# IoT Technology Tracking and Industrial Requirements for Security, Privacy and IPR/Data Protection

<table>
<thead>
<tr>
<th>Project Acronym</th>
<th>IoT4Industry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Document-Id</td>
<td>D.2</td>
</tr>
<tr>
<td>File name</td>
<td></td>
</tr>
<tr>
<td>Version</td>
<td>FINAL document</td>
</tr>
<tr>
<td>Date</td>
<td>Start: 01 October 2014</td>
</tr>
<tr>
<td></td>
<td>End: 31 September 2015</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Dr. Violeta Damjanovic-Behrendt</td>
</tr>
<tr>
<td></td>
<td>Robert Bob Mulrenin</td>
</tr>
<tr>
<td>QA Process</td>
<td>Georg Güntner</td>
</tr>
</tbody>
</table>
Table of Content

IoT4Industry in a Nutshell
IoT4Industry Task 2 Description
1. Problem Statement
   2.1. Privacy and Data Protection Challenges in the IoT
       2.1.1 Privacy models
       2.1.2 Privacy mechanisms
       2.1.3 Data Access Control in the IoT
   2.2. Security Challenges in the IoT
       2.2.1 Top security vulnerabilities in the IoT
       2.2.2 Detecting and preventing attacks in the IoT
       2.2.3 Securing Data Provenance in the IoT
           2.2.3.1 Data Provenance and Sensor Ontologies
           2.2.3.2 PROV-DM, PROV-AQ and PROV-O
           2.2.3.3 PROV-O and SSNO Ontology Alignment
       2.2.4 Game Theory to support IoT security
           2.2.4.1 Preventing DoS Attacks
           2.2.4.2 Intrusion Detection
           2.2.4.3 Strengthening Security
           2.2.4.4 Coexistence with Malicious Sensor Nodes
3. Industrial Requirements w.r.t. IoT Security, Privacy and IPR/ Data Protection
4. Conclusion
Appendix 1: Interview with Industrial Experts in Austria
References
IoT4Industry in a Nutshell

IoT4Industry is an exploratory project funded by the Austrian Research Promotion Agency, for the period October 2014 – September 2015. It stands for “Secure, Privacy-preserving Agents for the Industrial Internet”. The core research areas of IoT4Industry relate to security, privacy and IPR/data protection requirements, translated into game-theoretic agent behavior, which is simulated in a demo factory floor environment, supporting a supply chain negotiation. A demo factory floor is implemented within our IoT-lab, which is located at Salzburg Research.

While it is still not clear which protocols and technologies will become mainstream for the Industrial Internet, it has become clear that security and privacy will feature strongly in any risk assessment concerning the adoption of practices using the Industrial Internet. Thus, in IoT4Industry, we focus on the use of policy-enacting, multi-agent systems that securely manage machines and manufacturing cells, and we build a feasibility demonstrator based on open source tools and firmware. In addition, IoT4Industry helps us to prepare for internationally recognized contributions to science and technology in the field of Industrial Internet, notably in Horizon 2020.
IoT4Industry Task 2 Description
(from the IoT4Industry project proposal)

**D2: IoT Technology Tracking and Industrial Requirements for Security, Privacy and IPR/Data Protection**

**Goals:**
1. Report on IoT frameworks as they emerge
2. Catalogue of industrial needs w.r.t. IoT security
3. Understanding of security concerns from industrial practitioners

**Description of the content**
We will start our investigation by inviting three experts from production companies in order to elicit their expectations and security needs in an IoT/Industrial Internet scenario. These specific views will be balanced with desktop research on emerging requirements, gleaned from the on-line discourse in the field, and from following the relevant literature. In parallel with the requirements gathering (technology pull), we will also track developments in the technical and scientific fields, to maintain an up-to-date view of Industrial Internet throughout the duration of the project.

**Method**
a) Three interviews with invited practitioners to obtain an initial understanding of industry needs
b) Technology tracking activity over the course of the project to keep knowledge of emerging IoT platforms up to date
c) Desktop research following the scientific/technological discourse concerning industrial requirements and expectations vis-à-vis the Industrial Internet

**Milestones, results and deliverables**
01/2015: Technical report on industrial requirements is published
09/2015: Technology scouting report on emerging IoT platforms for Industrial Internet
1. Problem Statement

A significant growth in the number of objects that are connected with the existing Internet infrastructure through various standards and communication protocols, leads to a new technological paradigm, known as the **Internet of Things (IoT)**. The research area of the IoT embraces diverse technological fields such as electronic engineering, mobile and wireless communication, computer science, which are additionally supplemented by technologies such as Wireless Sensor Networks (WSN), Machine-to-Machine (M2M) interfaces, RFID, and more. The idea of the IoT paradigm is to connect zillions of objects from both physical and virtual environments, to provide data exchange between those objects, and support decision making in a more federated way, in order to create smart environments for logistics, environmental monitoring, manufacturing and industry.

However, the challenges around the deployment of the IoT are many, ranging from IoT architectures, sensor development, real time processing, data analytics and storage, services, and communication, to the privacy and security of services, processes, individuals and objects involved in IoT. For example, a recent report on the IoT, published in January 2015 by the Federal Trade Commission (FTC) “… *urges companies to adopt best practices to address consumer privacy and security risks.*” The FTC recommends a series of concrete steps that businesses need to take in order to protect consumers’ privacy and security [FTC15]:

- *build security into devices at the outset, rather than as an afterthought in the design process;*
- *train employees about the importance of security, and ensure that security is managed at an appropriate level in the organization;*
- *ensure that when outside service providers are hired, that those providers are capable of maintaining reasonable security, and provide reasonable oversight of the providers;*
- *when a security risk is identified, consider a "defense-in-depth" strategy whereby multiple layers of security may be used to defend against a particular risk;*
- *consider measures to keep unauthorized users from accessing a consumer's device, data, or personal information stored on the network;*
- *monitor connected devices throughout their expected life cycle, and where feasible, provide security patches to cover known risks.*
Furthermore, the authors in [FTC15] recommend the use of the **data minimization approach** for the companies, by limiting the collection of consumer data, and retaining the necessary information only for a set period of time: “...*companies can choose to collect no data, data limited to the categories required to provide the service offered by the device, less sensitive data; or choose to de-identify the data collected.*” This is not an easy step in IoT, where anyone can access ubiquitous services anyway and anytime, process information obtained from heterogeneous resources, and exchange the results interpreted from various perspectives. Therefore, the design and deployment of efficient security and data privacy mechanisms in IoT is considered to be of great importance.

In parallel, the real success and advancement of IoT systems depend on various security and privacy aspects of the IoT, bringing the protection of sensitive data exchanged through IoT as one of the biggest IoT challenges today. For example, the user data can be exposed to unlawful surveillance, data profiling, intrusion in private life, unauthorized access to personal data, infection by malware, etc. [WP29-14]. At the same time, user might easily lose track of the existing interconnected devices, or on intended usage of data created by IoT systems. Similarly, the industrial organisations strive to ensure that their services are equipped with the privacy-friendly default mechanisms, as expected by the EU citizens.

**Our motivation in IoT4Industry** is threefold: (i) to understand the IoT conceptually, (ii) to overview and elaborate on the current state of Internet security and privacy, and (iii) to explore how to bridge the current constraints with requirements that can assure security, privacy and data protection in the IoT. Above all, we want to investigate current technologies and challenges in privacy and cybersecurity with the aim to help users in industry gain a better understanding of IoT potentials, technological benefits, and dangers.

Herein, we explore various mechanism with the potential to control IoT devices, and we particularly focus on those solutions that could benefit from **knowledge-driven approaches**: for example, Semantic Web technologies that can facilitate data aggregation, cross-referencing, data analytics, and enable services and applications to make better decisions. However, quality of data may vary dramatically due to **data obfuscation** (c.f. hiding data due to policy constraints). To assess the quality, reliability and trustworthiness of connected objects and their data, we would explore the role of **data provenance** (information about contextual evidence of entities, services, objects in the IoT) based on the W3C PROV-O ontology (provenance ontology) and its data model supporting the interoperable interchange of provenance information in heterogeneous environments (c.f. [FTC15], [WP29-14]).
http://www.w3.org/TR/prov-overview/).

