Privacy risk analysis in the IoT domain

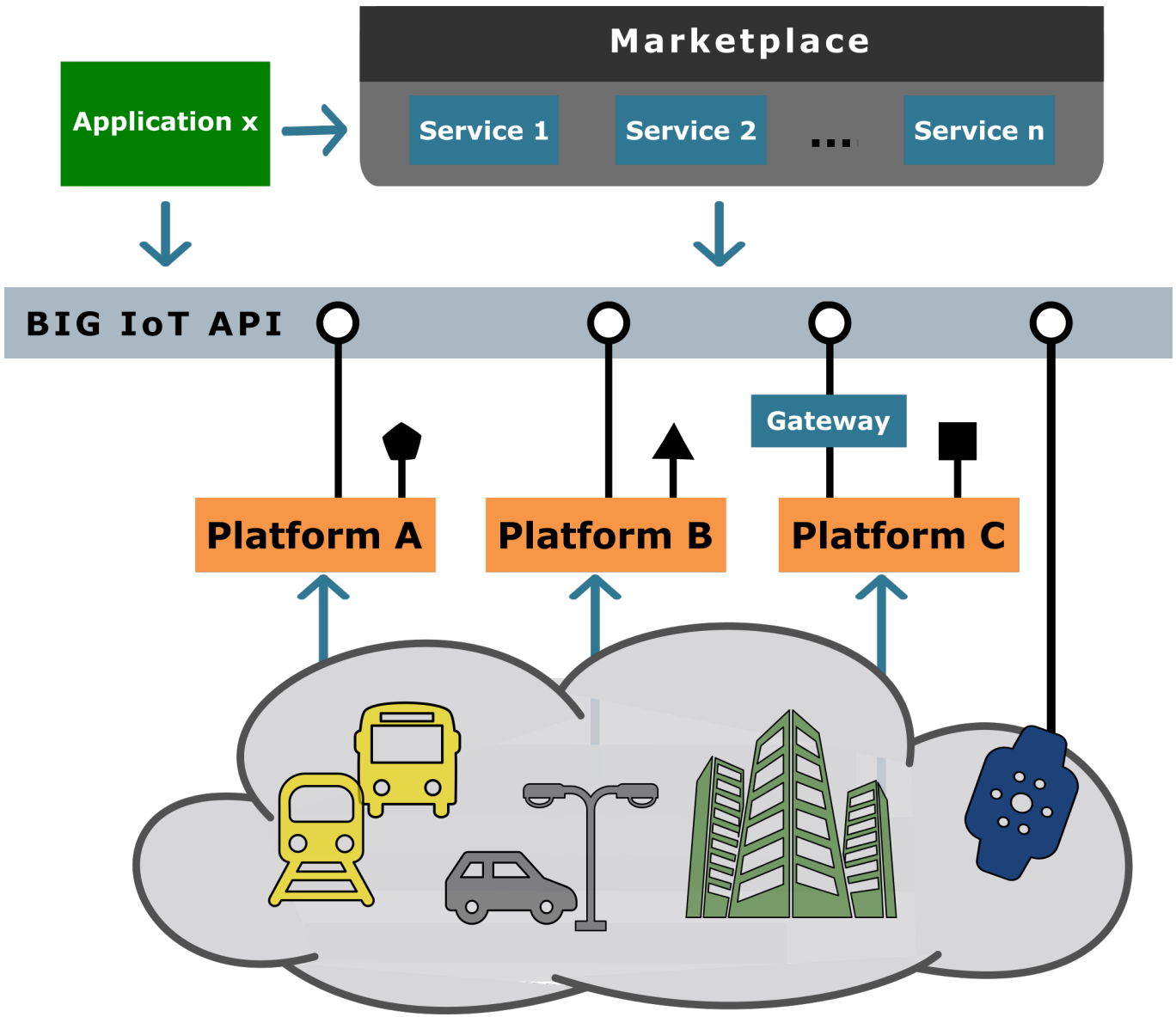
Juan Hernández-Serrano, Jose L. Muñoz, Olga León Lars Mikkelsen
Hans-Peter Schwefel & Arne Bröring
BIG IoT

Outline

Goal

- Most IoT systems exchange user data:
  - Privacy is a key factor
  - Is challenging because systems are comprised of many software and hardware distributed components
- This paper:
  - Proposes a risk rating methodology for identifying and rating privacy risks.
  - We also demonstrate how to apply this methodology in an IoT use case set in the context of the EU H2020 BIG IoT project.

