
- **Formal methods** as a link between law and ICT
- **Pragmatic** approach (goal directed)

Plan

1. Privacy by design
2. Exploring the design space: case study
3. Towards a systematic approach

Information and Privacy Commissioner of Ontario (Ann Cavoukian)

November 2008:

“The purpose of privacy by design is to give due consideration to privacy needs **prior to the development of new initiatives** – in other words, to consider the impact of a system or process on individuals’ privacy **and to do this throughout the systems lifecycle**, thus ensuring that appropriate controls are implemented and maintained. ”

Working Party 29

December 2009:

- “The principle of “Privacy by Design” should be introduced in the new framework: privacy and data protection should be integrated into the design of Information and Communication Technologies. ...
- This principle of “Privacy by Design” should not only be binding for data controllers, but also for technology designers and producers. On top of that, as the need arises, regulations for specific technological contexts should be adopted which require embedding data protection and privacy principles into such contexts.”

32nd International Conference of Data Protection and Privacy Commissioners

October 2010:

Privacy by design resolution :

- “Recognize Privacy by Design as an essential component of fundamental privacy protection; ...
- Proactively encourage research on Privacy by Design; ...
- ... ”

General principles

- Data minimisation
- Consent of the subject (informed, free, etc.)
- Limitation of use (purpose)
- Security (confidentiality, integrity, etc.)
- Transparency, traceability, accountability
- Quality
- Rights of the subjects (access, correction, deletion, etc.)

Plan

1. Privacy by design
2. Exploring the design space: case study
3. Towards a systematic approach

Case study: Pay as you drive

Impact of the minimisation principle on the architecture of the system



First option (centralized)

- On Board Equipment (OBE)
 - GPS
 - GSM
 - On board computer : sends all location data and vehicle identification to the server
- Operator
 - Computes the fee due for each car
 - Spot-checks the cars to detect misbehaviours or failures of the OBEs

First option (centralized)

Secure solution for the operator but

The operator knows all the whereabouts of all the vehicles \Rightarrow

Highly privacy intrusive

Second option (Vpriv)

- **On Board Equipment**
 - Commits to a fixed set of anonymous tags
 - Sends the location data to the server with anonymous tags
 - Adds the fees corresponding to its own tags and returns the sum to the server
- **Operator**
 - Spot-checks the cars to detect misbehaviours or failures of the OBEs
 - Computes the fee due for each location data received
 - Returns to each car all the individual fees (end of each quarter)
 - Conducts the verification protocol to check the sum returned by the OBEs (dedicated protocol for secure multi-party computation)

Second option (Vpriv)

Better solution for the driver but

- Requires anonymous communications
- Risks of de-anonymization
- Complexity and cost

Third option (Secure OBE)

- **On Board Equipment**

- Secure component
- Performs all the computations of the fees
- Sends the fee to the operator at the end of each quarter

- **Operator**

- Spot-checks the cars to detect misbehaviours or failures of the OBEs (two-way communications)

Third option (Secure OBE)

Excellent solution w.r.t. data minimisation but

- Requires more expensive OBEs
- Less secure for the operator
- Need to update the fee calculation software securely

Forth option (Commitments)

- **On Board Equipment**

- Sends vehicle identification and hashes of the location data to the server
- Performs the computations of the fees
- Sends the fee to the operator at the end of each quarter
- Discloses partial sums in case of spot-checks

- **Operator**

- Spot-checks the cars to detect misbehaviours or failures of the OBEs
- Conducts the verification protocol

Forth option (Commitments)

Flexible but

- Interactive verification protocol
- Non minimal disclosure of data during spot-checks

Enhancements:

- Commitment trees (de Jonge – Jacobs)
- Homomorphic commitments (Balash et. al. : PrETP)

Plan

1. Privacy by design
2. Exploring the design space: case study
3. Towards a systematic approach

Beyond specific technical choices

General strategy for data minimization :

1. **Divide** personal data (acceptable level of granularity)
2. **Transform** data items (acceptable level of information)
3. **Allocate** data items to the actors (possibly introducing third parties)

Towards a systematic approach

Situation:

- Service to be delivered
- Actors involved
- Requirements of each actor

Objective: exploration of the design space

- Find architectures which can both deliver the service and meet the requirements or
- Check if an architecture meets the requirements and can deliver the service

Formal model for data minimization

- **Service:** set of equations
- **Requirements of the parties:** constraints on sets of variables
- **Exploration of the design space:** inference system

Illustration : “Pay as you drive”

- **Actors:** VH, OP and Env
- **Service:** Eq : set of equations $X = F(Y_1, \dots, Y_n)$
defines the computation of the fee T (for a quarter)
- **Control function :** $\rho(X) = A$ actor controlling the computation of X

Control may differ from location : e.g. a secure component provided by OP can be part of the OBE while remaining under the control of OP (until its security is not compromised)

Formel system

Inference system :

- $C \vdash X$ in the context C , OP can detect any error in the computation of X
- $C = \{(R_i, D_i, G_i)\}$
 - R_i : set of variables which can be collected by OP in a run
 - D_i : set of variables which can be committed and discovered by OP in a run
 - G_i : set of variables which can be spot-checked by OP in a run

C : set of operations available to OP to perform his verifications

Formel system

Requirements of the parties:

- Constraints on contexts C
- Constraints on ρ
- $C \vdash T$: possibility for OP to detect any error in the computation of the final result (quarterly fee)

Examples:

- $\rho(T) = VH$ The computation of T must remain under the control of the vehicle
- $\forall (R,D,G) \in C, R \subseteq \{T\}$ The only data that OP can collect is T
- $\forall (R,D,G) \in C, \text{Card}(G) \leq 1$ No more than one spot-check in a quarter

Semantics

- **Semantics over distributed traces:** effect of each operation on the knowledge set of each actor
- **Assumptions on traces:** properties of cryptographic operations, notion of control, threat model (tampering with variables)
- **Proof of correctness of the inference system :**

$C \vdash X$ and $\text{Wrong}(X, \sigma) \Rightarrow$

OP can use the operations in C to extend the trace σ into a trace σ' which brings to OP the proof that X is not correct

Benefits of the formal approach

- Precise definitions of **assumptions** and **requirements**
- **Systematic exploration** of the design space
- **Guarantees** on the architectures
- Detection of **inconsistencies**

Privacy by design - Conclusion

- Consensus on the fact that the **privacy by design approach** should be promoted, supported by legal instruments and more widely adopted
- Consensus on **general principles** (data minimization, informed consent, transparency, etc.)
- **A range of techniques are already available** (anonymization, commitments, secure multiparty computation, homomorphic encryption, etc.)
- What is badly needed is a **systematic approach and tools to support it**
- Privacy is a **complex issue with potentially conflicting requirements** ⇒ **formal methods** can be play an instrumental role in this context

Bibliography

- J. Balasch, A. Rial, C. Troncoso, C. Geuens, B. Preneel, I. Verbauwhede, **PrETP: Privacy-Preserving Electronic Toll Pricing**, Proc. 19th USENIX Security Symposium, 2010.
- A. Cavoukian, **Privacy by design**, Report of the Information & Privacy Commissioner Ontario, Canada
- F. Garcia, B. Jacobs, **Privacy-friendly energy metering via homomorphic encryption**, Proc. 6th Workshop on Security and Trust Management, Springer Verlag LNCS, 2010.
- W. De Jonge, B. Jacobs, **Privacy-friendly electronic traffic pricing via commits**, Proc. Workshop of Formal Aspects of Security and Trust, Springer Verlag, LNCS 5491, 2009.
- D. Le Métayer, **Privacy by design: a matter of choice**, in Data Protection in a Profiled World, S. Gutwirth, Y Poullet, P. De Hert, Springer Verlag, 2010
- D. Le Métayer, **A formal privacy management framework**, Proc. Workshop of Formal Aspects of Security and Trust, Springer Verlag, LNCS 5491, 2009.
- A. Popa, H. Balakrishnan, A. Blumberg, **Vpriv: protecting privacy in location-based vehicular services**, Proc. 18th USENIX Security Symposium, 2009.