**Document organization.** After a brief description of the problem statement, we continue this report overviewing the state of the art in privacy, security, Intellectual Property Rights (IPR) and data protection in the IoT. **Section 2.1** explores the main privacy and data protection challenges in the IoT, which are related to privacy models, privacy mechanisms and data access control. **Section 2.2** discusses security challenges in IoT, including security vulnerabilities, the detection and prevention of attacks in IoT, data provenance and, game-theoretic approach to IoT security, by differentiating between game-theoretic model used to (i) prevent Denial of Services (DoS) attacks, (ii) detect intrusion, (iii) strengthen security, and (iv) detect coexistence with malicious sensor nodes. Please note that our next report, D.3 "Specifying Game-Theoretic Behaviour for Agents in Industrial Supply Chain", presents an additional set of game-theoretic algorithms targeting supply chain negotiation rather than privacy and security. Our challenge in D.5 “Security, Privacy, and IPR/Data Protection Requirements Translated into Game-Theoretic Agent Behaviour” is to work on consolidation of GT algorithms from both privacy and security area and from supply chain management field, in order to find a common GT model to be applied in IoT4Industry. **Section 3** looks at the main industrial requirements on security and privacy in IoT, which is based on the following: (i) our discussion with industrial partners during the two IoT Salzburg meetups held in March 10, 2015 and May 19, 2015 (c.f. http://www.meetup.com/IoT-Salzburg/), (ii) recently published IIRA (Industrial Internet Reference Architecture), (iii) the RAMI (Reference Architecture Model Industry 4.0), and (iv) the SCISSOR project (Horizon 2020 N644425, February 2015 - January 2018), especially SCISSOR’s report “Technological assessments, scenarios and requirements”. Salzburg Research is the project partner in SCISSOR, with the role to develop a data source registration service to which semantic capability descriptions can be submitted in a form of metadata about the registered sensors and actuators. Based on this registry, semantic browse and search services will be offered for further analysis and decision making. SCISSOR analyses the requirements related to security services, ranging from industrial control processes, pipelines, power generation, nuclear plants, water distribution, ambient control, etc. The analysis covers the real-world expertise of the consortium partners related to the ICS (Industrial Control System) technologies, devices, and network infrastructure. Finally, we would like to mention the “Global IoT Day Vienna” event, 09. April 2015 (for more: http://www.iot-vienna.at/global-iot-day-event/2015/doku.php).
Section 4 gives some conclusion remarks. Note that the document given in Appendix 1 of D.2 is the IoT4Industry interview template, which has been created to serve as a basis for the discussion with the industrial experts in Austria, during IoT Salzburg meetups.

The IoT allows people and objects (“things”), as well as data and virtual environments, to actively participate in virtual process, and interact with each other through various “technologically smart” environments, improving the quality of people’s lives, and increasing economic growth. However, the IoT introduces new challenges for both the security of systems and processes (e.g., water distribution) and the privacy of individuals (e.g., purchasing preferences, health condition). For example, the interaction of people with the IoT based environment calls for the protection of their privacy, while the interaction with the services and process in IoT calls for their safety. Hence, in this section, we address the privacy and security in IoT, perceiving possible threats and the ways of preventing attacks. Note that the following state-of-the-art analysis on security, privacy and data protection, described in Section 2.1, is based on the WP29's¹ document [WP29-14], which identifies the main data protection risks in the context of the EU legal data protection framework. In addition to WP29 document, we consider additional literature discussing challenges and risks related to the IoT systems, such as privacy preservation technologies in IoT, cyber security in IoT, etc.

2.1. Privacy and Data Protection Challenges in the IoT
The existing privacy and security mechanisms in the IoT, such as lightweight cryptography, privacy assurance, and secure protocols, are not protective enough. The authors in [RONL11] emphasize that “… researchers … must analyze current security protocols and mechanisms and decide if such approaches are worth integrating into the IoT as is or if adaptations or entirely new designs will better accomplish security goals.” The authors in [WP29-14] identify the following privacy and data protection challenges in the IoT:

¹ WP29 is an independent European advisory body on data protection and privacy.
• **Lack of data control and information asymmetry:** To prevent situations in which users experience loss of control over data, as a result of multiple interconnections and interactions in the IoT, user’s data flows need to be collected and processed in a transparent manner, ensuring privacy and data protection.

• **Quality of the user’s consent:** Many IoT devices (objects) are designed neither to provide information about their nature (e.g., many IoT devices cannot be distinguished from “non-connected” ones) nor to provide a mechanism for obtaining the user’s consent. At the same time, these devices can embed various motion sensors, cameras, microphones, etc., which can record and transfer data without users being aware of it. Such lack of information constitutes a significant barrier to demonstrating valid consent under EU law, according to which the data subject must be informed about the usage of their data. Some possible ways to offer a data subject real control over her data could be supported using:
  - *Privacy Proxies*, which define sensor-policy pairs that authorize, limit or reject access to sensor data by third parties;
  - *Sticky Policies*, as a way to establish the context of use of the data, the purposes, policies on third party access and a list of trusted users.

• **Inferences derived from data and repurposing of original processing:** The IoT systems allow for *sensor fusion*, which is a process of using sensor data collected by other parties and from different sources, for secondary uses. For example, the gyroscope from a smartphone can be used to reason about the individual’s driving habits, or motion sensors that count the number of steps, which can be used for reasoning about the user’s physical condition. Hence, it is important for IoT stakeholders to assure the users that the data is used only for predefined and known purposes.

• **Intrusive bringing out of behaviour patterns and profiling:** Generating knowledge from various sensor sources, analysing and reasoning on top of sensor data, can reveal private aspects of user’s habits, user’s behaviour patterns and preferences. Such a trend might be very intrusive on the private life and the intimacy of individuals and must be closely monitored.

• **Limitations on the possibility to remain anonymous when using services:** Clearly, remaining anonymous and preserving privacy in the IoT becomes increasingly difficult with the current proliferation of sensors. For instance,
wearables kept in close proximity of data subjects, result in the collection of multiple MAC addresses of multiple sensor devices, which can be used for a range of purposes, including location analytics or the analysis of movement patterns of crowds and individuals.

2.1.1 Privacy models
Regarding techniques designed to protect user privacy, we would like to mention models such as the k-anonymity [SWEE02], data transformation/ randomization, the cryptography, and cryptographic techniques, such as Secure Multiparty Computation.

The k-anonymity model distinguishes the following three entities [SWEE02]:
- individuals, whose privacy needs to be protected,
- the database owner, and
- the attacker.

This model assumes both the domain of the data (the nature of the domain, its attributes and their values, e.g. a height of an individual) and the algorithms used for anonymization are known to attackers. The k-anonymity model of privacy was studied intensively in the context of public data releases in [AFKM05] [BAAG05] [BOYD05] [IYEN02] [SWEE02]. The authors in [SEN11] introduce the idea of Secure Multiparty Computations (SMC) in the context of IoT. SMCs “are cryptographic techniques that enable computations on data received from different parties in a way that each party knows only its own input and the result of the computations” [GOLD97]. Practically, the SMC functions operate on top of private data that two or more parties contribute to the particular computation, without allowing other participants or third party to access data. The authors in [GOLD97] defines SMC “... as a mechanism to compute a function \( f(x_1, x_2, \ldots, x_n) \) with private inputs \( x_1, x_2, \ldots, x_n \) in a distributed environment with \( n \) participants, where each participant \( i \) knows only its input \( x_i \) and no other information, except the value of the output \( f(x_1, x_2, \ldots, x_n) \).”

The authors in [DUAT01] discuss several efficient solutions to the general SMC problem, for ubiquitous environments with either energy or computational or communication resources constraints. These solutions are often problem-specific and may be resource demanding. In the following, we summarize examples of resource efficient SMC:
● **the secure sum protocol** [CKVL02], which tries to find the best match between the user privacy protection and the service efficiency;

● **the homomorphic encryption** [BENA94] [PAIL99] [NACC98], which allows for a specific algebraic operation to be executed without decrypting the data;

● **the location- and identity-based encryption** of the 2G/3G mobile systems, which is based on three-way authentication [KOOL05];

● *privacy preserving online dating systems*, and/or *privacy preserving service discovery systems*, can be seen as *access control systems*, offering user’s identification without revealing her identity;

● **privacy preserving attribute-based access control system**, which protects user identity by enforcing access control based on required attributes [FRAL06];

● **private collaborative forecasting and benchmarking**, which is based on the computation of the aggregated results, while private data remain secret [ABLF04].

The authors in [SIN11] noticed that “...most of the SMC protocols require considerable computational and communication overhead, and are inefficient in real world ubiquitous applications.... The future research effort should attempt to develop security and privacy models that may provide low-cost practical solutions with acceptable security level for a given type of application.”