BIG IoT


Marketplace
Application x
Service 1
Service 2
...
Service n
BIG IoT API
Gateway
Platform A
Platform B
Platform C

Privacy in SDLC

- We have to consider privacy in the software development life cycle (SDLC).
- Identify risks and rate these risks.
- State of the Art:
  - Risk analysis in IoT: **JBC16RiskAnalysis**, **savola2015riskEhealthIoT**, **tai2015rrmWaterIoT**.
  - Privacy issues, challenges and implications: **Arabo2012PrivacyIT**, **Kozlov2012SecPrivIoT**, **WEBER2011AccountabilityIoT**.
- Not yet:
  - A clear methodology for the analysis and integration of privacy-related risk assessment results into the SDLC.
  - Properly assigning a priority to development issues/actions derived from the analysis.

Outline

RRM for IoT

- The OWASP RRM (Risk Rating Methodology) covers:
  1. Evaluation of the potential attacker
  2. The exploit evaluation and detection
  3. Evaluation of the technical and business impact

- In IoT:
  - Many small and low resource sensor units
  - Much data transfer and data aggregation
  - Services and applications using these data

Contribution I

- We take risk factors of OWASP RRM that are relevant for our scenario
- Propose new factors
- We assign weights to each factor from 1 to 9
- Using the weights and the scores, we obtain average values for:
  1. **Likelihood** factors: threat agent factors combined with vulnerability factors
  2. **Impact** factors: technical impact factors and business impact factors
- Values of factors are divided into 3 levels:
  1. Low: 0 to 3
  2. Medium: 3 to 6
  3. High: 6 to 9

Contribution II

Risks are evaluated by combining impact and likelihood:

	Likelihood		
Impact	low	medium	high
high	**medium**	**high**	**critical**
medium	**low**	**medium**	**high**
low	**negligible**	**low**	**medium**

Outline

Vulnerability factors

- **Ease of discovery**. Weight: 0.25.
  (1) Practically impossible
  (3) Difficult
  (9) Automated tools available
- **Ease of exploit**. Weight: 0.25.
  (1) Theoretical
  (3) Difficult
  (9) Automated tools available
- **Awareness of the exploit**. Weight: 0.25.
  (1) Unknown
  (4) Hidden
  (9) Public knowledge
- **Intrusion detection**. Weight: 1.
  (1) Active detection in application
  (3) Logged and reviewed
  (9) Not logged

Threat agent factors

- **Skill level of the attacker**. Weight: 1.
  (1) Security penetration skills
  (5) Advanced computer user
  (9) No technical skills
- **Attacker motivation** Weight: 0.5.
  (1) Low or no reward
  (4) Possible reward
  (9) High reward
- **Opportunity to perform the exploit**. Weight: 1.
  (0) Full access or expensive resources required
  (4) Special access or resources required
  (9) Can be done by readily available off-the-shelf equipment
- **Size (potencial victims)**. Weight: 1.
  (1) Just one
  (5) Tens of individuals
  (9) Thousands

Technical impact factors

- **Loss of privacy**. Weight: 1.
  (2) Minimal non-sensitive data disclosed
  (6) Minimal critical data disclosed to extensive non-sensitive data disclosed
  (9) Personal information that is directly linked to an individual

