STUDIES IN COMPUTER SCIENCE NO 14, DOCTORAL DISSERTATION

JOSEPH BUGEJA

ON PRIVACY AND SECURITY IN SMART CONNECTED HOMES

MALMÖ UNIVERSITY

ON PRIVACY AND SECURITY IN SMART CONNECTED HOMES

Malmö University,
Studies in Computer Science No 14,
Doctoral Dissertation

ISBN 978-91-7877-163-9 (print)
ISBN 978-91-7877-164-6 (pdf)
DOI 10.24834/978917877164-6
Holmbergs, Malmö 2021

JOSEPH BUGEJA

ON PRIVACY AND SECURITY IN SMART CONNECTED HOMES

Malmö University, 2021
Faculty of Technology and Society
Department of Computer Science and Media Technology

Studies in Computer Science

Faculty of Technology and Society
Malmö University

1. Jevinger, Åse. Toward intelligent goods: characteristics, architectures and applications, 2014, Doctoral dissertation.
2. Dahlskog, Steve. Patterns and procedural content generation in digital games: automatic level generation for digital games using game design patterns, 2016, Doctoral dissertation.
3. Fabijan, Aleksander. Developing the right features: the role and impact of customer and product data in software product development, 2016, Licentiate thesis.
4. Paraschakis, Dimitris. Algorithmic and ethical aspects of recommender systems in e-commerce, 2018, Licentiate thesis.
5. Hajinasab, Banafsheh. A Dynamic Approach to Multi Agent Based Simulation in Urban Transportation Planning, 2018, Doctoral dissertation.
6. Fabijan, Aleksander. Data-Driven Software Development at Large Scale, 2018, Doctoral dissertation.
7. Bugeja, Joseph. Smart Connected Homes: Concepts, Risks, and Challenges, 2018, Licentiate thesis.
8. Alkhabbas, Fahed. Towards Emergent Configurations in the Internet of Things, 2018. Licentiate thesis.
9. Paraschakis, Dimitris. Sociotechnical Aspects of Automated Recommendations: Algorithms, Ethics, and Evaluation, 2020, Doctoral dissertation.
10. Tegen, Agnes. Approaches to Interactive Online Machine Learning, 2020, Licentiate thesis.
11. Alvarez, Alberto. Exploring the Dynamic Properties of Interaction in Mixed-Initiative Procedural Content Generation, 2020, Licentiate thesis.
12. Alkhabbas, Fahed. Realizing Emergent Configurations in the Internet of Things, 2020, Doctoral dissertation.
13. Ashouri, Majid, Towards Supporting IoT System Designers in Edge Computing Deployment Decisions, 2021, Licentiate thesis.
14. Bugeja, Joseph, On Privacy and Security in Smart Connected Homes, 2021, Doctoral dissertation.

Electronically available at:
http://mau.diva-portal.org

In memory of my brother, Ryan.

ABSTRACT

The growth and presence of heterogeneous sensor-equipped Internet-connected devices inside the home can increase efficiency and quality of life for the residents. Simultaneously, these devices continuously collect, process, and transmit data about the residents and their daily lifestyle activities to unknown parties outside the home. Such data can be sensitive and personal, leading to increasingly intimate insights into private lives. This data allows for the implementation of services, personalization support, and benefits offered by smart home technologies. Alas, there has been a surge of cyberattacks on connected home devices that essentially compromise privacy and security of the residents.

Providing privacy and security is a critical issue in smart connected homes. Many residents are concerned about unauthorized access into their homes and about the privacy of their data. However, it is typically challenging to implement privacy and security in a smart connected home because of its heterogeneity of devices, the dynamic nature of the home network, and the fact that it is always connected to the Internet, amongst other things. As the numbers and types of smart home devices are increasing rapidly, so are the risks with these devices. Concurrently, it is also becoming increasingly challenging to gain a deeper understanding of the smart home. Such understanding is necessary to build a more privacy-preserving and secure smart connected home. Likewise, it is needed as a precursor to perform a comprehensive privacy and security analysis of the smart home.

In this dissertation, we render a comprehensive description and account of the smart connected home that can be used for conducting risk analysis. In doing so, we organize the underlying smart home devices according to their functionality, identify their data-collecting capabilities, and survey the data types being collected by them. Such is done using the technical specification of commercial devices, including their privacy policies. This description is then leveraged for identifying threats and for analyzing risks present in smart connected homes. Such is done by analyzing both scholarly literature and examples from the industry, and leveraging formal modeling. Additionally, we identify malicious threat

agents and mitigations that are relevant to smart connected homes. This is performed without limiting the research and results to a particular configuration and type of smart home.

This research led to three main findings. First, the majority of the surveyed commercial devices are collecting instances of sensitive and personal data but are prone to critical vulnerabilities. Second, there is a shortage of scientific models that capture the complexity and heterogeneity of real-world smart home deployments, especially those intended for privacy risk analysis. Finally, despite the increasing regulations and attention to privacy and security, there is a lack of proactive and integrative approaches intended to safeguard privacy and security of the residents. We contributed to addressing these three findings by developing a framework and models that enable early identification of threats, better planning for risk management scenarios, and mitigation of potential impacts caused by attacks before they reach the homes and compromise the lives of the residents.

Overall, the scientific contributions presented in this dissertation help deepen the understanding and reasoning about privacy and security concerns affecting smart connected homes, and contributes to advancing the research in the area of risk analysis as applied to such systems.

Keywords: Smart Connected Homes, Internet of Things, Smart Home Devices, Smart Home Data, Threat Identification, Risk Analysis, Privacy, Security, Vulnerability Assessment, Mitigations, Threat Agents.

PUBLICATIONS

Included Publications

Paper 1. **Bugeja, J.,** Jacobsson, A., Davidsson, P. (2016). On Privacy and Security Challenges in Smart Connected Homes (pp. 172–175). In: *Proceedings of the 2016 Intelligence and Security Informatics Conference (EISIC 2016)*. IEEE. https://doi.org/10.1109/EISIC.2016.044

Paper 2. **Bugeja, J.,** Jacobsson, A., Davidsson, P. (2017). An Analysis of Malicious Threat Agents for the Smart Connected Home (pp. 557–562). In: *Proceedings of the International Conference on Pervasive Computing and Communications Conference (PerCom Workshops 2017)*. IEEE. https://doi.org/10.1109/PERCOMW.2017.7917623

Paper 3. **Bugeja, J.,** Jacobsson, A., Davidsson, P. (2018). Smart Connected Homes. In: *Internet of Things A to Z Technologies and Applications* (1st ed., pp. 359–384). IEEE John Wiley and Sons. https://doi.org/10.1002/9781119456735.ch13

Paper 4. **Bugeja, J.,** Jönsson, D., Jacobsson, A. (2018). An Investigation of Vulnerabilities in Smart Connected Cameras (pp. 537–542). In: *Proceedings of the International Conference on Pervasive Computing and Communications Conference (PerCom Workshops 2018)*. IEEE. https://doi.org/10.1109/PERCOMW.2018.8480184

Paper 5. **Bugeja, J.,** Davidsson, P., Jacobsson, A. (2018). Functional Classification and Quantitative Analysis of Smart Connected Home Devices (pp. 1–6). In: *Proceedings of the Global IoT Summit (GIoTS 2018)*. IEEE. https://doi.org/10.1109/GIOTS.2018.8534563

Paper 6. **Bugeja, J.,** Jacobsson, A., Davidsson, P. (2018). An Empirical Analysis of Smart Connected Home Data (pp. 134–149). In: *Proceedings of the Internet of Things (ICIOT 2018)*. Lecture Notes in Computer Science, vol 10972. Springer. https://doi.org/10.1007/978-3-319-94370-1_10

Paper 7. **Bugeja, J.,** Jacobsson A. (2020). On the Design of a Privacy-Centered Data Lifecycle for Smart Living Spaces (pp. 126-141). In: Friedewald M., Önen M., Lievens E., Krenn S., Fricker S. (eds) Privacy and Identity Management. *Data for Better Living: AI and Privacy. Privacy and Identity 2019*. IFIP Advances in Information and Communication Technology, vol 576. Springer. https://doi.org/10.1007/978-3-030-42504-3_9

Paper 8. **Bugeja, J.,** Jacobsson, A., Davidsson, P. (2020). Is Your Home Becoming a Spy? A Data-Centered Analysis and Classification of Smart Connected Home Systems (pp. 1–8). In: *Proceedings of the 10th International Conference on the Internet of Things (IOT 2020)*. ACM. https://doi.org/10.1145/3410992.3411012

Paper 9. **Bugeja, J.,** Jacobsson, A., Davidsson, P. (2020). A Privacy-Centered System Model for Smart Connected Homes (pp. 1–4). In: *Proceedings of the International Conference on Pervasive Computing and Communications Conference (PerCom Workshops 2020)*. IEEE. https://doi.org/10.1109/PerComWorkshops48775.2020.9156246

Paper 10. **Bugeja, J.,** Jacobsson, A., Davidsson, P. (2020). PRASH: A Framework for Privacy Risk Analysis of Smart Homes. Submitted for journal publication.

Personal Contribution

For all publications above, the first author was the main contributor

in terms of research idea formulation, planning, execution, and manuscript writing. This also included the creation of design science artefacts, formal analyses, investigations, and data curation. For Paper 4, the second author was responsible for assisting with the data collection.

Other Publications

In addition to the papers incorporated in Part II of this thesis, there are also articles that may be relevant to mention as they have indirectly affected the compilation of this thesis. Thus, the following articles are related but not included in the dissertation.

Vogel, B., Kajtazi, M., **Bugeja, J.,** Varshney, R. (2020). Openness and Security Thinking Characteristics for IoT Ecosystems [Security and Privacy in IoT Systems (SPIoTS)]. *Information* 11, no. 12: 564. MDPI. https://doi.org/10.3390/info11120564

Bugeja, J., Jacobsson, A., Spalazzese R. (2020). On the Analysis of Semantic Denial-of-Service Attacks Affecting Smart Living Devices. In: *Arai K., Kapoor S., Bhatia R. (eds) Intelligent Computing. SAI 2020. Advances in Intelligent Systems and Computing, vol 1229*. Springer. https://doi.org/10.1007/978-3-030-52246-9_32

Kebande, V. R., Alawadi, S., **Bugeja, J.,** Persson, J. A., Olsson, C. M. (2020). Leveraging Federated Learning Blockchain to counter Adversarial Attacks in Incremental Learning. In: *Proceedings of the 10th International Conference on the Internet of Things (IoT 2020 Companion)*. ACM. https://doi.org/10.1145/3423423.3423425

Bugeja, J., Bahtijar, V., Jacobsson, A., Varshney R. (2019). IoTSM: An End-to-end Security Model for IoT Ecosystems. In: *Proceedings of the International Conference on Pervasive Computing and Communications Conference (PerCom Workshops)*. IEEE. https://doi.org/10.1109/PERCOMW.2019.8730672

Kebande, V. R., **Bugeja, J.,** Persson, J. A. (2019). Internet of Threats Introspection in Dynamic Intelligent Virtual Sensing (pp. 557–562). In: *Proceedings of the 9th International Conference on the Internet of Things (Workshop on Cyber-Physical Social Systems (CPSS2019))*. *ArXiv Preprint*. arXiv:2006.11801

Kajtazi, M., Vogel, B., **Bugeja, J.,** Varshney, R. (2018). State-of-the-Art in Security Thinking for the Internet of Things (IoT). In: *Proceedings of the 13th Pre-ICIS Workshop on Information Security and Privacy (WISP 2018)*. *5*. https://aisel.aisnet.org/wisp2018/5

ACKNOWLEDGEMENTS

A few exceptional opportunities arise in any person's life. For me, this PhD journey was one of them. I have leveraged my professional experience in and passion for the cybersecurity field into exploring the ramifications, in terms of privacy and security, of transforming a home into a smart Internet-connected home. The current and emerging market calls for an ever greater need for such an understanding. However, none of this would have been possible without the support of many people.

First, I would like to express my deepest gratitude and appreciation to my principal academic advisor and dean, Dr. Andreas Jacobsson, for his unwavering support, advice, guidance, and for affording me the chance to make this work a reality. Thank you for giving me so many opportunities to network with various researchers, participate in exchanges and internships abroad, and for granting me the privilege to deliver a course on information security at Malmö University in Sweden. I also like to particularly express my heartfelt gratitude and thanks to Prof. Paul Davidsson, director of the Internet of Things and People Research Center (IOTAP), who has also been my academic advisor. Despite your prolonged commitments, you always found the time to review my work and provide such excellent insight and information to help me further it from there. Thank you for sharing your wealth of knowledge and expertise with me and applying it to this dissertation. I am so grateful to have had you both, Andreas and Paul, as my study leaders. Thank you both for your dedication, professionalism, and for mentoring my academic development with great care and interest from start to finish.

