GDPR-Compliant Reputation System Based on Self-certifying Domain Signatures

Mirosław Kutyłowski, Jakub Lemiesz, Marta Słowik, Marcin Słowik, Kamil Kluczniak[1], Maciej Gebala

Wrocław University of Science and Technology, Wrocław, Poland

ISPEC, 2019

[1] currently Stanford University

Reputation system

Threats

GDPR

Solution architecture

Domain signatures

PADRE

PADRE1

PADRE2

PADRE3

Conclusions

Reputation systems

Reputation system

Purpose

estimate quality of service(s) or goods based on former experience of other people

Role of reputation systems

fundamental!

Reputation system
records

Typical contents of a reputation record

- evaluation object
- score and/or comments
- evaluation time
- author [optional]
- authenticating information [almost always missing]

Reputation systems assumptions

Assumption 1

It is unlikely that the crucial characteristics of the evaluation object change quickly in time. So the past experiences provide a good approximation what can be expected.

However, there are cases that the rogue parties build up a good reputation in order to cheat once the people start to trust them.

Reputation systems assumptions

Assumption 2

There is a certain degree of randomness and bias in the former reports, but taking into account many reviews compensates for the shortcomings of individual reports.

This may be untrue in case of systematic cyber attacks, troll farms, etc.

Reputation systems assumptions

Assumption 3

It is unnecessary to take all reviews into account. A random sample is enough.

In fact, the consumers read anyway the first few screens. A random sample is much better.

Alternative approach for reputation systems
trusted parties

Trusted evaluators

example: Stiftung Warentest from Germany

- non-profit organization
- comparative tests of consumer goods
- publishing the evaluation reports

Disadvantages

- lack of scalability,
- cost
- not suited for small scale cases

Reputation system

Threats

GDPR

Solution architecture

Domain signatures

PADRE

PADRE1

PADRE2

PADRE3

Conclusions

Threats for reputation systems

Threat: Deleting entries

Deleting entries

a moderator can delete reports

- sometimes **justified** (ethical issues, false informations, etc.)
- . . . however it can be **misused** for changing the evaluation outcome

bigskip

Problem

after deletion it is **impossible to judge** whether it was justified

Threat: Modifying entries

Modifying entries

- **blinding** some contents might be justified (e.g. personal data protection of third parties)
- . . . however this can be **misused**

Problem

- records can be **secured with digital signatures** but it means **(provable) lack of privacy** for evaluators
- distributed ledgers probably too expensive and too complicated

Threat: Flooding with biased reports

Flooding attack

Hide real reports in a big number of reports prepared by the attacker

- the attacker mimics a real diversity of views, mixing false and true data

Problem

- technique widely used in internet campaigns
- hard to fish out the fake reports,
  FRR and FAR is a problem

Threat: Sybil attacks

Scenarios

evaluation object:

- business done under a pseudonym
- after getting a bad reputation **restarting with a new pseudonym**

review author:

- reviews signed with a pseudonym
- **many pseudonyms used to increase own influence**
- ... or the **pseudonym changed in case of bad reputation**

Problem

Using real identities and digital signatures would solve the problem, but the users are **unlikely to give up their privacy**.

Threat: Unfair aggregation

Aggregating information

- the users are not likely to browse all reports
- so it seem to be useful to provide an **average score** and an **aggregated review**

Problem

- how to prove that aggregation was fair?

Threat: Information leakage

Privacy protection

identity of the evaluators should be under protection and not to be published

- preventing revenge for critical reviews
- preventing information misuse by third parties

Problem

for standard techniques: a **trade-off** between

- privacy of evaluators, and
- security and quality of evaluation records

Reputation system

Threats

GDPR

Solution architecture

Domain signatures

PADRE

PADRE1

PADRE2

PADRE3

Conclusions

General Data Protection Regulation

GDPR regulation

GDPR

- European General Data Protection Regulation:
- scope:
  - processing taking place in the EU
  - exporting data ...
  - activities concerning commercial services in the EU (regardless of processing site)
- GDPR concerns the filing systems (except for purely personal use)
- many other countries adopt similar rules ...

Personal data

GDPR Reputation System

M. Kutyłowski

Reputation system

Threats

GDPR

Solution architecture

Domain signatures

PADRE

PADRE1

PADRE2

PADRE3

Conclusions

Personal data

'personal data' means any information ***relating to an identified or identifiable natural person*** *('data subject'); an identifiable natural person is one who can be identified, directly or indirectly*

Problem

A system where

- evaluation object are services of identifiable persons,
- the evaluators are not fully anonymous

falls into the scope of GDPR.

The protected data need not to be sensitive. Example:

> *" I find the conference venue of ISPEC 2020 very nice – Miroslaw K."*

Profiling

GDPR definition of *profiling*

'profiling' means any form of automated processing of personal data consisting of the use of personal data

> ***to evaluate certain personal aspects relating to a natural person**,*
>
> *in particular to analyze or predict aspects concerning that natural person's*
>
> > ***performance at work**,*
> > ***economic situation,***
> > *health, personal preferences,*
> > ***interests**,*
> > ***reliability**,*
> > ***behavior**,*
> > *location or movements;*

So a reputation system falls into the category of "profiling", while profiling a central problem for GDPR.

