Inspecting Privacy in Electronic Services

Koen Decroix - Mobile and Secure
April 2015
Outline

• Introduction
• Methodologies
• Approach
• Privacy Modeling Concepts
• A Logic Based Modeling Framework for Analyzing Privacy
• Conclusions
• Questions
Introduction

Complex Electronic Services
Ever wondered what companies know about you?
… Max Schrems, an Austrian student, did!

Now he sues Facebook for their data practices on the personal data they collected about him.
A popular game platform from the late 70’s until the early 90’s
... evolved to multi-service platforms offering social gaming experience

SHARE Epic Moments
Show Off Your Greatness

Why keep your successes to yourself? Immortalise your favourite gaming triumphs and share them with your friends instantly at the tap of a button.
When registering to a service, you agree with the service provider’s terms and policies and give him your explicit consent for collecting, processing, and forwarding your personal information to collaborating third-parties.
... Non-transparent data handling practice declarations

Who is “who”?

WHO WE SHARE WITH:
We may share the personally identifying information of our website users with our affiliates in the Sony group family of companies and other third parties who assist us with fulfilling your requests, clear and verify transactions, deliver and administer products, content or services, manage and enhance customer data, store and maintain our database records, provide customer service, detect fraud or illegal activities, conduct customer research and surveys, develop new products and services and sell products and services to you.

... and what is “who” doing with your data?

We do not control our affiliates’ or third parties’ use of your information after we share it, but we use reasonable efforts to obtain our affiliates’ and third parties’ agreement to protect the confidentiality, security, and integrity of any personal information we share with them or that we permit them to collect directly. If consumers do not want their personal information made available to others in these ways, they should not provide their personal information to us.
... but do you realize what is declared?

... and what the consequences are concerning your privacy?
Many people do not!
The upcoming EU General Data Protection Regulation protects the privacy of EU citizens by means of different data protection principles.
Privacy is not a mere afterthought, but privacy safeguards must be built into all steps of the design process from the earliest design stage, i.e. privacy by design.
Obligation of data controllers to comply with regulations and to demonstrate this compliance and to implement mechanisms that ensure this compliance, i.e. service provider accountability
In the past, only functionality had to be considered during game development. Developers were able to handle this.
Today, system design is multi-disciplinary and requires expert-knowledge.
Need for computer-aided tools supporting the design of complex services.
Methodologies

A State-of-the-Art of Privacy Modeling Approaches
Privacy Requirement Engineering – identifying privacy requirements
  • Threat models
    • STRIDE/DREAD (Microsoft) – Data flow diagram
    • CORAS
  • Misuse cases – use cases
  • Attack trees
  • Problem frames

Quantitative approaches: metrics → measuring degree of anonymity
  • Anonymity networks
    • Anonymity sets
    • Information theoretic approaches – entropy based
  • Databases
    • $k$-anonymity $\rightarrow$ $l$-anonymity $\rightarrow$ $t$-closeness
  • Statistical databases:
    • Differential privacy $\rightarrow$ focus on used query algorithm
Quantitative Approaches

• **Policy-agnostic programming:** verification of program code compliance with privacy policies
  ➢ E.g. Jeeves, Hoare logic

• **Logic based modeling approaches:**
  • Conflict detection between privacy policies of entities in multi-tier systems
    ➢ E.g. Facebook apps
  • Conflict detection between privacy policies and privacy regulatory frameworks
  • Protocol verification

• **Markov chains:**
  • Verification if data collection serves certain goals

• **Process Algebras:** Applied $\pi$-calculus - ProVerif
  • Automated privacy analysis for protocols based on PETs
    ➢ E.g. e-voting system, e-auction system, electrical vehicle charging

• **Multi-paradigm:**
  • Designing controlled anonymous applications using ABCs.
Approach

A Logic Based Privacy Modeling Approach
A logic based modeling approach is used for the privacy analysis based on **user profiles** built from formal models representing the service under consideration. The feedback must be useful for **system designers** and **end-users** as well.
Privacy Modeling Concepts
Conceptual model of a composite service
Modeling properties of authentication technologies

X.509 Certificates

Calypso cards

Idemix

U-Prove
E.g. access to a service is only permitted if individual is older than 18y

Revealed attributes:
DoB, First name, Surname, SSN, Address, Gender, Card SN, ...