Before thanking any others, I would like to especially thank the research profile IOTAP funded by the Knowledge Foundation and Malmö University in collaboration with various industrial partners. I would also like to thank all the members of the research profile project "Intelligent Support for Privacy Management in Smart Homes", in particular Verisure, for their support at the beginning of my research.

Furthermore, I would like to convey my thanks and appreciation to my examiner Dr. Jan Persson, and PhD review group members – Prof. Bengt J. Nilsson and Prof. Helena Holmström Olsson – for their help in

reviewing this dissertation, assessing my individual study plan, and also for bestowing invaluable advice necessary to ensure the rigorousness and relevance of my work.

My sincere thanks also goes to Dr. Martin Boldt, senior lecturer in Computer Science at Blekinge Institute of Technology (BTH) in Sweden. Thank you for accepting the role to be my opponent for my penultimate PhD seminar. The lengthy discussion we had on my dissertation was core to improving the dissertation's quality and coherence. Thanks also for inviting me to BTH and for connecting me with your colleagues.

Thank you also to Prof. Mohan Kumar, professor and chair of the Department of Computer Science at Rochester Institute of Technology (RIT), for being my opponent at my licentiate seminar. I appreciate you traveling from the US to discuss my research. Your feedback was instrumental in helping me improve the overall significance of the dissertation. Thanks also for making my international research exchange at RIT so amazing. I have learned a lot from this experience.

I am also much obliged to Dr. Bo Peterson, head of the department at Malmö University, for stepping in to ensure a healthy work and research schedule on my part. The final part of this dissertation would not have been completed without your direct support. On that note, I would also like to thank, in particular, Dr. Åse Jevinger, director of studies and senior lecturer at Malmö University; Henriette Lucander, section head at Malmö University and licentiate of technology; and Susanne Lundborg, research liaison officer at Malmö University.

I would also like to thank Dr. Bahtijar Vogel, a senior lecturer at Malmö University. You have presented me with opportunities for collaboration with Lund University and the University of British Columbia. The various discussions we had were inspiring and truly important for my academic career and going forward.

A special thank you also to Dr. Victor Kebande, a previous postdoctoral researcher at the IOTAP, especially for helping me with the information security course management during my final year of studies. It was nice working with you, even though for brief.

I would like to thank all my fellow research scholars affiliated with the IOTAP and the Department of Computer Science and Media Technology at Malmö University for their kindness and friendship. In particular, I would like to mention Dr. Radu-Casian Mihailescu, Dr. Johan Holmgren, and Zahra Ghaffari, lecturer and MSc in Computer Science, for having been there since the beginning of my PhD journey.

Last, but certainly not the least, I would also like to acknowledge my appreciation to my parents, Michael and Rita, and my sister, Stephanie. Thanks for your love, support, and for always believing in me.

Malmö University, December, 2020 Joseph Bugeja

CONTENTS

Part I.

COMPREHENSIVE SUMMARY

CHAPTER

INTRODUCTION

Wireless cameras within a [smart] device such as the fridge may record the movement of suspects and owners.

Mark Stokes
Former Scotland Yard head of digital forensics

In 1991, Weiser, introduced the term of ubiquitous, also known as pervasive, computing in his seminal paper "The Computer for the 21st Century" [1]. His vision was that computing should be integrated seamlessly in the background, allowing people to employ it when needed without shifting their attention from their main tasks. Eight years later, the idea of the Internet of Things (IoT) was coined by Ashton while working on the Auto-ID Center at the Massachusetts Institute of Technology. Ashton originally coined the term "Internet of Things" in a presentation he made at Proctor Gamble (PG), where he made the first association between the new idea of Radio Frequency Identification in PG's supply chain and the emerging Internet [2].

The Internet of Things (IoT) can be thought of as a computing paradigm where physical objects (e.g., devices, vehicles, and buildings)

are augmented with identifying sensing/actuation, storing, networking, and processing capabilities, allowing them to communicate with each other and with other devices and services over the Internet to accomplish some objective [3]. These objects are typically referred to as smart objects, smart devices, or simply as connected things. Smart objects tend to be called "smart" to indicate their capability to make sense of and leverage their environment, including the ability to engage in autonomous decision-making [4]. By design, smart objects can interact with other smart devices and people, and in the process, collect and exchange data with each other, including remote servers on the Internet.

Over recent years, the technology behind the IoT has led to innovative applications broadly categorized under two domains, namely, industrial IoT and consumer IoT [5]. Industrial IoT concerns deployments in industrial and control environments such as Supervisory Control and Data Acquisition (SCADA), smart cities, water systems, and critical infrastructures in general. Consumer IoT refers to deployments that are targeted at individuals or families.

Smart homes belong to the consumer IoT domain. Essentially, a smart home is a residential space composed of a network of devices that provide "electronic, sensor, software, and network connectivity inside a home" [6]. This setup gives the residents the ability to get information, control, and automate different parts of the home and improve the quality of daily chores in a residence, possibly from anywhere and anytime, typically over the Internet through a smartphone application [7]. Effectively, smart homes are an application of the broader smart living concept which is focused on applying technologies to daily life to increase efficiency, affordability, and sustainability of resources [8].

As smart home technology has evolved, connected devices have been networked to form IoT systems of systems (SoS). These systems have enabled connected devices to provide different services to their users going beyond home automation. Services include that of enhancing the residents' overall security and safety, entertainment, health and fitness, and more. In this dissertation, we refer to IoT-based smart homes as smart connected homes.

In recent years, the development of the IoT and smart connected homes, have been gaining increasing momentum due to a range of advancements in wireless protocols, sensors, processors, data analytics, cloud technologies, and the widespread availability of smartphones. According to Statista, it is estimated that in 2025 the number of IoT connected devices will surpass 75 billion units [9], potentially generating $4T to $11T in economic value by 2025 [10], and by 2020 the digital data around the world, of which 10% of this amount would come from IoT devices, will reach 44 zettabytes [11]. The global smart home market was projected to reach approximately $53.45 billion in 2022 [12], with

an estimated compound growth of more than 14.5% from 2017 to 2022. This demonstrates the increasing consumer demand and rising adoption of this technology.

Noting the potential of the market, commercial information and communication technologies (ICT) organizations like Google, Apple, and Facebook, who previously did not have a presence in consumer home automation technologies, have launched their products, e.g., Nest smart thermostat, platforms, e.g., Apple HomeKit, and entertainment solutions, e.g., Facebook Portal, to compete on the market for building the next generation of smart connected homes. Today, the IoT is part of daily life, with smart assistants like Siri, Alexa, and Google Assistant, being added to everyday home appliances and utilities such as toasters, thermostats, lights, and the list goes on.

1.1. Research Setting

The home is a deeply meaningful and human place. It is considered a person's castle, sanctuary, refuge, and for many, a supportive environment in which to grow up and discover oneself. Fundamentally, a home is a place wherein one can expect core physical needs, including privacy and security, to be protected [13]. Nonetheless, the introduction of connected devices inside homes brings forth different social concerns [14]. Two of these concerns are privacy and security.

Privacy and security have been subject to long-term academic effort and have been recognized in technology regulation worldwide [15]. In the IoT context, privacy and security for connected devices have been identified as significant research challenges and priorities by the European Union (EU) Commission that have to be addressed for the benefit of society [16]. Accordingly, a growing body of work has sought to understand the privacy and security concerns associated with IoT devices, including their deployment in homes [17].

Despite the increasing efforts to make connected devices more privacy-preserving and secure, security experts have raised concerns about the privacy and security risks with connected devices in homes [18][19][20]. These concerns are supported by recent high-profile attacks, such as the Mirai distributed denial-of-service (DDoS) botnet attack that disrupted the Internet for millions of users [21]. However, a worrying concern is that even the most secure IoT devices tend to continuously and inconspicuously collect personal and sensitive data about the residents and their home and share it with third-parties located over the Internet [22]. Such can happen even when security-enhancing mechanisms are in place [23].

Privacy protection is a fundamental human right that is essential in the

functioning of democratic societies [24]. This is acknowledged by Article 12 of the Universal Declaration of Human Rights [25], which protects an individual from "arbitrary inferences with his privacy, family, home or correspondence," and "attacks upon his honour and reputation". Nonetheless, as the numbers, types, and sophistication of connected devices and the data being collected by them are increasing at a fast pace, this is threatening the inviolability of the home. Arguably, more sensors embedded in smart connected homes may signify the end of privacy in the home and may lead to different expectations of security for protecting smart home devices.

In this dissertation, we want to explore how the nature of privacy and security has been transformed as the home got connected to the Internet. Complicating this overarching research goal is that privacy, especially in comparison to security, has remained largely unexplored in the smart home context [26], privacy overlaps with security [27], and research work about smart homes tends to be segmented by multiple academic disciplines each bringing their own concepts and assumptions (cf. Paper 5).

The research questions related to the mentioned research goal are formulated in Chapter 3.

1.2. Research Scope

Researching privacy and security in the context of smart connected homes is challenging. Smart connected homes encompass a broad range of technologies and systems. Moreover, privacy and security have many facets that can broaden the extent to which smart connected homes are accordingly explored in this dissertation. Due to this breadth, scoping was essential to the research.

- **Smart connected homes.** The smart home is a SoS that incorporates a range of technologies. A SoS is "a collection of systems, each capable of independent operation, that interoperate together to achieve additional desired capabilities" [28]. In this dissertation, we focus the technical analysis work on commercial off-the-shelf (COTS) smart home systems. These systems can be installed in existing homes with relative ease making them accessible to a wide variety of users. Nonetheless, in modeling smart home systems, we generalize to also support Do-It-Yourself (DIY) systems and smart home lab projects.

- **Privacy.** Privacy is inherently subjective, cultural, and contextual [13]. Privacy in different interpretations has been referred to as a human right varying within different contexts. Four categories

of privacy have been identified by Clarke [29] as per different contexts: privacy of the person, privacy of personal data, privacy of personal behaviour, and privacy of personal communication. Information privacy is a construct that combines communication privacy and data privacy [29]. We primarily focus the research on information privacy (cf. Section 2.2.1). Nonetheless, we discuss the other categories of privacy when discussing the impact of some privacy violations.

- **Security.** Security is concerned with intentional failures, with the root cause of security problems being the human nature [30]. Discussions on the topic of security also tend to encompass other aspects, such as that of reliability and safety. We consider reliability and safety, which deal with accidental failures in a system, and physical harm, respectively, to be out of scope in this research. Instead, we focus primarily on information security (cf. Section 2.3.1). Information security is connected to the notion of information privacy. Nevertheless, when identifying certain privacy and security attacks we also discuss their implications on the safety of the occupants.

In this dissertation, the focus is on creating an understanding of privacy and security in smart connected homes. We postulate that such an understanding can contribute to policy improvements and lead to more privacy-preserving and secure smart homes.

1.3. Main Contributions

The research in this dissertation presents a number of contributions to science. In summary, these contributions are a combination of: theoretical contributions in the form of design science artefacts [31], namely, new models, a construct, and a method; empirical research contributions consisting of new findings based on systematically observed data; and survey contributions that concern review and synthesis of the work done in the smart home field to expose trends, themes, and gaps in the literature.

We present an overview of the main contributions in relation to the smart connected home, categorizing them according to their main research domain, as follows:

State-of-the-Art Devices and Data. We survey the connected devices and data being collected by them, and propose a construct that organizes the smart connected home devices.

- *Taxonomy and analysis of connected devices.* We propose a functional taxonomy of connected devices, including an analysis of their hardware and software capabilities. The taxonomy and analysis are developed empirically from the technical specifications of different commercial devices.
- *Classification and analysis of connected devices and their apps.* We propose a classification of smart connected home systems (i.e., the connected device and its corresponding app) according to their data collection capabilities. The classification and analysis are developed empirically using the embedded sensors found in connected devices and their accompanying apps.
- *Analysis and classification of collected data.* We analyze and categorize the data collected by smart connected home devices. The classification and analysis are developed empirically by investigating the privacy policies of different manufacturers of commercial smart home devices.

Threat Identification and Analysis. We create models that can help identify and analyze threats affecting smart connected homes.

