Knowledge-based System for Cybersecurity in IoT environments

- LISITE (Laboratoire d'Informatique, Signal et Image, Électronique et Télécommunication)
- ISEP (Institut Supérieur d'Electronique de Paris)

Advisors:

- Prof. Lina Mroueh (lina.mroueh@isep.fr)
- Dr. Nouredine TAMANI (nouredine.tamani@isep.fr)
- Dr. Saad EL JAOUHARI (saad.el-jaouhari@isep.fr)

Context

The progress in the area of embedded systems has favored the emergence of so called "smart objects" or "Things". The latter incorporate, in a context of low energy consumption, various wireless communication capabilities combined with a microcontroller driving sensors and / or actuators. The Internet of Things (IoT) conceptualizes this new environment based on traditional networks connected with objects as specific components of the real world. However, this also leads to inherit all the problems of security, privacy and trust already present in the Internet. These problems rest with stronger acuity in this new environment, because of its particular characteristics.

From cybersecurity point of view, Open Web Application Security Project (OWASP), as part of the Internet of Things Project, has published a list of IoT attack surface areas, which can be summarized into three categories [6] as follows:

- **Devices related threats**: IoT constrained devices can be the first line by which attacks can be initiated. The vulnerabilities can emerge from its firmware, physical interfaces, web interface, or from the enabled network services. Moreover, malicious peers can exploit the unsecure default settings, outdated components, and unsecure update mechanisms, etc.
- **Communication channels related threats**: Attacks can originate from the channels that connect IoT components with one another and with the external word (the Internet). Furthermore, the protocols used in IoT systems may contain some security issues that can threaten the entire systems. IoT systems are also targets to known network attacks such as denial of service (DoS), spoofing, eavesdropping and MITM attacks. An example of the amplitude of the attacks using or against IoT networks has been demonstrated in the Dyn cyberattack. A Mirai malware has infected IoT many devices such as printers, IP cameras, residential gateways and even baby monitors and used as botnets in order to raid the Dyn services [7].
- **Applications and software related threats** concerns mainly the vulnerabilities that can affect web applications and related software for IoT devices. Vulnerable web applications may be exploited in order to steal sensitive and valuable user credentials or to push malicious firmware updates which may compromise the whole network.

As the IoT domain is relatively new, the current available solutions are based on the adaptation of regular cybersecurity solutions from regular networks where the nodes are endowed with memory and computation resources that provide them the ability to perform complex processing that can help filter traffics, monitor events, and so on. It is not the case for IoT devices, which have very little resources. A security layer, generally implemented at the application level as a middleware, becomes then of a great importance to build trust in IoT environments.

There are many approaches to build such a middleware, such as adopting authentication and access control platforms to ensure IoT data confidentiality and integrity. But it is not sufficient to protect both the devices and the data from network attacks such as denial of service, for instance.

We explore, in this thesis, the potentialities of a knowledge-based system to protect a cyber-physical system by ensuring a one-in-all cybersecurity service for IoT environments.

Knowledge Based System (KBS) for Cybersecurity

Knowledge Representation and Knowledge Engineering have achieved major advancements at both theoretical and algorithmic/computation levels. From ontologies to existential rules, passing through linked open data, Knowledge representation languages and tools evolved to define Web 3.0.

A knowledge base founded on existential rules (KB) consists of 3 sets of elements: set of rules, set of facts et set of negative rules, seen also as constraints. The reasoning over an KB requires to derive all the possible facts via a derivation process called chasing, which is an NP-Complete problem.

A derivation process, limited to a subset of possible existential rules, has been designed and needs to be explored in the field of cybersecurity. From a theoretical point of view, decidability is still to be the main issue to deal with in the field.

From the state of the art, we can identify the engineering work introduced in [5] for event detection based on knowledge graphs. The research is focused on event detection for decision making in cyber-physical systems. These event are complex and characterized by their richly attributed signals, spatio-temporal correlation and contextual and environmental factors They propose Kronos (for Dynamic Knowledge extraction), which supports automatic extraction integration and search from the context rich events (in particular ones with anomaly-based event model) and semantic knowledge from sensors data streams, which can lead administrators to actionable knowledge. This is done via online event discovery via anomaly detection over data stream, a windows-based correlation inference which provides the users with analytical event queries to do the search in the rich context.

Besides, knowledge-based systems are modeled and harnessed in two different situations

1. Smart Building Monitoring [3]: the application, developed as the proof of concept, exploits the data collected by a Building Information Management (BIM) of a smart building and raises alerts in case of detection of anomalies. An anomaly is defined as the infringement of a negative rule.
2. IoT Personal data protection [2]: the application developed as a semantic firewall for IoT data protection in order to protect the data exchanged between a smart appliance or an IoT and an external entity, which consumes the collected data.