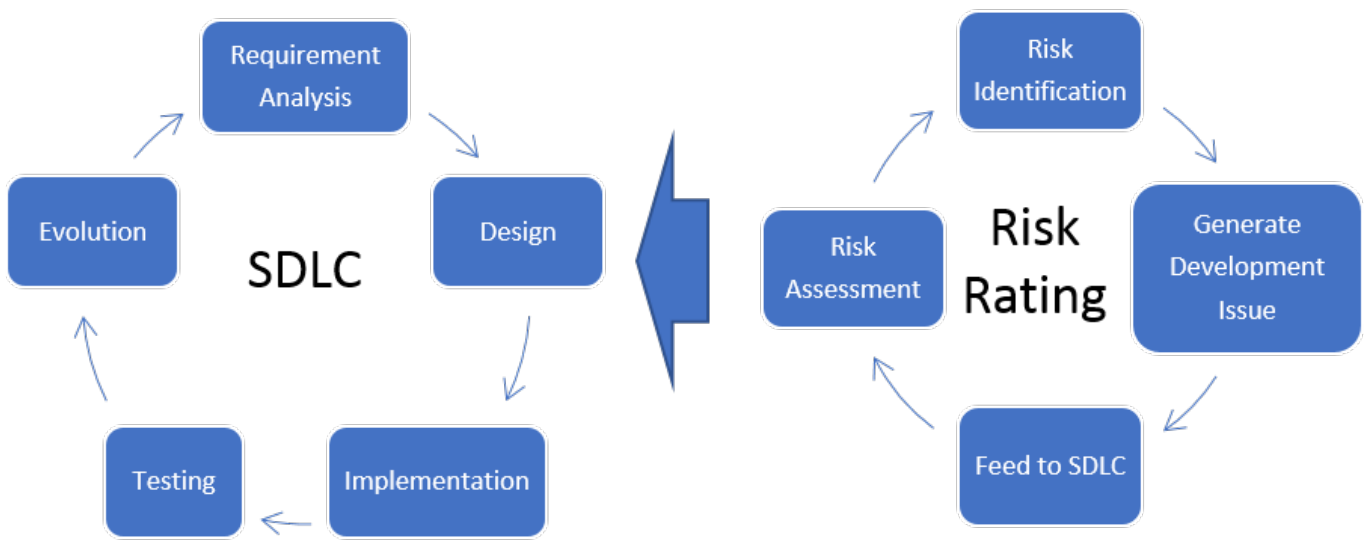
- **Loss of attacker accountability**. Weight: 0.5.
  (1) Fully traceable
  (7) Possibly traceable
  (9) Completely anonymous

Business impact factors

- **Financial damage**. Weight: 0.5.
  (1) Less than the cost to fix the vulnerability
  (3) Minor effect on annual profit
  (9) Bankruptcy
- **Reputation damage**. Weight: 0.5.
  (1) Minimal damage
  (5) Loss of goodwill
  (9) Brand damage
- **Privacy violation scale**. Weight: 1.
  (1) Minimal scale
  (5) Few hundreds
  (9) Millions

Integration with the SDLC

- The last step in the risk evaluation process is to feed the results to the development cycle.
- So risks can be mitigated during the next development iteration.



Outline

Environment Monitoring Service

- Environment Monitoring Service (EMS) is a BIG IoT use case.
- EMS provides data from **pollution** sensors in cars and **noise** metering stations at different city spots.
  1. **Selected private cars**. BOSCH's members in BIG IoT have installed Bezirk[1] devices in several cars provided by SEAT to the project. Installed devices can query the internal pollution sensors and offer the readings as well as the current position of the car.
  2. **Noise metering stations**. Cities commonly have a deployment of noise stations. These stations report current level of noise in specific spots. Noise data is used in the BIG IoT ecosystem to compute green routes and drive vehicles through the less-noisy paths. This fact allows to distribute the noise and keep it under a certain threshold all over the city.

[1] Bezirk is s Startup from Bosch that was established to cater to agility in a highly dynamic environment.

EMS Risks

1. **Vehicle/people tracking (*).**
   For cars are sharing their position, an attacker could track vehicles if the shared data is not properly anonymised. Even with no identifiers, an attacker could track a vehicle computing viable trajectories based on the pairs position-time. Moreover, if a vehicle is linked to a person (e.g. by visual contact), people can be actually tracked.
2. **Accessing/Hacking the in-vehicle internal bus.**
   Since the Bezirk “thing” is connected to the in-vehicle information bus, if an attacker gains access to the Bezirk device, it could also try to hack the internal bus and thus eavesdrop internal/not public vehicle data and even disrupt the vehicle operation by mangling those data.
3. **Noise level disruption.**
   If the attacker can tamper noise data, he can influence route recommendations since the noise is a factor in these recommendations. For example, the attacker can report that an area has no noise at all. Then, many vehicles would be routed through this area (and effectively, the attacker is creating a noisy area).