- *Privacy-centered system model.* We propose a system model that captures the dynamics of a smart connected home, including the properties and requirements for modeling privacy. The model is a formal description of the smart connected home allowing for the identification of privacy threats.
- *Privacy-centered data lifecycle.* We propose a model that captures the different data phases of a smart connected home. The model extends a standard data modeling technique with annotations and processes, allowing for identifying privacy threats and strategies for mitigating those.

Risk Modeling. We create a model and a method that can help analyze risks affecting smart connected homes.

- *Threat agent model.* We propose a model that identifies the different malicious human threat agents targeting the smart connected home, including their motivations and capabilities. The model can be used to understand the different kinds of attacks to expect when deploying IoT technologies inside homes.
- *Framework for modeling and analyzing privacy risks.* We propose a framework that can be used for dynamically discovering attack paths in a smart connected home deployment, including measuring privacy risks in a quantitative manner. The framework is a method that helps in automatically determining the privacy risk exposure of a smart connected home.

Mitigations and Challenges. We survey mitigations that help reduce risks in smart connected homes and identify challenges in designing privacy-preserving and secure smart connected homes.

- *Identification of security challenges and their mitigations.* We identify state-of-the-art challenges and their mitigations in smart connected homes. Different challenges are explored and mitigations that function at different architecture layers, including during the design and development phase, of smart connected homes.

Overall, the mentioned contributions are researcher-oriented but are of potential interest to practitioners, system analysts, and software developers working in the smart home domain. Specifically, the presented contributions add novelty to the areas of human and societal aspects of security and privacy, analysis and design of emerging devices and systems, and system security.

In this dissertation, other contributions that are considered side-contributions are included. These are described in the actual publications, i.e., in Part II of the dissertation. Moreover, the contributions are elaborated on in Chapter 5 of Part I.

1.4. Dissertation Outline

This dissertation is a compilation thesis that is divided into two parts – Part I and Part II. In Part I, we provide an extensive introduction to the dissertation area and summarize answers to the posed research questions. In Part II, we include the ten peer-reviewed publications that form the actual research of this dissertation.

The rest of Part I is organized as follows. In Chapter 2, we introduce the conceptual framework needed to understand the rest of the dissertation, including a description of the smart connected home and fundamental notions connected to privacy and security. In Chapter 3, we present the research questions addressed in this dissertation. In Chapter 4, we describe the methodology that has been applied during the research process of this dissertation. In Chapter 5, we summarize the answers to the posed research questions categorizing them according to their primary research topic. Finally, in Chapter 6, we conclude the dissertation and identify some opportunities for future work.

CHAPTER

CENTRAL CONCEPTS AND RELATED WORK

If you know the enemy and know yourself, you need not fear the result of a hundred battles.

Sun Tzu
The Art of War, 5 B.C.

In this chapter, we present an introduction to the main concepts relevant to this dissertation. Also, we outline the literature relevant to the research setting and scope.

We start this chapter by describing the smart connected home, particularly its enabling technologies and existing systems. Next, we describe privacy and security concepts, including the topic of threats, risks, and mitigations. Related literature work and gaps are presented when discussing privacy and security.

2.1. Smart Connected Homes

There is no generally standard definition or consensus of what a "smart

home" is. The definition of the term varies according to the technology or the functionality the home implements. Several alternative names have been used across the years to refer to the smart home, e.g., "intelligent living", "digital house", "smart environments", and more [32]. A common, simple, and established definition has been developed by the UK Department of Trade and Industry (DTI). The DTI's Smart Home project defined a smart home as "a dwelling incorporating a communication network that connects the key electrical appliances and services, and allows them to be remotely controlled, monitored or accessed" [33].

While DTI's definition works for most smart home scenarios, nowadays homes are evolving into smart living spaces where the living environments and type of services offered go beyond smart homes to include other aspects of human living such as education, work, and social life. Furthermore, in addition to the automation and control aspects, smart homes are also providing proactive services, e.g., providing timely physical support, to the residents through sensor technologies and algorithms based on Artificial Intelligence (AI) and machine learning.

2.1.1. Smart Home Evolution

The history of smart home technology goes back many years. In fact, the actual term "smart home" was originally coined by the American Association of House Builders in the year 1984 [34].

Although the concept of a smart home has been around for a while, the smart home has only taken momentum in recent years. Here, an essential milestone for making the development of smart home technology a reality was when electricity was brought to households at the beginning of the 20th century [35]. Electricity stimulated the introduction of new equipment in the home, e.g., electrical machines and domestic appliances.

Another important landmark introduced in the last quarter of the 20th century was the introduction of information technology in the homes. This created new possibilities for exchanging information sparking the evolution of smart home technology [35].

More recently, we observe another important milestone in the smart home evolution brought about by the IoT and the ensemble of technologies surrounding it, in particular innovations in sensors and microelectronic devices.

We group the smart home evolution into two phases: Pre-IoT smart connected homes and IoT smart connected homes.

Pre-IoT Smart Connected Homes. The first smart home devices emerged in the late 1960s with the invention of the Electronic Computing Home Operator (ECHO IV) and Kitchen Computer [36]. The

ECHO IV was used for family bookkeeping, inventory taking, and climate control [37]. A year later, the Kitchen Computer came out. This machine allowed people to store recipes (cf. Paper 1).

In the 1970s, X10 was established and used as a standard communication protocol for wiring houses for home automation. This is often touted as the ancestor of home automation.

When "personal computers" appeared in the consumer market in the late 1970s, controlling and automating home appliances was mainly conducted by hobbyists in DIY projects [38]. Here, some form of remote control was possible by decoding Dual-Tone Multi-Frequency (DTMF) signals through telephone lines [39]. However, the turning point in smart home development occurred when the domestic Internet, appeared on personal computers in the mid 1990s [40].

At the same time, in the 1990s, ubiquitous computing technologies arose. Using these technologies, researchers started developing smart home projects all across the globe [41]. In the majority of the cases, these homes were real-life living space testbeds [41].

We refer to these types of systems as "smart homes". Such systems tend to use proprietary protocols, offer no or somewhat limited integration facilities, and allow few control options to end-users, typically limited to local (in-house) control and using specific controllers.

IoT Smart Connected Homes. In recent years, the IoT became a commercial reality allowing for home devices to be remotely observed and controlled through the Internet. Hereunder is a chronological list of some of the most popular commercial smart home systems appearing in the consumer market in 2010 and onwards:

In 2010, the Nest Learning Thermostat[1] (nowadays owned by Google) entered the smart home scene. This device functions as a smart thermostat learning the residents' preferred house temperature and adjusting it automatically. Nest is sometimes identified as the flagship product that introduced the contemporary smart home [42].

In 2014, Amazon launched its first (1st generation) smart speaker system – Amazon Echo[2] – that could control the smart home by using the voice as an input channel and providing a full ecosystem of programmable skills (capabilities). In the same year, SmartThings (later acquired by Samsung) issued a device that functioned as a residential gateway (sometimes called a hub or home controller) linking together different connected devices at home [43].

In 2015, Apple released HomeKit[3]. This is a software framework and an interoperability protocol that allows different devices to communic-

[1] `https://nest.com/thermostat` [accessed December 31, 2020].

[2] `https://en.wikipedia.org/wiki/Amazon_Echo` [accessed December 31, 2020].

[3] `https://developer.apple.com/homekit` [accessed December 31, 2020].

ate with each other.

In 2016, Google released Google Home (nowadays called Google Nest), a smart speaker[4] with the Google Assistant built-in. Two years later, Apple released Apple HomePod[5], a smart speaker using Siri as a voice-assistant.

Today, as of 2020, the smart home market is filled with all kinds of devices. In particular, we observe devices equipped with sophisticated sensors, leveraging AI technologies, and harnessing the capabilities of other connected systems. For instance, we find Facebook Portal[6] allowing for smart video calling using integrated smart cameras; Samsung Ballie[7] acting as a personal robotic butler following the users around the home and helping them in their chores; and Verisure's connected home alarm[8] allowing for the detection of critical situations in homes, for example, a home intrusion attempt.

We refer to these types of systems as "smart connected homes". These systems tend to be Internet-connected, feature multimodal user interface channels, various networking protocols, and "intelligent" logic, making it possible to make some autonomous decisions.

The focus of this dissertation is on this category of smart homes.

2.1.2. Existing Smart Home Systems

Several smart home systems have been conducted over the last several decades. We divide these systems into two types: laboratory systems and commercial systems. Laboratory systems are fundamentally used for research purposes and often involve dedicated housing facilities, whereas commercial systems involve platforms and COTS products retrofitted into actual finished homes.

Laboratory Systems. These function as live-in labs or experimental houses commonly developed to study human behavior and in-home automation [34]. Typically, laboratory systems involve monitoring and recording of residents' activities and interactions in a purposely designed setup. Some prominent examples are: Aware Home project [44], MavHome project [45], GatorTech Smart House project [46], House_n project [47], and PlaceLab [48].

[4] `https://www.techradar.com/reviews/google-home` [accessed December 31, 2020].

[5] `https://www.apple.com/homepod-2018` [accessed December 31, 2020].

[6] `https://portal.facebook.com` [accessed December 31, 2020].

[7] `https://news.samsung.com/us/samsung-ballie-ces-2020` [accessed December 31, 2020].

[8] https://www.verisure.se/landingpages-blocks/verisure.html [accessed December 31, 2020].

Most of the mentioned systems are linked to the pre-IoT smart connected homes. These are essentially testbeds for technological components and an early attempt to bring the ubiquitous computing paradigm into the home.

Commercial Systems. Nowadays, there is a growing trend of developing ready-to-use COTS solutions. These are sometimes referred to as smart home gateway (hub) ecosystems. Here, the idea is to provide the residents with a central gateway capable of connecting and interacting with various connected devices present in a home.

Various large manufacturing companies have launched similar products such as Samsung Smart Home, Google Home, Apple HomePod, and many more. Most of these systems tend to leverage the cloud infrastructure to deploy their services. Another characteristic of these systems is that they support several different applications (beyond that of home automation), tend to be programmable, and allow end-users options to customize them according to their liking.

Main Observations and our Research Focus. Commercial systems, in comparison to laboratory systems, tend to be installed in actual residences. Here, the residents tend to have an active role in selecting and bringing into their household the technology they desire and often install it themselves without relying on a professional [49]. Moreover, commercial systems tend to bring forth some added complexities.

Some of these complexities are related to: the sophistication of the underlying and evolving technologies; new dynamics for example in relation to the ecosystem of stakeholders and services; and challenges for example given the variety of unregulated and unstandardized devices. Thereby, this raises interesting research opportunities for scholarly and industry communities.

Given these factors, in this dissertation, we put the attention on commercial systems. These systems are associated with the IoT-based smart connected homes we explored earlier.

2.1.3. Application Areas of Smart Connected Homes

The smart connected home encloses multiple services (applications) commonly belonging to energy, entertainment, security, and healthcare [50]. In smart connected homes, the connected devices form the core of the concept, as they create the foundation of the user experience.

There is a remarkable number of connected devices available in the consumer market. These devices, in particular through the use of sensors, collect data on which decisions are made. Connected devices deal with different types of data, some of which can be personal and sensitive. The

Table 1.: Smart connected home application areas and examples of devices and their corresponding data types.

Application area	Device type	Collected data types
Energy and resource management	Plug, light bulb, shower head water meter	Location data, consumption data
Entertainment systems	Music player, TV, audio speaker	Voice commands, features accessed, search queries
Health and wellness	Blood pressure monitor, scale	Body metrics, social networking services related
Networking and utilities	Gateway/hub, wireless signal extender	Network/connectivity-related data, personal preferences
Human-machine interface	Remote control	Battery charge level
Household appliances and kitchen aids	Vacuum cleaner, oven, floor mopper	Location data, operating schedules
Security and safety	Cloud camera, door bell, smoke detector	Contact preferences, location data, interaction data
Sensing	CO_2 sensors, rain sensor, air quality sensor	Sensor status

smart connected home application areas alongside examples of devices and types of data captured by each is summarized in Table 1.

In Part II of this dissertation, we elaborate on the application areas, device types, and the collected data types of devices.

2.1.4. Smart Connected Home Components

The technical composition of a smart connected home consists of various components that interact with each other, exchanging data about the state of the home, the environment, and its residents' activities and behavior. These components tend to be operated or managed by different stakeholders, typically serving three types of users: data subjects, data controllers, and data users (cf. Paper 7).

Data subjects, typically represent the smart home residents whose data are processed. Data controllers, sometimes also referred to as "data holders", "data curators", or "data processors", are the entities, typically service providers or device manufacturers, that collect, store, and process data generated by connected home devices and data subjects. Data users represent the entities that access the released data.