GDPR principles

Data minimality principle

a system should not gather more data than it is necessary to achieve its purpose

Purpose limitation principle

"personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes"

Storage limitation

data *"kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed"*

GDPR principles

Integrity and confidentiality

*personal data shall be processed in a manner that ensures **appropriate security** of the personal data, including protection against unauthorized or unlawful processing [. . .] using **appropriate technical or organizational measures**.*

Accountability

*The controller shall be responsible for, and be **able to demonstrate compliance** with [the principles stated in GDPR]*

GDPR
obligations of the parties running the system

GDPR Reputation System
M. Kutyłowski
Reputation system
Threats
GDPR
Solution architecture
Domain signatures
PADRE
PADRE1
PADRE2
PADRE3
Conclusions

1. Taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, the controller and the processor shall implement appropriate technical and organizational measures to ensure a level of security appropriate to the risk, including inter alia as appropriate:
   - (a) the pseudonymisation and encryption of personal data;
   - (b) the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;
   - (c) the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident;
   - (d) a process for regularly testing, assessing and evaluating the effectiveness of technical and organizational measures for ensuring the security of the processing.
2. In assessing the appropriate level of security account shall be taken in particular of the risks that are presented by processing, in particular from accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to personal data transmitted, stored or otherwise processed.
3. ...
4. The controller and processor shall take steps to ensure that any natural person acting under the authority of the controller or the processor who has access to personal data does not process them except on instructions from the controller, unless he or she is required to do so by Union or Member State law.

Consequences

Severe legal risks for running reputation systems: it's hard to fulfil all obligations with standard techniques

Reputation system

Threats

GDPR

Solution architecture

Domain signatures

PADRE

PADRE1

PADRE2

PADRE3

Conclusions

Our solution architecture

Traditional architecture

Centralized architecture

- the data are collected, processed, stored and presented by a single (trusted) organization
- all obligations and risks are concentrated there

Problems

- the *right-to-be-forgotten*
  – hard to balance the rights, frequently a complicated legal issue
- information obligations
  – on data subject's request a full report must be presented

Proposed architecture

Reputation record

- **kept by the evaluation subject himself**
  no need to report the data to physical persons
- **secured against manipulations**
- **a random sample over all transactions**
  a random sample has advantages even regarding reliability over a full report or an aggregated records
- **the evaluators pseudonymized but their identity may be uncovered** in case of law enforcement
  protection of evaluators' privacy and protection against misuse of anonymity

Typical interaction

A provides a service for *B*

1. *A* presents its reputation record

Typical interaction

A provides a service for *B*

1. *A* presents its reputation record
2. service or product provided by *A*

Typical interaction

GDPR Reputation System

M. Kutyłowski

Reputation system

Threats

GDPR

Solution architecture

Domain signatures

PADRE

PADRE1

PADRE2

PADRE3

Conclusions

A provides a service for *B*

1. *A* presents its reputation record
2. service or product provided by *A*
3. *B* computes its *domain specific pseudonym*, creates a report and a domain signature

Typical interaction

A provides a service for *B*

1. *A* presents its reputation record
2. service or product provided by *A*
3. *B* computes its *domain specific pseudonym*, creates a report and a domain signature
4. a pseudorandom deterministic value *i* derived

Typical interaction

A provides a service for *B*

1. *A* presents its reputation record
2. service or product provided by *A*
3. *B* computes its *domain specific pseudonym*, creates a report and a domain signature
4. a pseudorandom deterministic value *i* derived
5. depending on *i*, party *A* may be obliged to update its reputation record

Remarks

- *A* cannot predict if it will be necessary to update its reputation record
- *B* cannot change *i* and enforce including its evaluation report in the reputation record of *A*

Pseudonymous signatures

Domains

Domains

- domains correspond to disjoint activity areas
- each domain holds a public key which is created in an interaction with the Issuer [a]

[a] there is no corresponding secret key used by the domain, there are also schemes for ad hoc domains with no domain public keys

Remarks

- **in our application scenario each evaluation object defines a domain**

Signers

Joining the system

- each user must be registered by the Issuer
- by running the registration procedure a user gets
  - a private signing key
  - its master certificate [a]

[a] specific to the scheme used in this paper

Creating domain specific pseudonyms

Creating a pseudonym

for a domain D, a user A can create

- a single (domain specific) pseudonym $D(A)$
- a certificate for $D(A)$ [a]

the private key and the master certificate of A must be used

[a] specific to the scheme used in this paper

Creating domain specific signatures

Reputation system

Threats

GDPR

Solution architecture

Domain signatures

PADRE

PADRE1

PADRE2

PADRE3

Conclusions

Creating a signature

a signature corresponds to

- the signed message
- the **domain's public key**
- **the domain specific pseudonym** of the signatory