### 2.1.2 Privacy mechanisms

The authors in [RONL11] discuss several privacy mechanisms, such as: privacy by design, data transparency, data management, trust and governance. In our analysis, we add to it mechanisms such as obscurity by design, identity management, data obfuscation, tamper-proofing, watermarking or fingerprinting.

**Privacy by design mechanism** provides users with the tools created to support the management of their own data. For example, whenever users produce a data fragment, they can use dynamic consent tools that permit certain services to access as little or as much of that data as desired [RONL11]. The implementation of the privacy by design faces a number of challenges, including a lack of specificity, and weak market forces motivating adoption [RUBI12]. To date, it includes mainly back-end implementation principles, such as data minimization and security [GÜTD11] [KUFK11] [SPCR09].
Obscurity by design mechanism incorporates flexible set of principles, which designers should consider when building privacy mechanisms. The authors in [STHA12] discuss obscurity by design in the context of Social Media, and define obscurity as follows: “Information is obscure online if it exists in a context missing one or more key factors that are essential to discovery or comprehension. We have identified four of these factors: 1) search visibility, 2) unprotected access, 3) identification, and 4) clarity.”

Data transparency mechanism informs users about entities involved in managing their data. This mechanism will affect both stakeholders and users, in sense of the amount of user’s personal (or corporates) data, which are provided to IoT services.

Data management mechanism includes data management policies and policy-enforcement mechanisms, which “requires interpreting, translating, and optimally reconciling a series of rules, each of which might be in a different language. And any policies must align with legislation on data protection, which itself could change.” [RONL11]

A well known protocol and tool suite for specifying privacy policies of both Web sites and users is Platform for Privacy Preferences (P3P) [CRAN03] [LIBH05]. The authors in [REBE03] notice that privacy solutions for the Web (notably P3P-based) would be better utilized if they were a part of data they are supposed to protect. Hence, in [LIBH05], the idea of coupling privacy protection mechanisms with metadata has been proposed based on a work presented in [BHLX04] [REBE03].

Identity management mechanisms differ between several object identity principles, such as core and temporary identities, identity of a group of objects, identity based on owner’s identity or some specific features (quantity, ingredients, etc.), authorization and authentication. One approach to identify the device or its owner is known as digital shadowing [SAGI09], which only implicitly identifies the owner. In other words, digital shadowing is based on the notion of a digital shadow: user’s sensors do not store the owner’s real identity, only a virtual identity that contains information about the user’s attributes and their interactions with the IoT environment.

Digital shadows could be further linked to RFID tags that are accessible only to authorized readers. This implies that an unauthorized RFID reader is not able to track the user. This concept is known as - revocable access delegation [RERG10]. Another concept in which an RFID reader scans the tags embedded in some object, downloads the companion privacy policy, and compare two of them is called - the privacy coach [BROE10]. If the company’s privacy policy does not match the user’s preferences, the user can choose not
to use it at all. Additionally, whenever an RFID reader tries to read the mobile phone’s signal, the phone can check the reader’s privacy policy and ask for user’s consent.

**Trust and governance frameworks** are considered as the enforcing mechanisms towards simplification of data protection: trust should reduce the uncertainty of objects interacting in a dynamic and ubiquitous IoT environment, while governance framework should reduce liability of billions of IoT objects in the environment. Recently, the authors in [HGZZ14] presented a data-driven qualitative model for the evaluation of the trust value of sensed products in supply chains. In this model, a big volume of sensed data is used through a Bayesian network to calculate a quantitative trust value, rather than expert’s subjective experiences and knowledge. The authors in [HGZZ14] additionally explore other trust models designed for IoT, such as:

- TRM-IOT model uses only reputation as the main index of trust [CHCH11];
- A multi-dimensional trust evaluation model for large-scale P2P computing, as proposed in [LIZY11], combines several factors to evaluate the trust, e.g., assumptions, expectations, behaviors, etc.;
- Hexmoor et al. in [HEWB06] proposed a trust-based security model, in which evaluation requires behavior data and expert’s experiences;
- Wang et al. in [WAWU10] proposed a topology transform-based recommendation trust model, which can relieve the malicious effects on the accuracy of recommendation trust. This model uses recommendation mechanisms that are rather not suitable in the IoT;
- Shi et al. in [SHBA05] proposed a trust model that assists users and machines in online decision-making using past behavior as a predictor of likely future behavior.

Other approaches for protecting software clients or agents from a malicious host include [COTH00]: (a) data obfuscation, (b) tamper-proofing, and (c) watermarking or fingerprinting. **Data obfuscation** is “a mechanism for hiding private data by using misleading, false, or ambiguous information, with the intention of confusing an adversary” [BRNI11]. This mechanism acts as a noisy information channel between a user’s private data and an untrusted observer [CHPA08] [SHOK15]. Some problems related to data obfuscation that still need to be addressed in the IoT are as follows:

- how to design an obfuscation mechanisms that protects the privacy of the user and, at the same time, imposes a minimum utility cost, and
- how to guarantee the user's privacy, despite the lack of a single best metric for privacy.

Regarding privacy protection, there are two major metrics proposed in [SHOK15]:

- **Differential privacy**, which limits the information leakage through observation, but does not reflect the absolute privacy level of the user; and

- **Distortion privacy** (inference error), which measures the error of inferring user's secret from the observation. This requires assumption of a prior knowledge which enables quantification of absolute privacy, but is not robust to adversaries with arbitrary knowledge.

Thus, either of these metrics alone is incapable of capturing privacy as a whole. They rather complement each other: Differential privacy is sensitive to the likelihood of observation given data. Distortion privacy is sensitive to the joint probability of observation and data.

The authors in [SHOK15] model the optimal obfuscation mechanism that minimizes utility loss, satisfies differential privacy, and guarantees distortion privacy, given a public knowledge on prior leakage about the secrets. The authors measure the involved metrics based on separate distance functions defined on the set of secrets. They model prior leakage as a probability distribution over secrets, which can be estimated from the user's prior data.

The problem of optimizing the tradeoff between privacy and utility has been discussed in the literature, but notably in the context of statistical databases [BRNI10] [GHRS09] [GHRS12] [LHRM10] [BRSH08] [IMWB14]. A protection mechanism for distortion privacy metric can be designed to be independent of the adversary's inference algorithm (e.g., Bayesian inference [MCKY03] [BERG85] as privacy attacks). Distortion privacy, which evaluates privacy as the inference error [STBH11], is a follow-up of information theoretic metrics for anonymity and information leakage [DSCP03] [SEDA03]. This class of metrics is concerned with what can be inferred about the true secret of the user by combining the observation (or obfuscated information) and prior knowledge. The problem of maximizing privacy under utility constraint, assuming prior knowledge, is proven to be equivalent to the user's best strategy in a zero-sum game against adaptive adversaries [STTH12]. With this approach, one can find the optimal strategies using linear programming. However, if we want to guarantee a certain level of privacy for the user and maximize her utility, the problem cannot be modeled as a zero-sum game anymore and there has been no solution
for it so far. Hence, the authors in [SHOK15] adapt a game-theoretic notion of privacy for designing optimal obfuscation mechanisms against adaptive inference. They formulate this game as a Stackelberg game and solve it using linear programming. They add the differential privacy guarantee as a constraint in the linear program and solve it to construct the optimal mechanism. The result is that not only the perturbed data samples are indistinguishable from the true secret (due to differential privacy bound), but they cannot be used to accurately infer the secret using the prior leakage (due to distortion privacy measure). Finally, the authors in [MZAB12] explore examples combining security and game theory.

2.1.3 Data Access Control in the IoT
In the IoT, both the device manufacturer and the service provider may have access to more data that could be analyzed [WP29-14]. Third parties may want to make use of this data for different purposes in order to infer various information about users; for example, user’s habits, daily routines, health and physical conditions, political opinions, location analytics, etc. The Article 6 of Directive 95/4/EC constitute a cornerstone of EU data protection law, emphasizing the following principles:

- personal data should be collected and processed fairly (never without the individual being aware of it) and lawfully;
- data can only be collected for specified, explicit and legitimate purposes (the purpose limitation principle);
- data unnecessary for a specific purpose should not be collected and stored “just in case” or because “it might be useful later” (the data minimization principle);
- personal data collected and processed in the context of IoT is kept for no longer than is necessary.