Besides the aforementioned applications, anomaly detection is one of the main user application domains, which can be considered. Indeed, the KBS framework can be applied in diverse areas such as fraud detection in documents, in cybersecurity by analyzing both ingoing and outgoing connections within a network, in malevolent behaviors in a system to access both physical and logical sensitive resources such as network components, IoT devices, databases, and so on.

Starting from simple use-cases where the world is considered as closed (CWA), we need to develop decidable algorithms to reason about the situation the system faces. From the data collected and semantically modeled, we need to develop algorithms capable to:

1. **Learn normal situations**: in order to detect anomalies, we need to distinguish between what is normal in the system and what it is not. The data collected by sensors and any IoT objects about the environment can serve for training learning algorithms and build a first step to modeling what we can name « normal situations » in the system [2].
2. **Detect anomalies**: once normal situations are known; they could be interpreted as rules in the knowledge base. As long as the rules are not contradictory, the system is said to be consistent and by extension safe. The problem of anomaly detection in a given system is mapped to the consistency checking problem in a knowledge base. A first attempt towards such a mapping has been carried out in the context of smart building management in [3]. A thorough theoretical study of the aforementioned relationship between anomalies and inconsistencies is one of the objectives of the proposed PhD thesis, in order to identify all the relevant characteristics and properties of such a mapping [4] from computational point of view.

Challenges

From theoretical point of view, the proposed topic is three-fold:

- Knowledge representation to give semantics to the world considered in the study: IoT environment.
- Learning and reasoning within this framework to extract from the collected data and the attached semantics the main situations seen as normal, then perform reasoning in order to assess any evolution within the system to detect anomalies.
- To go beyond the detection and to start the prevention in general and for IoT environment in particular.

From application point of view, the objectives, challenges and tasks to achieve in this PhD thesis are as follows:

- To conceive a real IoT environment in order to experiment on and to produce datasets.
- To design and implement a knowledge-based anomaly detection to protect the system from attacks.
- Tests and assessment of the effectiveness of the proposed approach.

Required skills

- Engineering degree or a Master's degree major in cybersecurity
- Expertise in knowledge representation and data engineering
- Expertise in Machine Learning approaches
- Good programming skills in C++

Bibliography

[1] **A Guided Tour of Artificial Intelligence Research. Volume I: Knowledge Representation, Reasoning and Learning**. Chapter: **Reasoning with Ontologies**, by Meghyn Bienvenu, Michel Leclère, Marie-Laure Mugnier and Marie-Christine Rousset. Chapter: **Knowledge Engineering**, by Nathalie Aussenac-Gilles, Jean Charlet and Chantal Reynaud. Pierre Marquis, Odile Papini and Henri Prade Editors. Springer Nature Switzerland AG 2020. https://doi.org/10.1007/978-3-030-06164-7

[2] N. Tamani, Y. Ghamri-Doudane: **Towards a user privacy preservation system for IoT environments: a habit-based approach.** FUZZ-IEEE 2016: 2425-2432

[3] N. Tamani, S. Ahvar, G. Santos, B. Istasse, I. Praça, P-E. Brun, Y. Ghamri-Doudane, N. Crespi, A. Bécue: **Rule-Based Model for Smart Building Supervision and Management.** SCC 2018: 9-16

[4] A. Arioua, N. Tamani, M. Croitoru, J. Fortin, P. Buche: **Investigating the Mapping between Default Logic and Inconsistency-Tolerant Semantics.** ICAISC 2015: 554-564

[5] M. H. Namaki, X. Zhang, S. Singh, A. Ahmed, A. Foroutan, Y. Wu, A. K. Srivastava, A. Kocheturov, "**Kronos: Lightweight Knowledge-based Event Analysis in Cyber-Physical Data Streams**," *2020 IEEE 36th International Conference on Data Engineering (ICDE)*, Dallas, TX, USA, 2020, pp. 1766-1769, doi: 10.1109/ICDE48307.2020.00165.

[6] OWASP Internet of Things (IoT) Project
https://wiki.owasp.org/index.php/OWASP_Internet_of_Things_Project Last visited [15-03-2021]

[8] The Dyn Attack - How IoT Can Take Down the "Global Information Grid" Back Bone (Part I) https://www.rsa.com/en-us/blog/2016-10/the-dyn-attack-how-iot-can-take-down-the-global-information-grid-back-bone-part-i, Last visited [15-03-2021]