2017-00176 - PhD Position / Predictive Security Monitoring for Large-Scale Cyber-Physical Systems

Contract type: Public service fixed-term contract  
Level of qualifications required: Graduate degree or equivalent  
Fonction: PhD Position

About Inria
Inria, the French National Institute for computer science and applied mathematics, promotes "scientific excellence for technology transfer and society": Graduates from the world's top universities, Inria's 2,700 employees rise to the challenges of digital sciences. With its open, agile model, Inria is able to explore original approaches with its partners in industry and academia and provide an efficient response to the multidisciplinary and application challenges of the digital transformation. Inria is the source of many innovations that add value and create jobs.

Context
Within the framework of a partnership:
- European H2020 project SecureIoT

Assignment
- Scientific Context:
In last years, Internet-of-Things became a reality with numerous protocols, platforms and devices [8] developed and used to support the growing deployment of smart services: smart-home, -transport, -health, -city... and even the rather usual rigid systems with industry 4.0. Providing new services has required the development of new functionalities with as underlying goals to have more power- and compute-efficient devices which can embed various sensors. Obviously, IoT also supposes a full infrastructure to guarantee the efficiency of communications and processing of information. The embedded devices are thus completed by access points, routers, servers, etc. At the higher levels services are developed and provided to the users. This ecosystem is very rich and cannot be controlled by a unique entity, e.g. services are often developed by third parties, manufacturer of embed devices are different to those providing connectivity. As a result, such a complex system is naturally a source of potential threats and real cases recently demonstrates that IoT can be affected by naive weaknesses [7,6]. At Inria, we even demonstrated how simple and cheap IoT device can be used to take over the control of a Z-Wave home installation in a silent manner [2]. Therefore, security is paramount of importance. In the last decade, many IoT architectures have been proposed, such as the reference model IoT-A [3], including security modules. However, as highlighted before, security cannot be guaranteed without failure or by-design and this is true with evolving ecosystems such as IoT, with now the emerging trend of using fog-based architecture rather than well-established cloud models. To enhance security, one option is to redesign an IoT architecture with stronger security but this will face the same problems as before, since some security issues can appear afterwards and the trade-off between security features and their cost is hard to establish mainly multiple IoT devices. Maintaining the architecture with new security elements would be therefore required but a remaining problem is the large number protocols or platforms that already exist. Nowadays, the only viable solution is so to provide new security mechanisms that could be composed on demand and deployed in any IoT platforms by the operators, the integrators or the vendors rather than developing protocol- or architecture-centric security solutions.

- Bibliography:
Main activities

The objective of the thesis is to define and evaluate a predictive security framework for IoT in a context where the devices can be spread over multiple domains and could leverage heterogeneous protocols. This will be achieved in the context of the European H2020 project SecureIoT which provides three use cases with real platforms (industry, healthcare, connected and autonomous cars). The challenging goal is to (1) observe evidences of future attacks or misuse by (2) collecting and integrating heterogeneous data.

Several tasks need to be achieved to reach this objective and they are tightly coupled:

- Definition of a programmable monitoring plane for IoT devices in a fog scenario. Collected data needs to be meaningful for the predictive analysis but the monitoring and collection process needs to be scalable. The monitoring must not be protocol centric and has to gather information from multiple layers (network, system, services). In particular, discovering the full topology and interactions of the infrastructure is crucial for security analysis but cannot be easily retrieved due to the mobility of devices, the use of multiple protocols and limited capacity of devices. First, an hybrid passive-active monitoring of IoT devices will be developed to recover topology, network interactions, system and service information (fingerprinting) while keeping a low footprint. Machine learning will support such task as being proved to be efficient in similar context [4]. Second, based on gathered information, monitoring capabilities of the infrastructure will be automatically inferred and exposed to enable composition of monitoring functions. The inferred capabilities will be associated with an IoT security knowledge base to establish relationships between the observed events and existing attack patterns and vulnerabilities. A recurrent problem is then the quantity of data to handle and forward to decision points, e.g. to edge servers. Aggregation or fusion of monitoring data should occur in an efficient manner to support the security objectives while limiting the induced overhead and privacy risks.

- Predictive security analysis. A security engine needs to perform an in-depth analysis of all collected data for performing a security assessment of the observed system, orchestrate the monitoring plane itself (e.g. to request additional data and optimize the placement of probes and monitoring agents) and support decision on counter-measures to be applied. This decision includes identifying which counter-measures to be considered, but also determining how to compose and where to deploy them [9, 10]. Such element is the central piece of the security of the IoT architecture and supposes a real-time execution and preferably in a streaming mode by integrating new data on the fly and raising alerts when detecting a deviant pattern, which can lead to an erroneous or forbidden states. Machine learning and data mining techniques will be leveraged but have to carefully crafted to our context. Data is far from being normalized, properly formatted and easily comparable unlike numerical vectors. For instances, using traditional distances on retrieved data will be meaningless. A challenge resides in finding appropriate data representation, associated normalization techniques and metrics, which will allow us to handle all data in an holistic manner rather than analyzing each data type separately. For instance, network traffic and application logs should be analyzed together as they may reveal correlations. Graph- or process-mining [5] will be investigated to represent dependencies.

Skills

- Skills and profile:
  - Required qualification: Master degree, preferably in computer science
  - Knowledge and skills in the following fields will be appreciated: networking, security, machine learning

Benefits package

- Subsidised catering service
- Partially-reimbursed public transport
- Social security
- Paid leave
- Flexible working hours
- Sports facilities

Remuneration

1982,00€ brut per month