In terms of components, a generic smart connected home (see Figure 1 for an illustration) consists of the following:

- **House.** This represents the set of physical locations forming the residence area, e.g., the apartment, including all the areas that are within the curtilage[9].
- **Connected device.** These are hardware units, e.g., domestic appliances, lights, or sensors, that can sense, actuate, process data, and communicate. Three core devices are sensors, actuators, and end-user client devices. Sensors detect, monitor, and measure properties of objects such as room temperature. Actuators perform actions in the physical environment, such as switching on or off lights. End-user client devices such as smartphones are commonly used by the data subjects to interact and manage the smart connected home. We also refer to connected devices as smart devices or nodes.
- **Gateway.** The gateway (hub) is a specialized connected device that collects data from other connected devices and commonly acting as the central point of connectivity for end-users to manage the connected devices. Gateways connect the local IoT (home area) network to the Internet, often via the residential router. Moreover, gateways can act as network bridges translating between different communication protocols. Some connected devices, such as smart speakers, also provide built-in gateway functionality.
- **Cloud.** The cloud is used by some connected devices as a backend for storing and processing data, and sometimes also as a mechanism for integrating different standalone connected devices. De-

[9] The dwelling area — called the curtilage — is defined as an area that is attached to a house but extends outwards beyond the four walls of the house. Curtilage tends to be protected legally, and in some cases, also constitutionally (e.g., in the US through the Fourth Amendment [51]).

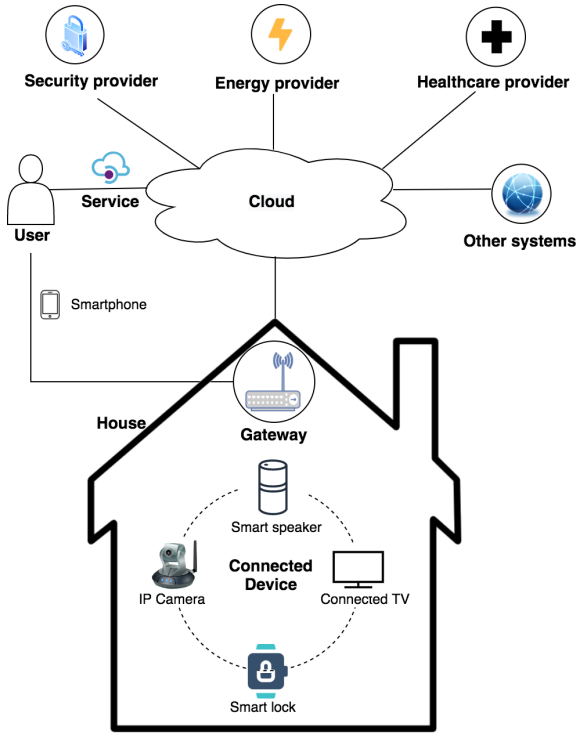

Figure 1.: A typical smart connected home architecture. Typically, data subjects (a type of user) access connected devices through the help of a smartphone. Commonly, the interaction between the smartphone and a connected device is mediated through a gateway device.

pending on the adopted architecture and communication model, some connected devices can send sensed data directly to the cloud. However, this is often facilitated through the gateway.

- **Service.** Software applications that provide the facility to control, manage, and operate the smart home system. Services provide the facility for the smart connected home to implement different application areas and to integrate with other connected systems. Some services may expose APIs (Application Programming Interfaces), allowing for controlling connected devices over standard Internet protocols. Cloud services, in particular IFTTT (If This Then That), allow the facility to interconnect different devices together and for running automations.
- **User.** The stakeholder that uses and benefits from the services offered by the smart connected home. Typically, this represents the data subjects but may also include other entities, for example, those satisfying the role of data controllers or data users.

More details about the composition and the architecture of a smart connected home are found in Paper 3. A more detailed description and formalization of the smart connected home is found in Paper 9 and Paper 10.

2.1.5. Technical Capabilities of Connected Devices

The smart connected home consists of a vast array of connected devices. Smart connected home devices vary significantly in terms of their hardware and software capabilities [52].

At one end, there are constrained devices, such as smart locks, with low CPU, memory, and battery power specification. At the other end, there are high-capacity devices [53], such as gateways, that are typically powered by the main supply and have higher specifications allowing them to support programmatic access, remote administration, and different communication options ranging from wired to wireless protocols, remote access, and more.

In Table 2, we show some of the capabilities of commercial smart connected home devices, in terms of their supported protocols, services, and embedded sensors. As shown in Table 2, the sophistication of connected devices varies between the different brands and types of devices. The supported capabilities of connected devices also affect the requirements for the effective deployment of privacy and security-enhancing mechanisms.

In this dissertation, we classify smart connected devices and analyze their capabilities, particularly in Paper 5 and Paper 8.

2.2. Privacy

In this section, we focus on privacy and its particular relevance to the smart connected home. Consequently, we start by describing the concept of privacy from a scholarly perspective and considering as well recent developments from a regulatory perspective. Then, we introduce the topic of privacy threats, risks, and privacy-enhancing mechanisms that can help mitigate risks.

2.2.1. The Concept of Privacy

Privacy is a concept that has been central throughout human history. Although many definitions of privacy have been put forth, there is no universally agreed-upon definition of this concept. Nonetheless, privacy

Table 2.: Specifications of smart connected home devices.

Device type	Network protocols	Services	Embedded sensors
Facebook Portal	Wi-Fi, Bluetooth	Alexa built-in, WhatsApp, Messenger	Microphone, Camera
Amazon Echo	Wi-Fi, Bluetooth	Alexa built-in, API, IFTTT, Web browser, mobile apps	Microphone
Nest Learning Thermostat	Wi-Fi, Bluetooth, Thread	API, IFTTT, mobile apps	Temperature, Humidity, Proximity, Motion, Ambient light, Optical
August Smart Lock	Bluetooth	IFTTT, mobile apps	Door sensor

concerns existed long before the advent of computers and cybertechnology. Indeed, the Code of Hammurabi, dating to about 1754 B.C., enshrining Ancient Babylonian law, already protected the home against intrusion by others [54]. In modern times, privacy scholars have argued about the definition and scope of "privacy" since at least the late 19th century [55].

Warren and Brandeis, in their seminal article "The Right to Privacy", published in 1890, articulate the right to privacy as "a right to be let alone" claiming individuals possess an "inviolate personality" in the face of media surveillance. At that time, this involved photography, a technology which they recognized as intruding private spaces [56].

Westin, as the computer era was emerging, developed Warren and Brandeis description of privacy into the broader notion of "informational privacy" defining it as the "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others" [57]. This mainly emphasized the notion of control of the data subjects over their data.

Altman expanded on Westin's idea of privacy to include a contextual notion that explained why individuals at times seek privacy and at other

times, disclose information [58]. Altman understood privacy as a type of "boundary regulation" process wherein people try to achieve their ideal privacy state by using some control mechanisms to regulate interaction with others dynamically.

Nissenbaum expanded Altman's notion of privacy further in her theory of Contextual Integrity (CI). Essentially, CI claims that privacy is always provided in context. Thus it may change its meaning or intent when it is used in another setting (context). Different contexts are governed by different social norms that govern information flow within and out of that context. CI asserts that privacy is violated if contextual norms of appropriateness or norms of information flow are breached [59].

In this dissertation, given the pragmatic and descriptive nature of CI, we adopt it in some of the recent publications, in particular in Paper 9 and Paper 10, for dealing with privacy violations.

2.2.2. Privacy Laws and Data Protection

There are many legislative and regulatory compliance issues regarding privacy, and government organizations are taking a significant interest in IoT privacy from a legal perspective. The cornerstone of most modern privacy laws and policies are the Fair Information Practices (FIPs) developed in the 1970s [60] [61]. Essentially, the FIPs are a set of internationally recognized practices that govern the collection and use of personal data, serving as a model of ethical treatment of consumer data. In recent years, there have been ongoing worldwide efforts to enact or update privacy laws to address the challenges posed by digital technologies.

An important legislative requirement is the EU General Data Protection Regulation (GDPR) [62] which went into effect on May 25, 2018. The GDPR is a regulation that aims to safeguard the personal data rights of EU citizens and residents. Personal data in this context can include data that describes the person's economic, mental, or physical status. Sensitive personal data includes ethnicity, political opinions, religious beliefs, health, and genetic and biometric data. Personal data is frequently referred to as personally identifiable information (PII) in a US context. PII is any data item that can be traced back to the person of origin or concern.

Beyond the GDPR, data privacy and residency regulations are also increasing. For instance, countries like Russia, China, and Indonesia, to name a few, require that their citizen's data must be stored on physical servers located within the country's borders, while Europe is, at the time of writing, discussing a new regulation called ePrivacy Regulation (ePR) [63] that covers individuals' privacy in relation to electronic communications. Effectively, ePR adopts the definitions of privacy and data

introduced within the GDPR but acts to enhance them.

In the US, the California Consumer Privacy Act (CCPA) [64]; the CCPA can be considered the US counterpart of the GDPR; which became effective in January 2020, is designed to enhance the privacy rights of consumers living in the state of California. Recently, in December 2020, in response to data breaches and security concerns involving IoT devices, a new US law – the Internet of Things Cybersecurity Improvement Act of 2020 [65] – was officially signed into law requiring minimum security standards for IoT devices owned or controlled by the US Federal Government.

The mentioned legislative and regulatory requirements help address the growing threats to user privacy. Accordingly, software engineers are increasingly expected to give appropriate consideration to privacy when developing IoT solutions. Nonetheless, we observe that there is still a lack of proactive and integrative approaches to help safeguard the privacy of data subjects. In some of the included publications, particularly in Paper 7, we discuss how some of the data subjects' rights identified in the GDPR can be achieved by enhancing the data lifecycle phases (e.g., the collection, processing, and disclosure) with specific privacy-preserving processes.

2.2.3. Privacy Threats in Smart Connected Homes

A threat can be described as any potential occurrence that may result in an unwanted outcome for an entity or for a specific asset (resource) [66]. Privacy threats in the IoT are characterized by data lifecycle phases and actions that violate the data subjects' expectations [67].

A categorization consisting of seven privacy threats affecting IoT systems is provided by Ziegeldorf et al. in [68]. Three examples of threats covered in [68] are identification, tracking, and profiling. Another commonly used modeling technique for finding privacy threats is LINDDUN [69]. The "LINDDUN" acronym is derived from the categories of privacy threats it identifies, namely: Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness, and Non-compliance. An alternative threat modeling methodology that focuses on three privacy-specific threat categories, linkability, unawareness, and intervenability, is the Quantitative Threat Modeling Methodology (QTMM) [70]. Unlike Ziegeldorf et al. [68], both LINDDUN and QTMM were developed before the IoT and thus may not deal with IoT-specific threats.

Especially to explore what is at stake if a connected device is affected by a threat and for the sake of privacy it is core to understand the data that are collected by smart home devices. It is particularly vital to investigate the data collection phase, as it is arguably at that phase that

privacy threats arise. This is because, at that point, data are released and transferred from data subjects to connected devices.

In this dissertation, we investigate the data collection practices of real-world smart connected home devices in Paper 6. Additionally, we identify privacy threats affecting smart connected homes in Paper 9.

2.2.4. Smart Connected Home Privacy Risks

The International Electrotechnical Commission (ISO/IEC) 27005:2011 defines risk as the potential for a threat to exploit a vulnerability (weakness in a system) to cause harm to an asset. Informally, risk is expressed as a function of assets, threats, and vulnerabilities [30]. Risks can be formally investigated through a process known as risk analysis [30]. Prior literature identified several privacy risks of smart connected homes.

Arabo et al. [71] identified different privacy risks of connected devices, including the possibility of identity theft, social engineering attacks, social threats, and more. Jacobsson et al. [72], in an empirical risk analysis study, with emphasis on smart home automation systems, concluded that risks related to the human factor or the software components pose the highest risk. Apthorpe et al. [73] demonstrated through an experiment how to infer with reasonable accuracy privacy-sensitive in-home activities from smart homes containing commercially-available IoT devices.

End-user privacy concerns have also been examined. Zheng et al. [74] conducted semi-structured interviews with smart home owners noting that the users' desire for convenience features and connectedness dictate their privacy-related behaviors for dealing with external entities, such as manufacturers, involved in collecting IoT data. Zeng et al. [75] conducted semi-structured interviews with smart home residents identifying several privacy concerns these users have, such as continuous audio/video recording, data collection and mining, spying by other users in the home, and more. Zimmerman et al. [76] conducted semi-structured interviews with potential smart home users, noting that most participants were not convinced that their data was kept secure within a smart home and expressed concerns about potential attacks.