In addition, the owners of data (data subjects) must have a possibility to revoke any prior consent given to a specific data processing and to object to the processing of data relating to them. The exercise of such rights must be possible without any technical or organizational constraints or hindrances and the tools provided to register this withdrawal should be accessible, visible and efficient. Furthermore, some wearable IoT devices (glasses, clothes, etc.) must include a possibility to disable the “connected” feature of the thing and allow it to work as the unconnected item. Today practices in the design of IoT devices still DO NOT PROVIDE ANY CONTROL of the data.
2.2. Security Challenges in the IoT

The development of the IoT entails significant security concerns. For example, the authors in [WP29-14] identify several challenges that force device manufacturers to balance the implementation of confidentiality, integrity and device security, with the need to optimise battery efficiency (the use of computational resources and energy). The current solutions are usually compromise solutions, offering low security, the ease of surveillance practices and data breaches, for the price of better battery efficiency. This comes with the risk that the IoT “may turn an everyday object into a potential privacy and information security target, while distributing those targets far more widely than the current version of the Internet.”

The security in the IoT is not only about securing devices; it might affect other IoT components, such as the communication links, storage infrastructure, data access controls, data itself. All IoT components must be designed and manufactured in a way that guarantees high security standards. The compliance with the Article 17 of Data Protection Directive 95/46/EC should be additionally presented.

Before trying to use the software to secure any system, it is necessary to proof the hardware itself, and to create a security protocol for selecting securable devices. It should include any device connected in the network, such as USB sticks or chargers, and power outlets, etc. IoT sensors that allow direct wireless connection might also be hopeless to secure the network; preferably routers or other gateways will enhance security of the sensor array.

The future communication between agents and various devices may use homographic encryption to encrypt all data flowing in the system. The encrypted data can be passed with an infrastructure and operations can be performed on the data. Complete homographic encryption libraries are not yet offered; there are only simple research based libraries.

In the following, we summarize several recommendations on securing the IoT:

- Identification of the security weaknesses of any hardware and IoT devices, including USB connectors, sticks, power extenders, etc.;
- IoT devices should support an API that can be secured either by the vendor mechanism or by an IoT stack local security gateway. Vendors offering “security by obscurity” should be encouraged to offer alternative solutions;
- Users or integrators may not use USB related devices unless USB protection is utilized to prevent attacks on the whole network;
The IoT stack must offer a local security gateway and registration mechanism for any registered IoT device;

- The IoT stack should simplify how components or agents interact. For example, peers (agents or system components) could use a secure proxy (a master server peer in P2P network) for communicating with other peers or agents.

### 2.2.1 Top security vulnerabilities in the IoT

The Open Web Application Security Project (OWASP)\(^2\) identifies 10 top vulnerabilities in the IoT, discusses their technical and business impacts, and the ways to prevent application attacks and reduce network vulnerabilities. Here is a brief overview of OWASP’s top IoT vulnerabilities:

- **Insecure Web interface** (e.g., ensuring that weak passwords are not allowed; account lockout after 3-5 failed login attempts, etc.);
- **Insufficient Authentication/Authorization** (e.g., ensuring that granular access control is in place when necessary; credentials are properly protected; password recovery mechanisms are secure, and implement two factor authentication where possible, etc.);
- **Insecure Network Services** (e.g., ensuring that only necessary ports are exposed and available; services are not vulnerable to buffer overflow and fuzzing attacks);
- **Lack of Transport Encryption** (e.g., ensuring that data is encrypted using protocols such as SSL and TLS while transiting network, etc.);
- **Privacy Concerns** (e.g., ensuring that only data critical to the functionality of the device is collected; not collecting sensitive data; data anonymization, etc.);
- **Insecure Cloud Interface** (e.g., ensuring that credentials are not exposed over the internet; two factor authentication implemented);
- **Insecure Mobile Interface** (e.g., ensuring that user accounts can not be enumerated using functionality such as password reset mechanisms; implementing two factor authentication if possible);
- **Insufficient Security Configurability** (e.g., ensuring the ability to separate normal users from administrative users, the ability to force strong password policies, etc.);
- **Insecure Software/Firmware** (e.g., ensuring that the device has the ability to

---

\(^2\) OWASP project:
https://www.owasp.org/index.php/OWASP_Internet_of_Things_Top_Ten_Project#tab=Main
update (very important); the update file is encrypted using accepted encryption methods, etc.);

- **Poor Physical Security** (e.g., ensuring that data storage medium can not be easily removed; device can not be easily disassembled, etc.).

### 2.2.2 Detecting and preventing attacks in the IoT

Attacks in the IoT may lead to disconnecting significant parts of the network, replication of strategic nodes in the network, and more. Existing approaches to sensor network security are mainly based on detection of distributed, replicated nodes, either by using the topology of the network to detect replication (**line-selected multicast**), or by exploiting the birthday paradox to detect replicated nodes (**randomized multicast**) [PAPG05].

The authors in [KAWA02] note that security issues present in ad-hoc networks are similar to those in sensor networks (as previously discussed in [ZHHA99] [STAN99]), but the defense mechanisms developed for ad-hoc networks are not directly applicable to sensor networks. The reasons for it could be the following:

- **Public key cryptography** used in some ad-hoc network security mechanisms for authentication and secure routing protocols [ZHHA99] [HUBC01] [KZLZ01] [ZAPA01] [BITR01] is too expensive to be implemented in sensor nodes;

- **Symmetric key cryptography** for ad-hoc networks in secure routing protocols have been proposed in [HUJP02] [HUPJ01] [BHRB01] [PAHA02]. These protocols are based on source routing or distance vector protocols, but they are unsuitable for sensor networks;

- Marti et al. in [MGLB00] and Buchegger and Boudec in [BUBO02] consider the problem of minimizing the effect of misbehaving or selfish nodes on routing through punishment, reporting, and holding grudges. They introduce two extensions to the Dynamic Source Routing Protocol (DSR), called **Watchdog** (detects misbehaving nodes by overhearing transmission, although it might not detect the presence of ambiguous collisions, receiver collisions, limited transmission power) and **Pathrater** (avoids routing packets through misbehavior nodes);

- Perrig et al. in [PSWC01] present two building block security protocols optimized for use in sensor networks, such as SNEP and (micro)TESLA. SNEP provides confidentiality, authentication, and freshness between nodes and the sink, and (micro)TESLA provides authenticated broadcast.
The authors in [KAWA02] analyze the security features of the major sensor network routing protocols, crippling attacks against those protocols, and suggest possible countermeasures. One of the major assertion here is that “sensor network routing protocols must be designed with security in mind, which is the only effective solution for secure routing in sensor networks”. Most network layer attacks against sensor networks falls into one of the following categories:

- **Spoofed, altered or replayed routing information**: Spoofing allows for many other forms of network attacks, such as creating routing loops, generating false error messages, partitioning the network, etc. The authors in [SRUM13] propose a method for detecting spoofing attacks that utilizes K-means cluster analysis (as localization algorithms), and integrate this method into a real time indoor localization system, which localize the position of the attacker.

- **Selective forwarding**: A simple form of this attack is a case when a malicious node behaves like a black hole, and refuses to forward every packet she sees. A more subtle form of selective forwarding attack is a case when an adversary selectively forwards packets.

- **Sinkhole attacks**: the adversary’s goal is to lure the traffic from a particular area through a compromised node, creating a metaphorical sinkhole with the adversary at the center.

- **Sybil attacks**: a single node that presents multiple identities to other nodes.

- **Wormholes**: an adversary is tunneling messages received in one part of the network over a low latency link and replays in a different part [HUPJ02] [KAWA02].

- **HELLO flood attacks**: this attack uses a single hop broadcast with the aim to transmit a message to a large number of receivers [KAWA02].

- **Acknowledgement spoofing**: an adversary is spoofing link layer acknowledgments for “overheard” packets addressed to neighboring nodes. Acknowledgment spoofing is about convincing the sender that a weak link is actually strong, or that a dead or disabled node is alive, then transmitting packets on exactly those links.

Protecting sensor network routing protocols against outsiders, Sybil attacks, HELLO floods, spoofing, can be achieved by augmenting existing protocols with the mechanisms such as: link-layer encryption and authentication using a globally shared key, multipath routing,
identity verification, bi-directional link verification, and authenticated broadcast [KAWA02]. Cryptography itself is not enough to defend against laptop-class adversaries. Hence, the design of the secure routing protocol need to be considered from the early design and the development of sensor network, in order to provide an effective detection and prevention of attacks in the IoT. Table below summarizes on various types of attacks and their countermeasures. For example, HELLO flood attack responds to bi-directional link verification, and/or authenticated broadcast, but cannot be solved through link-layer encryption.