The signature can be created only with a private key resulting from the registration procedure

Creating domain specific signatures

Creating a signature

a signature corresponds to

- the signed message
- the **domain's public key**
- **the domain specific pseudonym** of the signatory

The signature can be created only with a private key resulting from the registration procedure

Signature verification

- input: ... , the domain public key, the **domain specific** pseudonym and certificate,
- the result should be `invalid` if the signature was created for a different domain or pseudonym

Main Properties

Single key per user

A user holds **a single signing key** and a **single master certificate** [a]

[a] the signing key is universal as it is not known with whom the user will interact

Cross domain unlinkability -informally:

it is infeasible to determine whether two pseudonyms in different domains belong to the same person
even if signatures corresponding to them are available

exceptions:

- when the singing key is known, or
- the deanonymization trapdoors to the domain public keys are used

Pseudonymous signature scheme used

Marcin Słowik, Marta Wszoła:
An efficient verification of CL-LRSW signatures and a pseudonym certificate system.
ACM ASIA Public Key Cryptography. APKC'17

A few properties

- based on Camenisch-Lysyanskaya LRSW signatures
- certificates that can be re-randomized by the user
- pairing groups used

Reputation system

Threats

GDPR

Solution architecture

Domain signatures

PADRE

PADRE1

PADRE2

PADRE3

Conclusions

Privacy Aware Distributed Reputation Evaluation

PADRE1
Design highlights

Reputation tables

- each evaluated party holds two 1-dimensional tables: $\mathcal{N}$ for negative scores and $\mathcal{P}$ for positive scores.
- the size N of the tables is constant

Preparing an entry by B about A:

B computes

- $nym_{A,B}$ – the pseudonym of B for the domain of A,
- a signature s for: $nym_{A,B}$, $b \in \{0,1\}$ (score), t (transaction time),
- $i := \mathcal{H}(nym_{A,B}) \bmod N$,
- an entry $\eta := (nym_{A,B}, t, b, s)$

η inserted on position i into $\mathcal{N}$ (if $b = 0$), or into $\mathcal{P}$ (if $b = 1$)

PADRE1
properties

Main features

- each entry authenticated with a pseudonymous signature
- a user can insert a new score, but always at the same position (one cannot flood the tables)
- the stored transaction times give a rough estimation of the number of insertions in a table
- one can separately estimate the number of entries not older than Δ

PADRE1
properties

Main features

- each entry authenticated with a pseudonymous signature
- a user can insert a new score, but always at the same position (one cannot flood the tables)
- the stored transaction times give a rough estimation of the number of insertions in a table
- one can separately estimate the number of entries not older than Δ

Estimator

- concern the time period $[T_0 - \Delta, T_0]$, where T_0 is the current time,
- let $Y_\Delta =$ the number of positive scores entered in this period in table $\mathcal{X}$
- calculate V_Δ - the number of positions in $\mathcal{X}$ with $t \in [T_0 - \Delta, T_0]$
- $\bar{Y}_\Delta$ is an unbiased estimator of Y_Δ:

$$\bar{Y}_\Delta = -N \ln \frac{N - V_\Delta}{N} .$$

PADRE-2

Changes over PADRE-1

- an entry prepared as $E = (nym_{A,B}, h, b, s)$ with **the new component** h:

$$h = H(nym_{A,B}, s)$$

- insertion strategy (e.g. if $b = 1$)
  - if still there is an empty place in $\mathcal{P}$, then insert E in this place
  - else:
    1. if h in E is higher than the 2nd component of each entry stored in $\mathcal{P}$, then drop E,
    2. else: E replaces the entry in $\mathcal{P}$ with the highest h component.

Properties

Estimation of the number of entries

- let n denote the number of attempts to write a record E in table $\mathcal{X}$,
- over all entries E only N of them with the lowest h component are stored,
- each h component may be regarded as a random number $\in (0, 1)$,
- let u be the highest component h stored
- the estimator on n is

$$\hat{n} = \frac{N-1}{u} .$$

- Now older reviews are more likely to be present in the table.
- There is a better overview of the whole reputation history.
- The price is that the recent entries are less frequently represented.

PADRE3

Sketch

- N different registers
- in each register just one entry chosen in a pseudorandom way
- the choice depends deterministiccaly on the component h
- the choices in different registers are independent and may follow different probability distributions
- a wide range of choices for probability distributions: e.g. uniform, exponential, ...

now we assume that the counter for the number of evaluation results is maintained and we focus on a random sample

High level conclusions

It is possible to create a profiling system compliant with the GDPR regulations.

Protection of personal data is not based on organizational means. Instead, there are technical guarantees with provable properties.

Technical conclusions

Pseudonymous signatures and domain specific pseudonyms is a quite universal tool and source of pseudorandomness in cryptographic protocols.

Reputation system

Threats

GDPR

Solution architecture

Domain signatures

PADRE

PADRE1

PADRE2

PADRE3

Conclusions

Thanks for your attention!