Other studies focused on specific smart home devices. Moody and Hunter [77] investigated how attackers relying solely on publicly available sources posted on the Internet can take advantage of weakly protected smart thermostats and exploit them to predict when a user is at home and more. Malkin et al.'s survey [78] about smart TVs revealed their respondents' uncertainty of data collection and usage as well as the common non-acceptance of data being repurposed (for advertising or other uses) or shared with third-parties. Huang et al. [79] studied pri-

vacy concerns of smart speaker technology users, outlining threats such as unauthorized voice purchases, unauthorized access to calendars and reminders, overheard call conversations, and concerns about external entities accessing and misusing collected data.

In this dissertation, we investigate vulnerabilities in commercial smart connected home cameras in Paper 4. Moreover, in Paper 10, we propose a generic framework for analyzing risks affecting smart connected homes.

2.2.5. Smart Connected Home Privacy-Enhancing Mechanisms

Several mechanisms can be implemented to mitigate privacy risks posed by connected devices, including those installed in homes.

Moncrieff et al. [80] proposed a framework to reduce improper access to residents' data in smart homes by dynamically managing access privileges based on contextual factors (e.g., the user's location and content of the ongoing conversation). Nurse et al. [81] outlined a framework for modeling risks in the smart home with a central goal of providing everyday users of IoT technologies intuitive ways to model risks and potentially assisting them with risk awareness in the context of smart homes. Apthorpe et al. [73] leveraged network traffic shaping through independent link padding to decrease the inference of privacy-sensitive activities captured by smart home devices.

More broadly, there are also various Privacy-Enhancing Technologies (PETs) to protect privacy. PETs, e.g., homomorphic encryption, can be described as technologies that help enforce legal privacy principles to protect and enhance the privacy of data subjects and users of information systems [82]. A different class of technologies aimed at safeguarding privacy is called Transparency-Enhancing Tools (TETs). Instead of focusing on data minimization, as is the case of PETs, TETs focus on providing users with increased visibility over aspects relevant to their data [83].

Privacy by Design (PbD) is an approach to systems engineering, initially developed by Cavoukian, that enables privacy to be incorporated throughout the entire engineering process from the earliest design stages to the operation of the productive system [84]. The notion of embedding PbD into the design of systems goes back to 2001, where Langheinrich [85] developed six principles (notice, choice and consent, proximity and locality, anonymity and pseudonymity, security, and access and recourse) for guiding system design based on fair information practices. Langheinrich discusses these generic principles in the context of ubiquitous computing in detail. The PbD approach has also been included in the GDPR, specifically in Article 32, obliging entities responsible for pro-

cessing personal data to implement appropriate technical and organisational measures.

In this dissertation, in Paper 7, we integrate to the smart connected home data lifecycle PbD strategies.

2.3. Security

Security is a critical requirement in a smart connected home. It is needed to prevent unauthorized access to people's homes, and consequently access to their personal and sensitive data. The topic of security, while sharing similarities and overlaps with that of privacy, is different.

In this section, we introduce the concept of security and explore its overlaps with privacy and other relevant terms. Following that, we discuss security threats, risks, and security-enhancing mechanisms that can help mitigate risks.

2.3.1. The Concept of Security

Security tends to be a critical aspect of information systems. It is especially important in IoT systems due to their connection with the physical world. Indeed, the coupling of sensors and actuators with the physical world together with the heterogeneous number of connected devices can amplify the severity and scale of security concerns more than in a typical ICT system. An ICT system tends to have more limited interaction with the physical world than an IoT system. Like privacy, security – particularly in the context of computing and cybertechnology – has no universally agreed-upon definition [86].

In 1975, Saltzer and Schroeder described security as "mechanisms and techniques that control who may use or modify the computer or the information stored in it" [87]. The authors' work describes eight design principles that are useful in secure system design and operation [88]. Regardless of the system being investigated, almost from its inception, the key objectives of security have been threefold: confidentiality, integrity, and availability protection for critical assets. This is known as the CIA triad (also called the AIC triad) of security [89]. All security controls, mechanisms, and safeguards are implemented to provide one or more of these protection types.

The purpose of confidentiality is to ensure that only authorized users can view information [30][90]. For instance, an attack on a robotic vacuum cleaner may result in disclosing a detailed map of a house's layout to unauthorized individuals. Integrity ensures that only authorized individuals can modify or delete information [30][90]. An integrity compromise may result, for instance, in a smart thermostat to increase in-

stead of decreasing the room temperature. The goal of availability is to ensure that the information or resource are accessible upon demand by an authorized user [30][90]. In the smart home context, a lack of availability to a connected device, for example, to a smart lock, may prevent the residents from entering their house.

Beyond the CIA model, security is frequently supplemented with other security objectives, e.g., authentication, accountability, and non-repudiation, to deal with more novel threats such as those hitting cyber-physical systems and for dealing with the increased use of networks [91]. Privacy is also sometimes added as a security goal by some researchers [89].

The Relationship between Security and Privacy. The concepts of security and privacy are not easy to separate. Indeed, there are some overlaps between security and privacy. Security's confidentiality requirement overlaps with privacy, mainly when the data are both personal and non-public. Security's integrity requirement overlaps with privacy's accuracy requirements. This is as both need to ensure that data are not altered without authorization. Security's availability overlaps with privacy since if the data are not available, then it cannot be accessed. Moreover, implementing some of the objectives of security, e.g., that of confidentiality, helps make a system privacy-preserving. However, by safeguarding confidentiality, integrity, and availability, that may not be sufficient for achieving privacy.

Indeed, some researchers [27] [24] argue that the concept of privacy has a broader set of concerns than security – for example, dealing with information flows, exposure, and identifiability, and subsequent use of personal data. Nonetheless, from a technology standpoint, privacy is reliant on security [92]. Security tends to be associated with the protection of confidentiality, integrity, and availability of assets, whereas privacy tends to be associated with the protection of data of (private) individuals across its lifecycle. As a conclusion, one can observe that the protection of data is a common theme in both security and privacy doctrines.

Security Overlaps with Safety. Sometimes the meaning of security overlaps with safety. Informally, safety is a broad term that is focused on the protection from all kinds of things. In contrast, security is a more specific term concentrating on the protection from malicious threat agents (bad actors). A more refined definition of safety associates it with the impact of system failures on their environment [30], including physical damage to a person. Even though security and safety are distinct concepts, there are overlaps between the two.

In the case of the smart connected home, some attacks, e.g., a data

poisoning attack on a connected health device (e.g., a wireless insulin pump), may lead to severe, possibly fatal, safety concerns to their human users, but also may affect the surrounding environment, e.g., by causing a fire due to overheating (e.g., if the connected temperature sensor does not detect rising indoor temperatures). Protection against safety threats is critical in the IoT, given the cyber-physical nature of these systems. Nonetheless, this is arguably more important in the industrial IoT context than in the consumer IoT context.

2.3.2. Security Threats in Smart Connected Homes

A threat is imposed or created on a specific asset by a threat agent [90]. There are essentially three different classes of threat agents: humans, technological, and environmental threat agents [93]. In terms of the human threat agent, instances of this can range from hackers to nation states, insiders to outsiders, and including those that can cause deliberate and accidental threats. Typically, for security analysis, we are interested in threat agents that cause deliberate threats.

To identify threats and threat scenarios different models or methodologies can be used. A threat model is a structured approach that allows a systematic identification and rating of threats that are likely to affect the system under consideration [94]. Complementary to LINDDUN [69] for privacy threats, STRIDE [95] is a threat model for security. In STRIDE, threats are categorized by the goals and purposes of the attacks, and namely, these are: Spoofing, Tampering, Repudiation, Information disclosure, DoS, and Elevation of privilege [95]. A different threat methodology that focuses on the threat agents, instead of the system or its assets, is the Threat Agent Risk Assessment (TARA) [96]. The TARA methodology identifies the threats that pose the greatest risk to a system by using a library of threat agent archetypes.

In this dissertation, we utilize a combination of system-centered and attacker-centered methods for identifying threats [95]. The attacker-centered method is used in Paper 2 and relies on the Industrial Control Systems Cyber Emergency Response Team (ICS-CERT) classification scheme [97] for identifying threat agents targeting the smart connected home. Moreover, we leverage, in particular in Paper 7 and Paper 9, Data Flow Diagrams (DFDs) [98] and black-box modeling as a type of formal modeling technique [99] for describing the smart connected home entities, modeling data flows, and identifying threats.

2.3.3. Vulnerabilities and Attacks

There have been several attacks against connected devices ranging from

proof-of-concept attacks to real-life breaches. Some examples of attacks that exploit the vulnerabilities of connected devices are Hajime, IoT Reaper, BrickerBot, and Mirai [21]. The Open Web Application Security Project compiled a list, called IoT Top 10 vulnerabilities [100], of the most common and most critical vulnerabilities in the IoT.

Typically, to discover vulnerabilities in a system, vulnerability analysis can be performed; vulnerability analysis is a process used by both attackers and defenders. For example, in the context of access control, vulnerability analysis attempts to identify the strengths and weaknesses of the different access control mechanisms and the potential of an attack to exploit a vulnerability therein. While threat modeling works at a higher abstraction level, vulnerability analysis works at a lower detail-oriented level. While we do not attempt to actively attack systems, we execute vulnerability scans to understand a system's exposure to certain threats (particularly information disclosure).

Different threat agents use different tools and methods for conducting vulnerability analysis and to conduct attacks. These can range from using specialized security distributions like Kali Linux[10] to online databases and search engines such as Shodan[11] and Censys[12], and more. Shodan – designed by Matherly in 2009 – is a vulnerability assessment tool [101] that crawls the Internet looking for Internet-connected devices (e.g., routers, printers, webcams) probing for their ports and indexing the retrieved banners and metadata. Censys has similar goals to that of Shodan but it uses different tools and methods to retrieve and document connected devices.

In this dissertation, given the flexibility, extensive documentation, and intuitive interfaces, we rely on Shodan to conduct a vulnerability assessment related to smart connected cameras (cf. Paper 4). Moreover, we use attack trees [102] to represent attacks from a formal modeling perspective (cf. Paper 10). Attack trees provide a formal and methodical tool for describing the security of systems by focusing on the attacker's perspective.

2.3.4. Smart Connected Home Security Risks

Cybersecurity risks arise from unauthorized activity related to the loss of confidentiality, integrity, or availability of a system or information asset [103].

Roman et al. [104] discussed the security risks resulting from the ever-increasing influx of IoT devices. The authors critically analyzed

[10] `https://www.kali.org` [accessed December 31, 2020].
[11] `https://www.shodan.io` [accessed December 31, 2020].
[12] `https://censys.io` [accessed December 31, 2020].

such emerging risks, their root causes, and mitigation mechanisms. Denning et al. [105] surveyed the landscape for connected devices in homes, identifying security attacks against such devices, and presented a framework for articulating key risks associated with particular devices in the home. Bastos et al. [106] identified several security risks in IoT technologies and protocols for smart home and smart city environments.

Security risk analysis and assessment; risk analysis is typically a risk assessment component, for the IoT domain has also been investigated in prior research. Risk analysis for smart home automation systems was proposed by Jacobsson et al. [72], emphasizing the security risks and mitigation mechanisms for such IoT deployments. Nurse et al. [107] presented different methodologies for assessing risks in the IoT context by considering the dynamics and unique features of IoT systems. Mohsin et al. [108] propose the IoTRiskAnalyzer framework to quantitatively analyze security risks of IoT systems based on a non-deterministic behavior model representing the threat agent.

In Paper 10, we present a framework that can be used for analyzing and evaluating risks affecting smart connected homes. While the proposed framework is concentrated on privacy, it also covers some security risks, given that some cybersecurity attacks may affect both privacy and security.

2.3.5. Smart Connected Home Security-Enhancing Mechanisms

As in privacy, information security can be enforced by technologies, which we will refer to as security-enhancing mechanisms. Even though these are not new, we will now go through some instances of security-enhancing mechanisms that are related to the work documented in this dissertation. These can be bundled together to provide a secure environment.