<table>
<thead>
<tr>
<th>Attack Type</th>
<th>Link-layer encryption and authentication using a globally shared key</th>
<th>Multipath routing</th>
<th>Identity verification</th>
<th>Bidirectional link verification</th>
<th>Authenticated broadcast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sybil attack</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Selective forwarding</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sinkhole attack</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Wormhole attack</td>
<td>-</td>
<td>+</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>HELLO flood attack</td>
<td>-</td>
<td></td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Spoofing attack</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

### 2.2.3 Securing Data Provenance in the IoT

The provenance, which is also known as Information lineage, or archival principles, has been extensively studied in archival theory. The authors in [HASW07] define data provenance as a summary of "the history of the ownership of the item, as well as the actions performed on it". Furthermore, the authors in [HASW07] explore the secure provenance problems and challenges in the area of computing and legal proceedings that
involves data. They define the security components model, which consists of confidentiality, integrity and availability (CIA) as the main facets of data security.

The authors in [SIPG05] present taxonomy of the application of provenance in science, and classifies provenance systems into the following three categories: database-oriented, service-oriented, and miscellaneous. For example, [WIDO05] presents a system called Trio, which purpose is to provide lineage and provenance in databases. The examples of provenance management systems for scientific computing include: Chimera [FVWZ02] (physics and astronomy), myGrid [ZGSB04] (biology), CMCS [MYER04] (chemical sciences), and ESSW [FRBO01] (earth sciences).

Buneman et al. in [BUKT00] explore various aspects of provenance of electronic records, and introduce the notions of why-provenance (process that generated the data) and where-provenance (the origin of data). Workflow provenance in scientific computing has been studied in [BADI06]. Provenance of data streams has been presented in [VIPL06], while the use of provenance in generating trust in social networks is explored in [GOLB06].

The authors in [LKMS04] emphasized the need for trust and provenance in information retrieval. Provenance as a mechanism for establishing trust in data, as defined by the W3C Provenance Working Group, represents "... information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability or trustworthiness." 3

The existing implementations of the provenance mechanisms include mainly systems for collecting provenance information, such as Provenance Aware Storage System (PASS) [MURE06], and Lineage File System [SACA05]. In parallel, secure file systems have been developed to store and secure data files in untrusted servers [HASA05]. Existing techniques for detecting breach of integrity, as described in secure file systems, can be utilized to provide trustworthy provenance.

Corruption of a provenance may lead to incorrect and possible harmful conclusions, leakage of privacy-related information (e.g., revealing information such as social security numbers) (see http://www.w3.org/TR/2013/NOTE-prov-aq-20130430/), etc. Publishing information about provenance must include access controls and mechanisms for enforcement and auditing of privacy policies (for example, standard HTTP authorization mechanisms can restrict access to resources, returning 401 Unauthorized, 403 Forbidden, or 404 Not Found as appropriate, or SSL, rate-limiting, spam filters, moderation queues,

---

3 W3C Provenance Working Group: http://www.w3.org/TR/2013/NOTE-prov-overview-20130430/
Reasonable precautions should be taken to prevent malicious use, such as Distributed Denial of Service (DDoS) attacks, Cross-Site Request Forgery (CSRF), spamming and hosting of inappropriate materials, or provenance pingback with a malicious site that presents a pingback URI executing an instruction on a different web site.

Provenance in the IoT brings many novel applications, as summarized in [COCT14]:

- supporting the understanding and reuse of sensor data, including data that has been aggregated or processed [PAHE08];
- ensuring the correct attribution of publicly available sensor data [CHSV10];
- supporting users trace the involvement of sensor data in experiments for reproducibility purposes [HESN13];
- searching for, and identifying sensor data within data stores [LNHM05];
- verifying data transmitted through nodes in a sensor network [SSGB12];
- supporting quality [CEBM13] and trustworthiness assessments [UMBR12], and many more.

2.2.3.1 Data Provenance and Sensor Ontologies

There exists a significant research track about the data provenance and sensor ontologies, for example:

- **The Open Provenance Model (OPM)** [MCFF11] is used by Lie et al. [LFMR10] to record the provenance of sensors within Tupelo, a semantic Content Management System (CMS). In Tupelo, data is modeled as the OPM equivalent of prov:Entity and actions as prov:Activity; no further alignments are described.

- Patni et al. [PSHS10] use the Provenir ontology [SBGS09] to capture and store the provenance of sensor data in their sensor management system. Alignments between Provenir ontology and the Sensor ontology are defined through series of rdfs:subClassOf and rdfs:subPropertyOf relationships that model e.g., observation values as the Provenir equivalent of prov:Entity, and the sampling time property as the equivalent of prov:atTime.

- Stasch et al. in [SSAK11] describe their extension of the Semantic Sensor Network Ontology (SSNO) to represent aggregations of observations, and use the
Provenance Vocabulary and OPM to record each aggregated provenance. However, it is unclear if the remaining SSNO concepts have been aligned to a provenance model, or if these alignments have been explicitly defined in an ontology to enable their reuse.

- Corsar et al. in [CEBM13] define an alignment between the SSNO and PROV-O (W3C Provenance Ontology), which covers only relationships between the SSNO observation, sensor, and sensing concepts and the PROV-O entity, agent, and activity concepts, respectively. The alignment is used to integrate SSNO data with other data via the PROV-O model to support quality assessment of observations from citizen sensors.

- Finally, the W3C Provenance Working Group defines the following:
  - PROV-DM, data model for provenance, which introduces a set of concepts to represent provenance information in a variety of application domains, in order “to enable the interoperable interchange of provenance information in heterogeneous environments such as the Web”,
  - PROV-O ontology, which maps the PROV-DM to RDF (see: http://www.w3.org/ns/prov),
  - PROV-AQ mechanisms for accessing and querying provenance,
  - PROV-CONSTRAINTS, a set of constraints applied to the PROV data model, and more.

In addition, the W3C Provenance Working Group developed an alignment between the SSNO and PROV-O ontologies, as two major reference models.

2.2.3.2 PROV-DM, PROV-AQ and PROV-O
In the following, we summarize the main idea behind the W3C Provenance Working Group’s definitions of (i) PROV-DM, (ii) PROV-O ontology, (iii) PROV-AQ mechanisms.

**PROV-DM.** Types and relations in PROV-DM are shown in Figure 1.

- **C1: Entities and Activities.** It consists of entities, activities, and concepts such as generation, usage, start, end.
- **C2: Derivations.** It contains derivations and derivation subtypes.
- **C3: Agents, Responsibility, and Influence.** It consists of agents and concepts defining responsibility of agents.
- **C4: Bundles.** It is concerned with bundles, which is a mechanism to support
provenance.

- **C5: Alternate.** It consists of relations between entities, referring to the same thing.
- **C6: Collections.** This component is about collections.

![PROV-DM components](http://www.w3.org/TR/2013/REC-prov-dm-20130430/#component2)

**Figure 1:** PROV-DM components (source: http://www.w3.org/TR/2013/REC-prov-dm-20130430/#component2)

**PROV-AQ** (c.f. http://www.w3.org/TR/2012/WD-prov-aq-20120619/) defines mechanisms for accessing, locating and querying provenance information. For example, there are three mechanisms for the resource provider to indicate a provenance-URI or service-URI along with a target-URI, such as:

- The requester knows the resource URI and the resource is accessible via **HTTP**
- The requester has a copy of a resource represented as **HTML** or **XHTML**
- The requester has a copy of a resource represented as RDF (or **HTML with embedded RDFa**).

**PROV-O.** The PROV-O ontology encodes the PROV-DM to RDF, using the OWL2 Web Ontology Language. PROV-O is a lightweight ontology that can be adopted in a wide range of applications. It can be used for modeling application-specific provenance details, or as a reference model for creating other domain-specific provenance ontologies.

The PROV-O terms (classes and properties) are grouped into three categories:

- **Starting Point classes and properties** provide the basis for the PROV-O (e.g. classes: `prov:Entity`, `prov:Agent`, `prov:Activity`; properties: `prov:wasGeneratedBy`, `prov:wasDerivedFrom`, `prov:used`, ...). These terms are used to create simple provenance descriptions that can be elaborated using terms from other categories.
- **Expanded classes and properties** provide additional terms that can be used to relate classes in the *Starting Point* category (e.g. classes: `prov:Collection`, `prov:Bundle`, `prov:Person`; properties: `prov:wasQuotedFrom`, `prov:wasInvalidatedBy`, `prov:value`, ...).