Overall Likelihood I

- **Skill level**. An attacker with some technical skills can track vehicles using ids provided by the EMS.
  Weight: 1 Score: 7
- **Motivation**. Tracking could allow spying on specific people. Depending on the target, the motivation could be high (potential reward).
  Weight: .5 Score: 5
- **Opportunity**. The attack could be operated from any standard computer connected to the EMS.
  Weight: 1 Score: 9
- **Size**. The attack could affect thousands to millions of people.
  Weight: 1 Score: 9
- **Ease of discovery**. It's easy for attackers to detect by basic input/output verification that the EMS is providing identifiers or pseudo-identifiers along with the pollution and position data.
  Weight: 0.25 Score: 7

Overall Likelihood II

- **Ease of exploit**. Once the attacker has an array of positions and dates it is a matter of simple data processing to get the path of the victim.
  Weight: 0.25 Score: 8
- **Awareness**. Since the format of data provided by the service is published on the BIG IoT marketplace as an offering description, it is obvious to be aware of the presence of identifiers or pseudo-identifiers.
  Weight: 0.25 Score: 6
- **Intrusion detection**. The attack will not be logged since the attacker just consumes the same data as any other standard consumer.
  Weight: 1 Score: 9

Final score: **7.95 (HIGH)**

Technical Impact

- **Loss of privacy**. An attacker can potentially track a person if it could be linked to a vehicle by another external source (e.g. by direct visual observation or by publicly accessible street cameras). Therefore, the attacker would be able to infer about the behavior about a person or a group of people.
  Weight: 1 Score: 7
- **Loss of attacker accountability**. Since the EMS is an authenticated service, based on the attacker queries, the attacker could be traceable.
  Weight: 0.5 Score: 5

Final score: **6.33 (HIGH)**

Business Impact

- **Financial damage**. The knowledge of the vulnerability would probably encourage people to leave the service or to explicitly demand to not be accounted. Therefore, a strong effect on annual profit should be expected.
  Weight: 0.5 Score: 7
- **Reputation damage**. A loss of goodwill is expected in the short term regarding to this service. A full brand damage is not expected since most of the affected individuals won't see affected their privacy.
  Weight: 0.5 Score: 5
- **Privacy violation scale**. While thousands to millions tracking data of vehicles will be disclosed. Linking a vehicle identifier with a person requires very specific external information (e.g. visual contact). Therefore, the breach would likely affect no more than hundreds of people.
  Weight: 1 Score: 5

Final score: **5.50 (MEDIUM)**

RRM Result

Overall likelihood	high (7.95)
Technical impact	high (6.33)
Business impact	medium (5.50)
Technical risk severity	critical
Business risk severity	high

Actions

- Data provided by the cars should be anonymised.
- Identifiers or pseudo-identifiers should not be used.
- In order to mitigate these risks the following actions should be carried out[2]:
  1. Instead of providing a measure at a given position (with GPS accuracy), provide pollution at a given area (e.g. street segment).
  2. Review device and service code to avoid sending potential identifiers or pseudo identifiers from the cars to the EMS.
  3. Review service code to ensure that no potential identifiers or pseudo-identifiers are stored and/or provided by the EMS.
  4. Check data already stored by the EMS (before implementing action 3) to filter out identifiers or pseudo-identifiers.

[2]The four actions became issues that got into the SDLC of the EMS with priority `critical`, since in BIG IoT we have agreed to use the worst-case scenario (either technical or business risk).

Outline

Conclusions

- With IoT becoming more widespread and integrated as parts of big complex systems, the aspect of maintaining the privacy of data is more relevant than ever.
- This work proposes an approach for identifying and rating privacy risks in the IoT domain.
- The approach is exemplified by applying it to a typical IoT use case namely the Environment Monitoring Service.
- Future work include applying the proposed RRM in more use cases and evaluating its applicability.
- Furthermore, while doing this it should be formalized how the remediation actions that is the outcome of the RRM should be included in the SDLC.