- **Encryption.** It is a mechanism that encodes information to hide the meaning or intent of a communication from unintended recipients. It can take many forms and apply to different electronic communication types, including text, audio, video files, and more. Encryption is often seen as a base premise for confidentiality.
- **Authorization.** It is a process of deciding whether a particular entity is allowed to perform a specific action on a system. Authorization is a fundamental concept in security as it ensures that only legitimate users and services can access protected data. Some examples of authentication methods are passwords, passcodes, and biometrics.
- **Firewall.** It is a technology that can filter network traffic, possibly

helping in preventing malicious traffic from the Internet from entering into a private network, as is the smart connected home network. Typically, firewalls filter traffic based on a defined set of rules, also called filters or access control lists.

- **Intrusion detection system.** It is a technology that can monitor and analyze network and host activity, possibly detecting when an attack has entered a system. An alternative to an intrusion detection system is an intrusion prevention system. An intrusion prevention system augments an intrusion detection system with the ability to automatically block detected attacks.

Mechanisms are put into place to mitigate potential risk. The mentioned mechanisms, when implemented and deployed properly, can eliminate a vulnerability or reduce the likelihood of a malicious threat agent to exploit a vulnerability, and thus reduce risk. Several types of other mechanisms exist, including non-technical, that can work together with the mentioned mechanisms to help implement security and contribute to implement privacy. We survey security-enhancing mechanisms that are of particular relevance to the smart connected home in Paper 1.

CHAPTER

3

RESEARCH QUESTIONS

Nothing is particularly hard if you divide it into small parts.

Henry Ford

In this chapter, we present the four research questions that have been tackled in this dissertation. These research questions aim to fill the gaps identified in previous research and to address real-world problems currently affecting smart connected homes.

As the number and heterogeneity of connected devices in homes increase rapidly, it becomes progressively challenging to gain a deep understanding of the smart connected home. Without a proper account of the smart connected home assets, particularly its connected devices and their data, it is challenging to recognize what is at stake when a smart home device gets compromised. Moreover, failing to identify and organize the assets makes it difficult to determine the threats associated with them. Nonetheless, complicating the efforts towards a common understanding are two factors: first, that multiple disciplines segment

the smart home field (e.g., networking, ubiquitous computing, and mobile computing); and second, that there is no standard taxonomy or classification of devices and data that takes into account the complexity and diversity of actual commercial devices. These complications lead to the first research question.

RQ1. How can smart connected home devices and the data collected by them be categorized?

Expanding on the findings of RQ1, it is beneficial to generalize the description of a smart connected home to allow for the systematic identification of threats, especially as relates to privacy. Nonetheless, despite considerable theoretical and practical contributions from the scholarly and industry communities, a standard representation, i.e., a system model, that accounts for the different smart connected home components and their data flows is missing. Without a system model, it is challenging to identify threats affecting smart connected homes methodically. Also, the lack of a system model further hinders understanding, at which point a threat arises. A system model can help find privacy and security issues early on and contribute to the development of threat models. This requirement leads to the following research question.

RQ2. How can smart connected homes be modeled to support threat identification?

Following the identification of different threats as part of RQ2, it is also useful to investigate the risks that smart home technologies pose to users. Providing methods for automatically conducting a risk analysis of IoT-based systems is necessary, given the IoT technologies' dynamic and evolving features. However, there is a lack of such methods, specifically those designed for smart connected homes. Also, the existing body of work on risk analysis tends to be focused more on security than on privacy, and most of the risk analysis methods require much manual work, making them prone to biases. Adding to this is also a shortage of attacker-centric models, i.e., models structured around threat agents and their capabilities. Failing to recognize the different threat agents may lead to the design and deployment of ineffective risk strategies. This knowledge gap leads to the following research question.

RQ3. How can privacy and security risks affecting smart connected homes be modeled and analyzed?

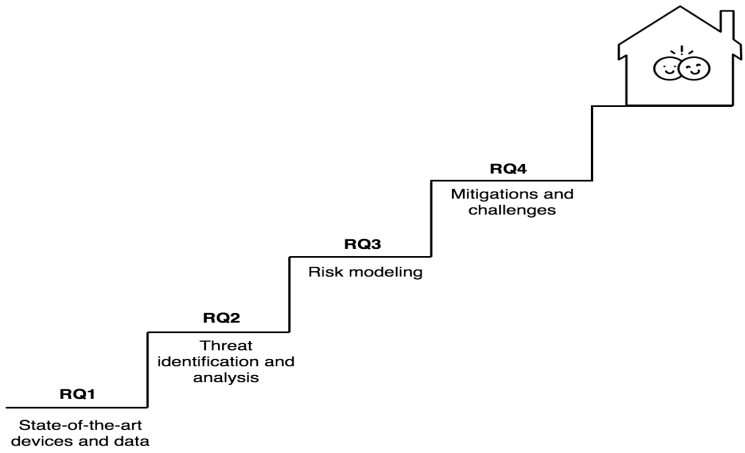

Figure 2.: An illustration depicting the research questions categorized according to their primary research domain. The answers to the research questions can progressively improve the knowledge to safeguard the inviolability of the residents inside their homes.

Having analyzed the risks of smart connected home technologies as part of RQ3, it is beneficial to identify challenges that make the design of privacy and security mitigations particularly challenging to implement in these environments. Privacy- and security-enhancing mechanisms have existed for decades, but homes are increasingly prone to cyberattacks leading to potential breaches of privacy and security. Accordingly, it is useful to survey existing mechanisms relevant to the smart connected home, including those that can be incorporated in the development lifecycle of a smart connected home. Having mechanisms integrated into the development lifecycle can minimize risks before targeting actual households or users. This requirement leads to the following research question.

RQ4. What are the challenges in mitigating privacy and security risks in smart connected homes?

In Figure 2, we illustrate the logical progression from RQ1–RQ4 indicating the main research domain being tackled by each research question. By answering the posed research questions, we can understand the risks posed by connected home technologies. Overall, this understanding may lead to the design of solutions that help preserve the inviolability of the home.

In this dissertation, although we include some aspects of behavioral science in some of the research, the focus is primarily on design re-

search. Specifically, we seek to provide an understanding that can help strengthen the foundations for developing more privacy-preserving and secure mechanisms intended for the smart connected homes.

CHAPTER

4

RESEARCH METHODOLOGY

There are as many scientific methods as there are individual scientists.

Percy W. Bridgman
Reflections of a Physicist, 1955

The objective of this chapter is to describe the strategy of inquiry used to answer the posed research questions. We start this chapter by identifying the research approach that was followed, and then introduce the adopted research strategies, data generation methods, and data analysis techniques used for studying the research questions. We conclude this chapter by identifying some of the followed ethical considerations.

4.1. Research Approach

There are, in general, two distinct research approaches: quantitative and qualitative [109]. Quantitative approaches assume a positivist (or post-positivist) paradigm [110][111][112]. Research conducted from positivism is expected to be objective, free of values, hypothesis driven, and

measurable. Qualitative approaches assume a constructivist (or interpretivist) paradigm, which supports the notion that there are many realities that are constructed as the researcher engages with participants [111][110][112]. In the last decades, mixed methods research has evolved as the third central research approach combining qualitative and quantitative research approaches, and rooting itself to the pragmatic paradigm [110][113][114]. The pragmatic paradigm is based on the proposition that there are several explanation of realities and researchers are free to choose all approaches available to understand the research problem [110].

In this dissertation, we primarily rely on mixed methods research as the overarching research approach for answering the research questions. From a qualitative perspective, we especially use it as existing theory is limited. As an example, when it comes to describing the smart connected home and its data, data flows, and processes. An explanation can be that the development of smart home technologies is more rapid and extensive that the range of research efforts can meet. Nevertheless, we also adopt research strategies that tend to be associated with the positivist paradigm. An example of this strategy is a survey. We use a survey, for example, to study the nature of the smart connected home, performing statistics to understand the technical characteristics of these homes.

A research strategy is the method used to answer the posed research questions. Typical examples of research strategies used in information systems and computing research include: survey, design and creation, experiment, case study, and action research [109]. In this dissertation, we use survey, design and creation, and case study as research strategies.

The adopted research strategies are explained in Section 4.2–Section 4.4.

4.2. Survey

Survey research strategy tries to systematically identify patterns in data so as to generalize to a larger population than the group being targeted [109]. They can serve different purposes – exploratory, description, or explanation [115]. Exploratory surveys are useful for attaining familiarity with a certain topic of interest. Description surveys focus on finding about the situation, events, attitudes, etc., that are occurring in a population. Explanatory surveys question the relations between variables.

In this dissertation, surveys were used for both exploratory purposes, e.g., to uncover challenges and risks in the form of a traditional literature survey, but also for description purposes, e.g., to describe the technical capabilities of devices.

There are different techniques for conducting surveys. In the study

of RQ4, we conducted a traditional literature review by examining documents manually. A different technique, namely, web and data mining, was used in the study of RQ1. The different surveys used for answering RQ1 are described as follows.

Device functionality and capabilities. The purpose of this survey (cf. Paper 5) was to identify the total number of distinct smart home functional categories, including the number of connected devices (and their percentile distribution) for each identified category. This survey involved a sample size consisting of 1,193. This number represented the entire dataset (as of May 2017) of smart home devices that were available in the utilized data source (SmartHomeDB[13]). After the functional categories were identified, we analyzed the technical capabilities belonging to each functional category. No sampling technique was used in this survey as statistics were computed on the entire dataset.

Device data. The purpose of this survey (cf. Paper 6) was to identify the type of data being collected by smart home devices, alongside the total number of device types that are associated with each. This survey involved a sample size of 87. As a sampling technique for selecting the devices to investigate, purposive sampling was used. Purposive sampling is a non-probabilistic sampling technique where the cases (devices) are selected based on their properties of interest [116]. The properties of interest for this survey reflected devices having distinct types and that feature the most reviews from the SmartHomeDB user community. The device types are derived from the device functionality and capabilities survey.

Device and app capabilities. The purpose of this survey (cf. Paper 8) was to develop a classification of smart connected home systems based on their data collection capabilities. A sample size consisting of 81 smart connected home systems was used for the analysis. This number represented the entire database (as of January 2020) of smart connected home systems that were available in the utilized data source (PrivacyNotIncluded[14]). Consequently, a number of systems were selected and their properties investigated. As a sampling technique for selecting the systems to investigate, stratified sampling was used. Stratified sampling is a probabilistic sampling technique in which the population is divided into separate subgroups and then samples are created by drawing subsamples from each of those subgroups [116]. This sampling technique was applied based on the functional category

[13] `https://www.smarthomedb.com` [accessed December 31, 2020].

[14] `https://foundation.mozilla.org/privacynotincluded` [accessed December 31, 2020].

of a system, and was used for selecting an unbiased sample of systems to investigate. The functional category of a system is derived from the device functionality and capabilities survey.

In order to collect data for the device functionality and capabilities survey and for the device and app capabilities survey, different software tools were used. Specifically, Python[15] programming language and scripts for that language, were developed to crawl SmartHomeDB and PrivacyNotIncluded product databases, and to extract capabilities.

4.3. Design and Creation

Design and creation research strategy focuses on the development of new information technology artefacts as a contribution to knowledge [109]. An artefact describes something artificial or something constructed by humans. Types of design science artefacts identified by March and Smith [117] and Hevner et al. [31] include constructs, models, methods, and instantiations.

In this dissertation, we propose a construct, models, and a method. These artefacts are described as follows.

- **Constructs.** These are concepts or vocabulary used in a particular IT-related domain. A new construct in the form of a taxonomy for classifying home devices was presented in relation to RQ1 (cf. Paper 5).

- **Models.** These combine constructs to abstract or represent a situation in such a way that it aids in problem understanding and solution development. Four new models were proposed for answering RQ1–RQ3. First, for RQ1, a model (cf. Paper 6) that categorizes the data collected by a smart connected home was introduced. Next, in relation to RQ2, we presented two models (cf. Paper 7 and Paper 9) that help in the identification and management of threats. Finally, for answering RQ3, we proposed a model (cf. Paper 2) that identifies the motivations and capabilities of different malicious threat agents.

- **Methods.** These provide guidance on the models to be produced and process stages to be followed to solve problems using IT. A new method was proposed in relation to RQ3. Essentially, an algorithm (cf. Paper 10) was proposed for generating attack trees given a system model of the smart connected home. This

[15] `https://www.python.org` [accessed December 31, 2020].

algorithm was used for computing attack metrics for each vulnerability. Overall, the method proposed for RQ3 is a framework for conducting privacy risk analysis of smart connected home systems.

The presented artefacts contribute to the scientific body of knowledge connected to the smart connected home, particularly that related to threats and risks. In Chapter 5, we explain these artefacts.

4.4. Case Study

Case study research strategy aims to obtain a detailed insight of one of the instances of a problem [109]. It is fundamentally an empirical investigation of a contemporary fact or situation within its real-life context [118]. Similar to survey research, there are three types of case studies: exploratory, descriptive, and explanatory [109].