- **Qualified classes and properties** provide elaborated information about binary relations asserted using *Starting Point* and *Expanded* properties (e.g. classes: `prov:Plan`, `prov:Role`, `prov:Communication`; properties: `prov:wasInfluencedBy`, `prov:hadUsage`, `prov:hadRole`, ...). The terms in this category are applied using a pattern that differs from those in the *Starting Point* and *Expanded* categories. While the relations from the previous two categories are applied as direct, binary assertions, the terms in this category are used to provide additional attributes of the binary relations.

### 2.2.3.3 PROV-O and SSNO Ontology Alignment

The authors in [COCT14] present an ontology alignment between PROV-O (http://www.w3.org/ns/prov) and SSNO (Semantic Sensor Networks Ontology) (http://purl.oclc.org/NET/ssnx/ssn), which are both W3C’s ontologies. While the PROV-O ontology is a W3C recommendation for the representation and interchange of provenance information on the Web [LSGB14], the SSNO ontology [CBBG12][LHTB12] represents a de-facto standard for the semantic specification of sensors, sensor devices, systems, processes, and observations. The alignment between PROV-O and SSNO considers PROV-O as an upper ontology, which further extends SSNO concepts and properties. This allows for representation of observation details and sensor deployments that are not possible in the SSNO alone, and gives a basis for alignment with Open Geospatial Consortium Observations & Measurements ontologies.

SSNO is built around the *Stimulus-Sensor-Observation pattern* [JACO10] that describes the relationship between an observing `ssn:Sensor`, the `ssn:Property` that is measured, the real-world `ssn:Stimulus`, the `ssn:Sensing` procedure followed and the resultant `ssn:Observation`. The Stimulus-Sensor-Observation pattern is a basis for the alignment to PROV-O.
2.2.4 Game Theory to support IoT security

Symantec’s security study from 2012 indicates that cyber attacks are continuing to raise! Only in 2012, cyber attacks increased by 42 percent (c.f. http://bit.ly/1Sz8w2O). In 2013, the Spamhaus/Cyberbunker attack clogged servers with dummy internet traffic at a rate of about 300 gigabits per second. This happened 13 years after the publication of best practices on preventing Distributed Denial of Service (DDoS) attacks, and motivated cyber-security professionals to search for novel approaches to IoT security.

In IoT4Industry, we specifically explore a game-theoretic setting for interaction among benign and malicious users. Therefore, we start by examining the recent work on using Game Theory (GT) for both the IoT and Wireless Sensor Networks. Game Theory is a mathematical method with the potential to describe both cooperation and conflicts between rational decision-makers [SWKC12]. It can be used to analyze system operations in decentralized and self-organizing networks, and describe the behavior of players involved in a game. For example, GT addresses problems where multiple players with different objectives compete and interact with each other. From the perspective of Wireless Sensor Networks (WSNs) and the IoT, GT has been proved to be useful in improving WSN performance objectives caused by the limited capabilities of sensor nodes in terms of computation, communication, and energy (e.g., GT can maximize sensing coverage, extend operation periods, or improve security aspects via adequate mathematical generalization). Furthermore, GT mechanisms can be used to monitor behavior of sensor nodes and evaluate reputation of each node based on collected observations. Reputation can be further used to predict the future behavior of other nodes in WSN.

There is already a plethora of research papers describing GT approach to WSNs: [KASI04] [KAIY04] [URBG03] [FEHB06] [BUHS05] [WALI06]. In addition, Machado and Tekinay summarize 29 publications with the focus on the use of GT approaches to formulate problems related to security and energy efficiency [MATE08]. Shen et al. in [SYCY11] summarize another 30 publications of the existing GT approaches to strengthen WSN security.

Commonly used GT methods to solve problems in WSNs can be classified as follows:

- **cooperative GT** - a class of games that studies the behavior of cooperative players and considers the utility of all the players involved in the game. This category of games is also known as **coalitional GT**, and in WSN particularly, allows for a
reduction of power consumption by forming coalitions. For example, some nodes in WSN can form a coalition by transferring data coordinately with the aim to minimize energy consumption;

- **non-cooperative GT** - a class of games that studies behavior of non-cooperative, competing players, and focuses on each user’s individual utility, rather than on the utility of the whole network. Mechanism design in non-cooperative GT helps to find the right incentives and ensures that the participants tell the truth (agent obtains maximum profit only by submitting a real bid/true value);

- **repeated GT** - a class of dynamic games, in which a game is played for numerous times and the players can observe the outcome of the previous game before attending the next repetition [OWEN01]. The interaction between sensor nodes can be modeled by using repeated GT, to prevent the problem of dropping packets attacks in WSN [YAMC08], or prevent passive DoS attacks at routing layer in WSN [AGDA07], etc.

- **evolutionary GT**, and more.

Hackers could attack WSN, and sensitive corporate or private data could be easily compromised. Hence, WSN need to be designed to prevent DoS attacks and detect intrusions, provide secure localization, etc. The authors in [SWKC12] survey the recent developments in the area of GT and GT-based applications in WSNs, such as: routing protocol design, topology control, power control and energy saving, packet forwarding, data collection, spectrum allocation, bandwidth allocation, quality of service control, coverage optimization, and other sensor management tasks (task scheduling, etc.). Shen et al. in [SYCY11] classify current GT approaches in WSN security into four categories: (i) preventing DoS attacks, (ii) detecting intrusion, (iii) strengthening security, and (iv) coexistence with malicious sensor nodes that is based on signaling games described in [WACK09]. Shen et al. pointed out on several future research areas for ensuring WSN security based on GT, including Base Station credibility, Intrusion Detection System (IDS) efficiency, WSN mobility, WSN QoS, real-world applicability, energy consumption, sensor nodes learning.

In the rest of this section, we present the above mentioned categories of GT methods for WSN security.
2.2.4.1 Preventing DoS Attacks
Denial of Services (DoS) attacks can be passive and active [AGDA07]. In passive attacks, selfish sensor nodes use the network, but do not cooperate with other sensor nodes in order to save battery life for their own communications. In that way, they do not intend to damage other nodes, while in active attacks, malicious nodes damage other nodes by causing network outage by partitioning. The existing non GT mechanisms to prevent DoS attacks in sensor networks are suggested in [AGDA07]:

- **Watchdog scheme**: it identifies misbehaving nodes [MGLB00] using two concepts:
  - *Watchdog* (every node implements constant monitoring). The disadvantage of the Watchdog scheme is that it doesn’t solve the collision between malicious nodes [HUSH03]), and
  - *Pathrater* (rates the transmission reliability of all alternative routes to a particular destination node).
- **Rating scheme**: the neighbors of any single node collaborate in rating the node, based on how well the node execute requested functions [MIMO02a] [MIMO02b] [MIMO02c]. The disadvantages of this approach are that evaluated node may be able to cheat easily and the result of the function may require significant overhead to be communicated to the evaluating node [HUSH03].
- **Virtual currency**: This scheme introduces a special type of selfish nodes called *nuglets* [BBCG01] [BUHU01]. The disadvantages of this scheme are that intermediate nodes are able to take out more nuglets than they are supposed to, and overhead [HUSH03].
- **Route DoS prevention**: It attempts to prevent DoS in the routing layer via cooperation of multiple nodes [BUBO02].

Apart non-GT mechanisms to prevent DoS attacks in sensor networks, there are GT-based models such as:

- **non-cooperative game** [AGBD05a] [AGBD05b] [SAEP09],
- **cooperative game** [AGBD04], and
- **repeated game** [AGDA07] [YAMC08].

The authors in [AGDA07] present a protocol based on repeated GT, which achieves the design objectives of truthfulness by recognizing the presence of nodes that agree to forward packets but fail to do it (passive DoS). The jamming and anti-jamming attacks are
modeled as a zero-sum stochastic game in [LILQ11], designed to defend DoS attack. Dong et al. [DLLX09] established an attacking-defending game model for detecting active DoS attacks, in which the strategy space and payoff matrix are given to both the Intrusion Detection System (IDS) and the malicious nodes.

2.2.4.2 Intrusion Detection

Research in the area of IDS for WSN has a relatively short history [PAPA06]:

- Marti et al. in [MGLB00] introduce the concept of *Watchdog* and *Pathrater* to identify routing misbehavior;
- Buchegger & Boudec in [BUBO2] extend the work of Marti et al. by replacing the Watchdog with a *Neighborhood Watch* paradigm. They also introduce a Trust Manager, a Reputation System and a Path Manager paradigms;
- Zhang & Lee in [ZHLE00] propose a general intrusion detection and response mechanism for the Mobile Ad hoc NETwork (MANET), in which each IDS agent participates in the intrusion detection and response independently;
- Huang et al. in [HFLY03] extend the work done by Zhang & Lee, by analyzing the routing activities and improving the anomaly detection process by providing more details about the attack types and attack sources;
- Sun et al. in [SUWP03] proposed a Markov chain-based anomaly detection approach for MANETs.