Revealed attributes:
e.g. DoB

e.g. Age > 18

e.g. 20 < Age < 25
Different types of users → Different trust perceptions

He just don’t care about his privacy, he trusts everybody

She cares about her privacy, only trusts companies with a good reputation
Trust perceptions are modeled in terms of storage and data forwarding (i.e. distribution)

An organization part of the set of storage/distribution trusted organizations, only stores/forwards the attributes that are declared in storage/distribution policy, else the organization is supposed to store/distribute all attributes it can collect.
Modeling identifiers linkable to an individual

**Pseudonymous:** group of attributes referring to individual without actual revealing his identity
  - e.g. e-mail address
  - e.g. username,
  - e.g. browser fingerprint, i.e. a unique combination of attributes representing browser configuration

**Identity:** group of attributes *sufficiently* revealing identity of an individual
  - e.g. first name, surname and address (in case of Homer)
  - e.g. first name, surname (in case of Lisa)
A Logic Based Modeling Framework for Analyzing Privacy
System Independent Model

- Vocabulary (Concepts)
- Behavior
- Inference Rules

Logic Component

User Model
- Trust Perception
- Initial State
  - Credentials
  - Profiles

System Model
- Organizations
- Services
- Service Policies
  - Access Control
  - Storage
  - Distribution
  - Output

Conclusions

Input Model

- Identifiability Model
  - Identities
  - Pseudonyms
Properties

• Declarative logic programming system
• Knowledge base system
• Intuitive modeling using predicate logic
• IDP language: FO logic enriched with types, aggregate, inductive definitions, partial functions
• Supports reasoning on incomplete knowledge
• Supports modular programming
• Integrated Lua

Structure of IDP program

• Vocabulary: contains non-logic modeling symbols = types, predicates, functions
• Theory: set of constraint rules and definitions
• Structure over the vocabulary: a partial valuation of the vocabulary elements that satisfies the theory
Model expansion = extending input model

Input model = partial structure satisfying theory

Output model = complete structure satisfying theory

Theory:
Generic behavior
Inference
Computation of User Profiles

Behavior
Inference Rules

Service Invocation Graph
Service Invocation Graph

Sub-graphs represent invoked services of different alternatives
Feedback that can be derived from user profiles

Query the user’s anonymity level:
→ Users is anonymous, pseudonymous, identifiable

Query the attributes released to organizations:
→ Detecting violations: e.g. an organization is not allowed to collect name and address
→ Verify if attributes required for the functionality can be collected by an organization

Query the impact of collaborations between organizations:
→ e.g. can a user be identified when organization x and y collaborate?

Querying the required trust between organizations:
→ e.g. Y receives name and address from X. X collects it from user’s X.509 based identity card. Y must only trust the issuer of the identity card.
Conclusions
• A qualitative modeling approach complementary to other approaches such as RE and quantitative approaches.

• Flexible approach → analyze privacy of services from different domains:
  o Travel reservation system
  o Web shop
  o Loyalty Systems
  o Ticketing systems in public transport

• Result are publicly available at: https://github.com/decroik/inspect-privacy-and-trust
• Privacy related feedback useful for designers
  o Impact of design decisions: e.g. using X.509 certificate instead of Anonymous credential (Idemix).
  o Impact of collaborations.
  o Automated conflict detection with privacy preferences of prototypical users.

• Privacy related feedback useful for end-users
  o Anonymity level
  o Conflict detection with personal privacy preferences. E.g. commercial organizations are not permitted to collect my SSN.
• Realization Accountability:
  o https://github.com/inferring-accountability/inferring-accountability
Main Publications


Questions