In this dissertation, we performed a descriptive short-term case study in relation to RQ3 (cf. Paper 4). This was done to obtain an insight on vulnerabilities posed by real-life instances of smart connected cameras. Such devices are popular in smart connected homes found all across the world. Cameras were especially relevant to study as images/video feeds are often perceived as the most privacy-invasive technologies [119].

As a vulnerability identification and assessment tool, Shodan was primarily used together with a comprehensive database of security vulnerabilities – the Common Vulnerabilities and Exposures (CVE) system [120]. Here, we developed a proof-of-concept application in Python programming language that interfaced with Shodan API. This was built to identify the total number of smart connected cameras, including metadata being transmitted from them.

To determine the severity (risk) levels of each identified vulnerability, the NVD[16] was utilized. The NVD is a widely used database containing millions of records about software vulnerabilities. Furthermore, it ranks vulnerabilities using qualitative labels, e.g., “Low”, “Medium”, and “High”. In this dissertation, we used the NVD to grade the identified vulnerabilities of smart connected cameras.

4.5. Data Generation Methods

A data generation method is the means by which empirical data or evidence is produced [109]. Four common examples of data generation

[16] `https://nvd.nist.gov` [accessed December 31, 2020].

methods are interviews, observations, questionnaires, and documents [109].

In this dissertation, we consult multiple document types as a means for generating data. A summary of these document types include:

- **Books.** These were used to gain an initial understanding of smart homes, cyber-physical systems, and as a generic reference connected to the security and privacy aspects of this dissertation.
- **Reports and news articles.** These were used to identify statistics and trends about smart home technologies and products. Also, they were used to identify real-life vulnerabilities and attacks on IoT devices. Specifically, penetration testing reports were leveraged in some of the studies. Some news portals, e.g., The Hacker News[17], were also consulted to get the latest information about vulnerabilities and attacks.
- **Journals, conferences, and workshop proceedings.** These were used to attain updated theories, emerging concepts, and methods used by researchers working on similar domains and research problems.
- **Privacy policies.** These were used to identify smart home data and data collection practices of commercial organizations, including device manufacturers.
- **Product databases and manuals.** These were used to identify technical information about COTS smart connected homes. Particularly, the databases SmartHomeDB and PrivacyNotIncluded, were used. SmartHomeDB is a comprehensive and community-supported database covering the technical specifications of commercial smart home devices. PrivacyNotIncluded is an alternative to SmartHomeDB that includes features related to the sensors utilized by devices, including their accompanying apps. Product manuals were also consulted to understand the technical composition of actual devices, e.g., in terms of their sensors.
- **Vulnerability databases.** These were used to find documentation about vulnerabilities. A key vulnerability database is the NVD. This database contains information about the severity of vulnerabilities, including, often, pointers to available exploit code.

In addition to the mentioned document types, we use simulation as an alternative data generation method. This is used to generate test data especially for demonstrating the usefulness of a proposed contribution. As an example, we use Alloy[18] for generating a random instance of a smart connected home in Paper 10. Alloy is a formal specification and analysis language.

[17] `https://thehackernews.com` [accessed December 31, 2020].
[18] `https://alloytools.org` [accessed December 31, 2020].

To retrieve relevant documents for the literature review various search terms were used. A few examples of these terms that are used across the different papers, along with other search terms, are: "smart home", "connected home", and "home automation". We also use literature snowballing as a technique for finding additional literature. Snowballing is a technique where the search for literature is based on what a particular author has cited, and citations of a research paper. All the utilized sources, including the type of documents cited, are identified in the actual publications (cf. Part II).

In terms of scholarly databases, we used IEEE Xplore, ACM Digital Library, and SpringerLink as the primary databases. Most of the retrieved publications were also indexed and available over Google Scholar. When reviewing existing work connected with a specific research question, we primarily rely on scholarly sources but also refer to industry sources. This is done to render a more comprehensive view of a given domain or problem area.

4.6. Data Analysis Techniques

There are two main methods used for analyzing data: quantitative data analysis and qualitative data analysis [109]. Quantitative data analysis uses mathematical techniques such as statistics to examine and interpret data. Qualitative data analysis looks for themes and categories typically within the words or images people use or create.

In this dissertation, we employ different techniques for analyzing data associated with both quantitative and qualitative data analysis. An overview of the adopted data analysis techniques is as follows.

- **Content analysis.** This is a systematic examination of the contents of a particular material for identifying patterns or themes. We used content analysis across the different studies, particularly, for survey research. As an example, we used content analysis to examine the presence, frequency, and centrality of concepts, often represented as words, e.g., as unigrams or bigrams, in the case of privacy policies (cf. Paper 6).

- **Thematic analysis.** This is a technique that helps identify, analyze, and interpret patterns within data. We used thematic analysis to uncover key smart home functional areas as they emerge from the data, and then to group them into higher-level categories as are used in the taxonomy of smart connected home devices (cf. Paper 5). Also, we used thematic analysis to identify the state-of-the-art challenges and mitigations (cf. Paper 1).

- **Statistical analysis.** This is a tool that helps find patterns and differences in the data and identify relationships between variables. We compute statistics about the occurrence of connected devices in each of the identified functional groups, and for calculating the distribution of technical capabilities across each smart home functional category (cf. Paper 5). Also, we use descriptive statistics to indicate the mean and variance of reviews of smart connected home devices (cf. Paper 6). Another statistical method used is the gap statistic method. This was used to determine the optimal number of clusters when using k-means for partitioning smart connected home systems (cf. Paper 8).

- **Formal methods.** This involves the application of different theoretical computer science fundamentals for the specification, development, and verification of software and hardware systems. We used a n-tuple structure, functions and relations, and Extended Backus–Naur, for describing the smart connected home in a generic form (cf. Paper 10). First-order logic, a type of predicate logic, was used for finding privacy threats in the smart connected home (cf. Paper 9). Attack trees (cf. Paper 10) and DFDs (cf. Paper 7) were used for representing attacks and for modeling data flows in a smart connected home, respectively.

Different software programs were used for conducting data analysis. The software tool, SPSS[19], was used to compute statistics about the occurrence of connected devices in each of the identified functional categories, and for calculating the distribution of technical capabilities in those categories. The programming language, R[20], was used to analyze the privacy policies for their collected data types, and also for applying unsupervised machine learning to identify clusters of smart connected home systems. The formal specification and analysis tool, Alloy, was used to capture the generic specification of the smart connected home and the structural relations between its various components.

The adopted research strategies, data generation methods, and data analysis techniques for each research question are summarized in Table 3.

4.7. Ethical Considerations

Ethical considerations are essential to ensure that no harm is done to

[19] `https://www.ibm.com/products/spss-statistics` [accessed December 31, 2020].
[20] `https://www.r-project.org` [accessed December 31, 2020].

Table 3.: An overview of the research design adopted for answering the research questions posed in this dissertation.

Research question	Research strategies	Data generation methods	Data analysis techniques
RQ1	Survey (Paper: 3,5,6,8), design and creation (Paper: 5,6,8)	Product databases and manuals, privacy policies, literature[21]	Content analysis, statistical analysis, thematic analysis
RQ2	Design and creation (Paper: 7,9)	Literature, simulated data	Formal methods
RQ3	Design and creation (Paper: 2, 10), case study (Paper 4)	Literature, vulnerability databases, reports and news articles, simulated data	Formal methods, statistical analysis, thematic analysis
RQ4	Survey (Paper: 1, 3), design and creation (Paper 7)	Literature	Thematic analysis, formal methods

[18] The value "literature" indicates that scholarly and industry literature were consulted for data generation.

any individual, group, organization, or environment by collecting data and publication of research studies [121]. In this dissertation, we did not conduct research that involved direct interaction with human participants. Nonetheless, there are other aspects of ethics related to digital data collection that we have undertaken responsibly.

When data involved vulnerabilities in specific connected devices or categories of connected devices that could cause potential harm to individuals, groups or organizations, we ensured that we only reported details of vulnerabilities that have been disclosed in public vulnerability databases. This especially concerned the use of Shodan in Paper 4. Moreover, we did not include any information, e.g., IP addresses, that could potentially identify a household or an individual.

As a technology, Shodan only collects metadata of connected devices

that are already broadcasted and available online. Moreover, we did not use Shodan for attack preparation. Attack preparation may be considered, in some countries or states, a criminal offence, for example, in the US by the Computer Fraud and Abuse Act (CFAA). Should we have opted to attack a host, then we would have sought to obtain an informed consent by the organization or individual that are in scope and only collect data in compliance with legal and contractual requirements.

Another dimension that may cause ethical concern revolves around data scraping. Data scraping was used as a technique for extracting capabilities, e.g., sensors of connected devices, from product databases. To proceed with ethical caution, we provided a User Agent string that makes the intentions clear and provided contact information. Also, we requested data at a reasonable rate, as otherwise the scraping process might be interpreted by the recipient, potentially, as a DDoS attack.

Moreover, we remark that some of the presented contributions may be used unethically. In particular, the attack algorithm that is included in Paper 10. While for demonstration purposes, we demonstrate the use of the algorithm to generate attack paths in a smart connected home, in the future, extensions of the algorithm could be developed for unethical reasons. Thus, future researchers and policymakers need to create a code of ethics or statement of ethical practices that proactively safeguard against the risk of harm (physical, physiological, or emotional) to participants.

Finally, we remark that, to the best of our knowledge, all the data collected in this dissertation are cited and reflected upon from our perspective. Additionally, all the datasets, tools, and methods used are discussed in detail in the actual publications.

CHAPTER

CONTRIBUTIONS

Data is the pollution problem of the information age, and protecting privacy is the environmental challenge.

Bruce Schneier,
Data and Goliath, 2015

In this dissertation, four research questions were studied. These research questions resulted in eight main contributions (C1–C8) that help answer them. The first three contributions (C1–C3) correspond to RQ1 and are related to state-of-the-art devices and data. The second set of contributions (C4, C5) correspond to RQ2 and are related to the domain of threat identification and analysis. The third set of contributions (C6, C7) correspond to RQ3 and are related to the domain of risk modeling. The fourth contribution (C8) is related to mitigations and challenges of smart connected homes. The contributions, alongside their associated research questions and papers, are graphically illustrated in Figure 3.

In the next sections, we explain each research contribution in detail.

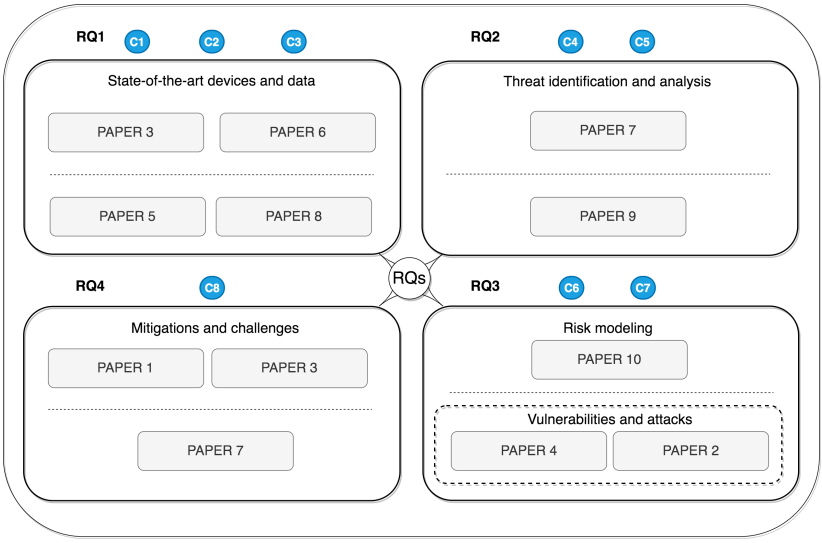

Figure 3.: Overview of the research questions mapped to their contributions and paper numbers.

5.1. State-of-the-Art Devices and Data

In the study of RQ1, we developed three main contributions (C1–C3). These contributions are elaborated below and are detailed in Paper 3, Paper 5, Paper 6, and Paper 8. Paper 3 mainly sets the foundation for the aforementioned papers.

[C1] Taxonomy and analysis of connected devices. We propose a hierarchical taxonomy of smart connected home devices that categorizes devices according to their functionality. The taxonomy is an attempt to homogenize the fragmented and multidisciplinary smart home space by utilizing the actual specifications of 1,193 commercial connected devices.