The use of mobile agents in the context of intrusion detection has been explored in both the LIDS project [PPCJ03] and the SPARTA project [KRT002]. Kachirski & Guha in [KAGU02] propose distributed IDS for the MANET environment [MINP04]. Several GT frameworks for modeling IDS has been presented in [KOLA03] (based on sampling in communications networks), [ALBA04], [PAPA06] (the interaction between an attacker and a host-based IDS is analyzed as a **dynamic two player non-cooperative game**). Reddy in [REDD09] discusses currently available IDS techniques, models attacks using GT and proposes a framework to detect malicious nodes by using **zero sum approach** for nodes in the forward data path. Reddy & Srivathsan in [RESR09] developed a framework to detect malicious nodes using zero sum game approach and selective node acknowledgements in the forward data path. A GT-based scheme is introduced to find out the vulnerable areas in WSN [AGBD04] [AGDB04]. Mohi et al. [MOMZ09] models the interaction of nodes in WSN
and IDS as a **Bayesian game** and use the same concept to make a secure routing protocol. Finally, Michiardi et al. [MIMO02] use **cooperative and non-cooperative GT** to develop a reputation based architecture to enforce cooperation. **Non-cooperative GT-based** schemes could assist detection schemes in improving their performance as well as an efficiency [XHTP11]. Qiu et al. in [QICX10] propose an active defense model for WSN in which the nodes have limited rationality of evolutionary learning ability based on **evolutionary GT**. The nodes can achieve the most effective defense by adjusting their defensive strategies active and dynamic. Jing & Ruiying in [CHDU09] analyze the cooperation stimulation and security in self-organized WSN under a GT framework.

### 2.2.4.3 Strengthening Security

GT-based approach to strengthening security considers auction theory [AGBD06] and coalitional game [LILY08], as presented in [SYCA11].

**Auction Theory for Strengthening Security.** There are four major auction formats: English, Dutch, First-Price and Second-Price sealed auctions. The authors in [AGBD06] propose a **Secure Auction-based Routing (SAR) protocol** using the First-Price auction format. Here, the bidder with the highest bid wins the auction and reaches equilibrium, and thereafter the truth bidding becomes a dominant strategy for sensor nodes. Both malicious and truthful sensor nodes compete against each other in order to forward incoming packets and, by doing so, each sensor node improves its reputation among other sensor nodes.

> “**The payoff of each sensor node is calculated based on battery power and reputation. The authors consider communication and computation of sensor nodes when computing the required power for each sensor node. If a sensor node saves its power by not forwarding incoming packets or dropping them for its selfishness, then it will be isolated from the network and get a bad reputation; otherwise, it will get a good reputation. As a result, sensor nodes prefer to participate in forwarding incoming packets and gaining reputation. Experiment results show that the SAR protocol can guarantee more reliable delivery, and can observe the behavior of sensor nodes and isolate suspicious sensor nodes by defining an acceptable threshold for reputation of sensor nodes.**” [AGBD06]

**Coalition Game for Strengthening Security.** The authors in [LILY08] study how to strengthen security in WSNs by using coalitional game. They propose a throughput characteristic function that describes the expected gain of a coalition from the cooperation. For finding a reliable routing path, a coalition to a weighted and directed graph is

---

30
formulated. The number of routing paths is related to the size of the coalition. In order to fairly distribute the gains among all sensor nodes, the authors in [LILY08] consider Shapley value method and prove that it is applicable to the payoff allocation inside coalitions given their proposed throughput characteristic function. Some game rules are defined as follows:

- A sensor node will join into a coalition only if it can create more payoff than being alone;
- A sensor node will deviate from the current coalition and join another coalition only if it can create more payoff in the future than current;
- A coalition will refuse a sensor node if the sensor node cannot increase the total payoff of the coalition;
- A coalition will exclude a sensor node if the sensor node can not benefit the coalition;
- Sensor nodes failing to join into any coalitions will be denied from the WSNs.

Under the above rules, the selfish sensor nodes that do not forward others’ data packets will hardly be admitted into coalitions because of their poor reputation. In such rules, sensor nodes are enforced to participate in a coalition and those that cannot join into any coalitions are under high suspicion of being malicious.

The coalitional game model can be integrated with various routing protocols; for example, AODV routing protocol presented in [PBRD03]. However, as noted in [SYCY11], there is a lack of simulation experiments for validating the performance of coalitional game integrated with the routing protocol, such as AODV and DSR [JOHM07].

2.2.4.4 Coexistence with Malicious Sensor Nodes
Malicious and regular sensor nodes can coexist as long as the destruction they bring is less than the contribution they make [SYCY11]. For example, the authors in [WACK09] study interactions between a malicious and a regular sensor node in WSN, defining two separated sets of sender’s and receiver’s strategies:

- The set of strategies of the sender is \{Attack, Forward\}. Attack means the sender is malicious, while Forward means the sender is regular or a malicious sensor node disguises itself;
- The set of strategies of the receiver is \{Monitor, Idle\}, which means the receiver monitors the sender or not, respectively.

The authors in [WACK09] additionally formalize the interaction between sensor nodes into
two games. The first game, called *malicious sensor node detection game*, is a signaling game that is a special category of Bayesian game with imperfect information. The second game, called *post-detection game*, is played when the regular sensor node knows confidently that its opponent is a malicious sensor node. Based on various combinations of the strategies used, the net utility is calculated. As noted in [SYCY11], the authors do not mention where their simulations are done and how the simulation environment has been configured. This is what motivates our research in *IoT4Industry*.

3. Industrial Requirements w.r.t. IoT Security, Privacy and IPR/ Data Protection

The Industrial Internet (of Things) brings new opportunities, along with new risks, to business and society in general. Many questions about the Industrial Internet still remain open; for example, how it will impact business models, policies, interaction, protection. To answer these questions and their impacts facing business and government leaders, the World Economic Forum’s IT Governors launched the *Industrial Internet initiative* at the Annual Meeting 2014 in Davos, Switzerland. This initiative conducted a series of research activities, including in-person workshops, virtual working group sessions, interviews of key thought leaders, and a survey of innovators and early adopters. The summary on this initiative is presented in the form of a study published in [WEF15], in which security and data privacy (data breaches, attacks, espionage) has been recognized as one of the most important hurdles to solve in order to realize the full potential of the Industrial Internet. The authors in [WEF15] conclude that cybersecurity for the IoT requires “new security frameworks that span the entire cyber physical stack, from device-level authentication and application security, to system-wide assurance, resiliency and incidence response models.” Figure 2 summarizes the greatest barriers that largely inhibit industries from adopting the Industrial Internet in EU, North America, and the rest of the world. Security and privacy concerns are ranked amongst major inhibitors. Therefore, the *Industrial Internet Reference Architecture (IIRA)*, as presented in its first published version in [IIRA15] “…initiates a process to create broad industry consensus to drive product interoperability and simplify development of Industrial Internet systems that are better built and integrated with shorter time to market, and at the end better fulfill their intended uses.” With a special focus on major architecture question in Industrial Internet, IIRA gives privacy and security a
strong emphasis. For example, it integrates the business viewpoint (the RoI for security), the usage viewpoint (transparent security to the user), the functional viewpoint (security for the system as a whole) and the implementation viewpoint (common architecture patterns and system components).

Figure 2. The greatest barriers that largely inhibit business from adopting the Industrial Internet (source [WEF15])

Furthermore, common security activities are described below [IIRA15]:

- **Security monitoring** gathers and analyzes security-related data continuously as activities are performed. It may take different forms depending on the context of operations and on security events. For example, different tasks are appropriate before, during and after an attack.
- **Security auditing** collects, stores and analyzes security information.
- **Security policy management** manages both automated and human-driven administrative security tasks by documenting their usage and constraints.
- **Cryptographic** support management consists of globally interoperable key management, secure credential storage and revocation.

To secure an Industrial Internet System, the authors in [IIRA15] outline a number of important and common security issues to be addressed in its implementation:
- **End-to-end security**: the implementation of an IIS must provide: protected device-to-device communications, confidentiality and privacy of the data collected, remote security management and monitoring, simultaneously addressing both existing technologies as well as new technologies, and seamlessly spanning both information technology (IT) and operational technology (OT) subsystems and processes without interfering with operational business processes.

- **Securing legacy systems**: the use of security gateways is an approach to secure legacy endpoints and their protocols. A security gateway acting as proxies for the legacy endpoints bridges the legacy protocols supported by the legacy endpoints and their counterparts used by new endpoints. A security gateway isolates the attack surface introduced by the legacy endpoints and their protocols to the links from these endpoints.

- **End-point security**: it provides embedding security capabilities and policy enforcement directly in end-point devices. The end-point should have the ability and autonomy to defend itself and must remain resilient even when disconnected from the external security management systems.

- **Information exchange security**: Communication and data exchanges within an IIS must be protected for authenticity, confidentiality, integrity and non-repudiation.

- **Management and monitoring security** of both the endpoints and the communication mechanisms, involves the following areas: identity management, provisioning and commissioning, security policy management, situational awareness, remote update, credential management, etc.

- **Data distribution and secure storage**, including data security, data centric policies, data analysis and privacy, etc.

Our analysis of industrial requirements for IoT security, privacy and IPR/data protection in *IoT4Industry*, is additionally based on the RAMI (Reference Architecture Model Industry 4.0), as well as the SCISSOR project (Horizon 2020 N644425, February 2015 - January 2018) and its project report "Technological assessments, scenarios and requirements". Salzburg Research is the project partner in SCISSOR. In Task 2.1, the project examines the real-world expertise of the consortium partners related to the ICS (Industrial Control System) technologies, devices, and network infrastructure. It considers the requirements related to security services, including a wide range of possible deployment and solutions,
from industrial control processes, pipelines, power generation, nuclear plants, water distribution, ambient control, etc.

Industrial infrastructures are often built on top of complex and heterogeneous systems (in sense of heterogeneous protocols in use), with many features strictly connected to the specific activity sectors, and strategically deployed, to manage their complexity. For example, literature on Smart Grid refers on Advanced Distribution Automation (ADA), which goal is to provide the real-time adjustment of the distribution system to changing loads, generation, and failure conditions, usually without operator intervention. In order to achieve real-time performances, many control applications will play a central role covering the full automation of all the controllable equipment and functions in the distribution system [EPRI04] [SCISSOR15a].

In order to achieve the goals of ADA, new applications and technology are expected to be developed, such as Fault Detection Isolation and Restoration (FDIR), Topology Processor (TP), Distribution Power Flow (DPF), Integrated Voltage/Var Control (IVVC), Optimal Feeder Configuration (OFC), Distribution Contingency Analysis (DCA), Distribution State Estimation (DSE), Distribution Load Forecasting and Estimation (DLF/DLE), etc. Intelligent Electronic Devices (IED) are key enablers of these applications, by receiving and sending data from/to electronic multifunction meters, digital relays, controllers, etc. [EPRI04] In addition, it is expected that ADA will bring new sophisticated algorithms, providing more intelligence and coordination at the central SCADA/DMS systems, as well as distributed control capabilities of the grid.

The evolution of the smart grids is based on the Advanced Metering Infrastructures (AMI), which heavily depends on Automated Meter Reading (AMR) devices, which support the following functionality [SCISSOR15a]:

- Measure power usage in real-time;
- Monitor and inform the Distributed System Operator (DSO), the customer and third parties about power quality;
- Track and keep a record about customer usage parameters;
- Remotely connect and disconnect customers from the power grid;
- Send out alarms to the DSO in case of technical issues and failures;
- Energy prepayment;
- Remotely receive and install firmware upgrades to incorporate new functionality;
- Anti-tampering and fraud detection;
● Remotely customizable load limit feature.

Despite broad use of SCADA systems for running critical industrial processes (such as, railway signaling systems, food production, wind turbines, oil refineries, etc.), SCADA has not been designed for security, which consequently might lead to intellectual property theft, loss of visibility of the system sensor readings, data integrity, loss of control of the plants, etc. Hence, the authors in [SCISSOR15a] survey SCADA literature related to security:

● threats analysis (physical threats, human threats, network incident taxonomy [HOLO98] [BLAC09] [HAHU03], SCADA attacks taxonomy [ZHJS11]),
● protection methods and tools (e.g. network monitoring, IDS),
● security risks (e.g. ICT network, or DSS control center compromised; modification, deletion, or disclosure of customer private data; shutdown of the corporate LAN).

In addition, the authors in [SCISSOR15a] perform the security risk analysis, with the aim to define possible attacks in smart grids and ICS/SCADA systems, and to find ways to protect these systems within the existing constraints. Finally, our analysis of the industrial requirements for security, privacy and IPR/data protection takes into consideration findings from the interviews and presentations given by the network of IoT experts, industrial leaders and IoT early adopters in Austria, that we met during IoT Salzburg meetups held in March 10, 2015 and May 19, 2015, in Salzburg.

4. Conclusion
This report presents our first results on IoT technology tracking, plus the investigation of the major industrial requirements for security, privacy and IPR/data protection in the IoT. The activities carried during the project follow up (1) IoT Salzburg meetups, and our discussions with the experts from several production companies in Austria, which help us understand the expectations and security needs of the industrial practitioners. (2) We additionally consult the report “Technological assessments, scenarios and requirements” of the SCISSOR project (Horizon 2020 N644425, February 2015 - January 2018), particularly its Task 2.1 that examines the real-world expertise of the consortium partners related to the ICS (Industrial Control System) technologies, devices, and network infrastructure. SCISSOR considers the requirements related to security services. (3) Finally, we base our investigation on the relevant literature from the areas such as privacy, security, IPR and
data protection in the IoT. We explore the main privacy and data protection challenges in the IoT, security challenges (security vulnerabilities, detection and prevention of attacks in IoT, data provenance, game-theoretic algorithmic approach to IoT security), and finally, security concerns and requirements from the side of the industrial practitioners.

In parallel with the requirements gathering, we also tracked developments in the technical and scientific fields. Our summary on emerging IoT platforms, over the course of the project is summarized in the additional report: “Report on emerging IoT platforms for Industrial Internet”. 
Appendix 1: Interview with Industrial Experts in Austria

This is a template prepared to support the Interview with Industrial Experts in Austria, to identify major expectations about Industry 4.0 and their security needs related to the IoT. It has been created to serve as a basis for the discussion with the industrial experts in Austria, during IoT Salzburg meetups.

(IoT4Industry project; full name of the project: Secure, Privacy-Preserving Agents for the Industrial Internet, project website: http://iot4industry.salzburgresearch.at/)

Interviewee: _____________________________________________
Date: ___________________________________________________

1. Which of the following barriers are the main inhibitors of adoption of the Industrial Internet in your company? Please rank from 5 (least relevant) to 1 (most relevant).

☐ Security concerns

☐ Privacy concerns

☐ Technology immaturity

☐ Interoperability, or lack of standards

☐ ____________________________________________________ (e.g. lack of skilled workers, lack of education in the area, lack of business models (or how do company profit from the IoT))

2. What is likely to be the primary M2M protocol used in your company for the IoT in 5 years?

__________________________________________________________
__________________________________________________________
3. What are your major privacy concerns?
☐ Loss of control over data
☐ Inferences derived from the data
☐ Repurposing of original data processing
☐ Bringing out behavior patterns and user profiling
☐ Remaining anonymous when using IoT based services
☐ Quality of the user’s consent (lack of information about the use of the data)
☐
☐

4. What are your major security concerns?
☐ Security weaknesses of any hardware and IoT devices
☐ Security vulnerabilities (e.g. insecure Web interface, insecure network services)
☐ Security attacks
☐ Insufficient authentication/authorization
☐ Lack of transport encryption
☐ Insecure mobile interface
☐ Insecure cloud interface
5. Have you experienced any of the following security attacks (please, ignore this question if you find it being sensitive information)

☐ Selective forwarding

☐ Sinkhole attacks

☐ Sybil attacks

☐ Wormholes

☐ HELLO flood attacks

☐ Acknowledgement spoofing
References


Workshop on Selected Areas of Cryptography Kingston, ON, 1994, pp. 120-128.


routing protocol for ad hoc networks," Department of CS, Rice University, TR01-383.


(KOOL05) G. M. Køien and V. A. Oleshchuk, 2005. "Location Privacy for Cellular Systems;


http://www.w3.org/TR/prov-o/ (last accessed April 2014)


[RAML15] Referenzarchitekturemodell Industrie 4.0 (RAMI4.0), APRil 2015.