In relation to this, we also derive from the parsed data a set of 12 capabilities, grouped under 4 categories, that can describe and generically compare smart connected home devices. Three examples of capabilities are related to a device's support for wireless protocols, device's embedded gateway functionality, and device's support for Application Programming Interfaces (API).

Alongside the classification, we analyze the entire spectrum of commercial smart home devices according to their functionality and implemented capabilities. A key finding of the analysis is the dominant support for wireless protocols across the entire surveyed systems and,

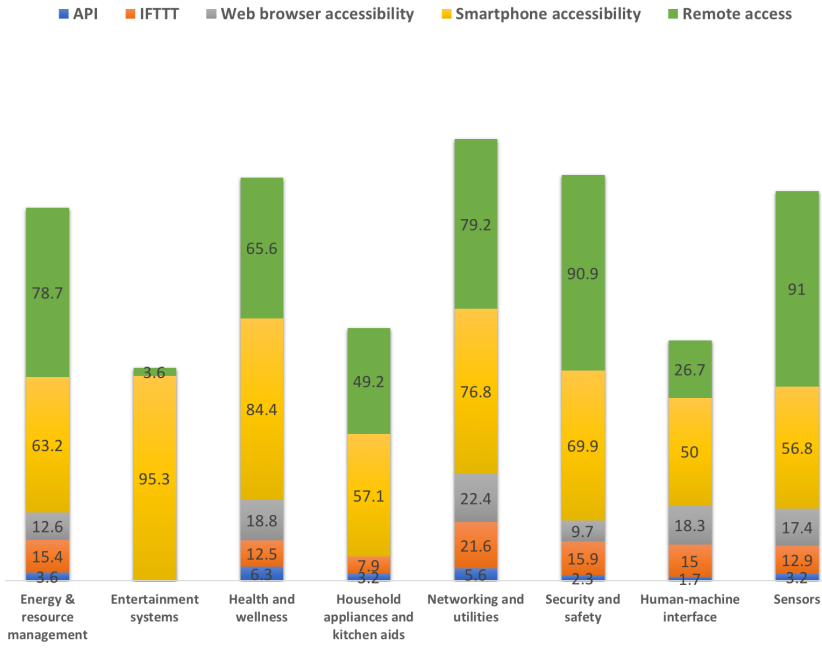

Figure 4.: Stacked bar chart showing the percentage distribution (n=1,193) of supported capabilities (API, IFTTT, web browser accessibility, smartphone accessibility, and remote access) for each smart home category. The chart indicates that overall the least implemented capability is the API, and that the majority of surveyed devices allow for remote access.

on the contrary, the lack of global support for APIs. In Figure 4, we present a graphical overview of some of the capabilities that have been implemented by the surveyed devices. Details about this contribution are found in Paper 5.

[C2] Classification and analysis of connected devices and their apps. We propose a classification of smart connected home systems that is focused on personal data exposure. This classification is built using k-means clustering, taking into account the technical specification of 81 commercial smart connected home systems. Building the classification through unsupervised learning helps reduce possible biases and subjectivity when selecting and labeling data.

The classification uses the data collecting capabilities of connected devices and their accompanying mobile devices, as input for building the classification. Specifically, the sensors that are considered to the most privacy-invasive, namely cameras, audio, and location, are used as features for computing the actual classes.

Alongside the classification, we analyze the distribution of data

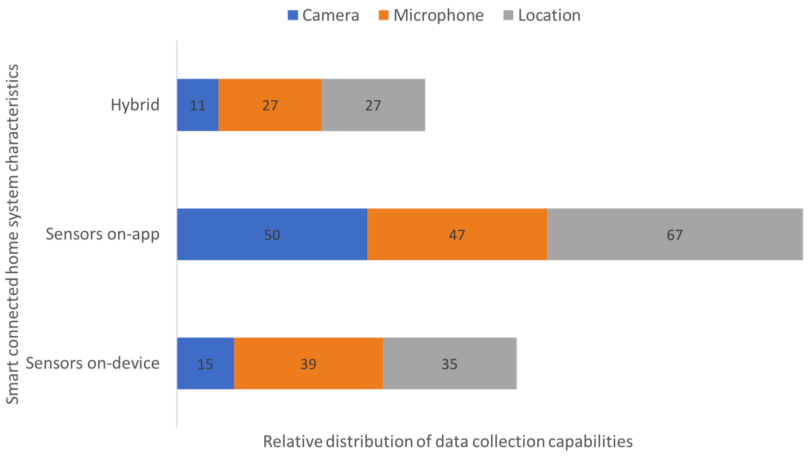

Figure 5.: The availability of different sensor types across the entire distribution of surveyed systems (n=81). Values inside cells represent the number of systems supporting a particular data collecting capability. Hybrid indicates systems that can simultaneously read a data collecting capability from the device and mobile app. The majority of smart connected home systems rely on apps for collecting data, with the location being the most collected attribute.

collection capabilities as supported by the surveyed systems. The results of the analysis is graphically depicted in Figure 5. An outcome of this analysis, is that the majority of sensory data are collected through mobile apps, instead of natively from connected devices. Details about this contribution are found in Paper 8.

[C3] **Analysis and classification of collected data.** We analyze the privacy policies of 87 different types of devices issued from 64 manufacturers of commercial smart home devices. A key finding of the analysis is that all the surveyed smart home devices collect personal and account details from users, with some devices, in particular gateways and a music player, collecting all the potential data types identified from policies.

In the analysis, we identify 10 different smart home data categories, which correspond to data being generated by the user, typically the smart home residents, and data being generated by the connected device. These categories are empirically derived by analyzing privacy policies. Also, we propose a model that groups together smart connected home data. This model categorizes collected data from different functional categories of devices according to their collection mode, collection method, and collection phase. This model is an effort at identifying and categorizing the different data types of an entire smart home system using privacy policies as a medium for doing so. Details about this contribution are found in Paper 6.

5.2. Threat Identification and Analysis

In the study of RQ2, we developed two main contributions (C4, C5). These contributions are elaborated below and are detailed in Paper 7 and Paper 9.

[C4] Privacy-centered system model. We propose a system model that formalizes the description of a smart connected home in a generic manner, capturing the dynamics, including the properties and requirements for modeling privacy in such a context. The model uses the CI theory as an overarching framework for describing information privacy and for identifying the different entities making up the model.

Together with a set of formulas that are based on first-order logic and functions, the model can be used to automatically and systematically identify privacy threats. Additionally, the model can help in recognizing privacy issues early-on when designing smart connected homes, potentially easing the path for achieving compliance with privacy regulations.

Details about this contribution are found in Paper 9, and an expansion of this model for use in privacy risk analysis is found in Paper 10.

[C5] Privacy-centered data lifecycle. We propose a model that captures the generic function of an IoT system to model privacy so that threats affecting such contexts can be identified and categorized at the system design stage. The model extends DFDs with new processes and annotations.

The model can help identify privacy threats indicating during which point of the data lifecycle a particular privacy threat is likely to occur, and knowledge of the corresponding data protection goals that are affected. The model is especially beneficial for smart connected home developers, particularly for data controllers, to better plan in identifying and consequently implementing privacy measures during the early stages of the software development process.

As a side-contribution of the privacy-centered data lifecycle are proposed data lifecycle enhancements represented as complex processes in DFDs. These enhancements can help secure the IoT development against privacy threats. The mitigations are represented as PbD strategies intended for both the data subjects and data controllers. In Figure 6, we depict the proposed DFD extensions.

Details about this contribution and how the proposed extensions can be used are found in Paper 7[22].

[22] In Paper 7, we use the term "smart living space" instead of "smart connected home". The former has arguably a broader scope. However, it includes smart connected homes, as we understand them, as a concept.

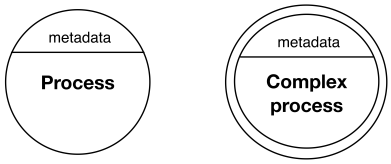

Figure 6.: A DFD extension to the process and complex process symbol. The metadata may represent a purpose statement indicating a reason for collecting and using personal data.

5.3. Risk Modeling

In the study of RQ3, we developed two main contributions (C6, C7). These contributions are elaborated below and are detailed in Paper 2, Paper 4, and Paper 10.

[C6] Threat agent model. We propose a model for the smart connected home that identifies different malicious intruders, risks, and typical compromise methods used by a range of threat agents. This model uses the ICS-CERT as a basis for identifying threat agents, motivations based on the FBI cyber-attack data, and explores both hypothetical and real-life examples of attacks targeting smart connected homes.

Identifying the malicious threat agents, including their motivations and capabilities, gives smart home researchers and developers an alternative approach to reason about risk exposure and a possible foundation for building effective protection strategies for smart connected homes. Details about this contribution can be found in Paper 2.

The threat agent model was instantiated using the hacker as the threat agent in Paper 4 to investigate the global vulnerability state of smart connected cameras in use around the world. A finding of that study is the extent of course-grained data available over the Internet, typically in banner information, that can be used to identify and potentially exploit connected devices with minimal technical skills. Part of the threat agent model was adopted and formalized in Paper 10.

[C7] Framework for modeling and analyzing privacy risks. We propose a framework for dynamically modeling and analyzing privacy risks affecting smart connected homes. The framework uses as input a system model of the smart connected home, a threat model, and a set of metrics that together with an algorithm help quantify the attack success likelihood and impact of privacy attacks affecting smart connected homes.

The framework contributes to automating the process of attack discovery and evaluation of cyberattacks targeting households. Providing

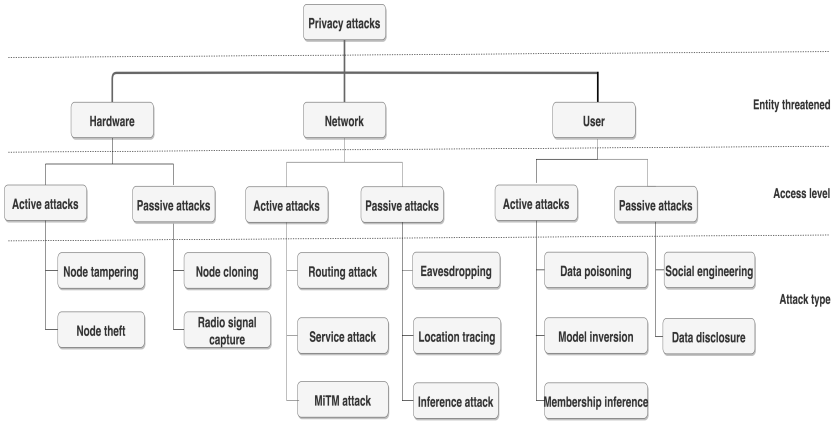

Figure 7.: Taxonomy of privacy attacks in the home structured according to the entity they target. Hardware layer attacks target the physical components; network layer attacks target the communication and connectivity; and user layer attacks target the smart home users.

automatic methods for conducting risk evaluation in the IoT is an important requirement given the dynamic and evolving features of IoT technologies.

As a side-contribution, an attack taxonomy that identifies and categorizes privacy attacks targeting the smart connected home[23] based on the entity they target was proposed. The resulting taxonomy (see Figure 7) is used in the proposed framework as a component for determining the susceptibility to privacy attacks. Details about this contribution are found in Paper 10.

5.4. Mitigations and Challenges

In the study of RQ4, we developed C8 as the main contribution. This contribution is elaborated below and is detailed in Paper 1, Paper 3, and Paper 7.

[C8] Identification of security challenges and their mitigations. We identify the state-of-the-art challenges and mitigations for smart connected homes. Specifically, we identify the challenges, particularly those related to security and privacy, that need priority attention from smart

[23] The term "smart home" instead of "smart connected home" was used in Paper 10. Although, in this case, we regard the terms to be equivalent, the term "smart connected home" was not used in this context following the journal's double-blind author submission guidelines. These guidelines prevent the use of terms that can be used to identify the contributing authors.

home developers and researchers. Other challenges, e.g., that of interoperability, are also identified as core challenges that adversely affect the design of typical smart home solutions. Details about this contribution are found in Paper 3. An analysis of the challenges related to privacy and security is found in Paper 1.

In Paper 1, we also identify the different mitigations working at different architecture layers of the smart connected home. Here, we identify technological contributions from both the industry and academia to mitigate security and privacy risks focusing on the device, network, and services. Some examples of these were introduced in Section 2.3.5, however, we also introduce specific examples focused on smart connected homes, including mentioning some of the actual consumer products.

Additional mitigations that can be applied to increase the security and privacy, specifically during the design phase of smart connected home systems, are found in Paper 7.

CHAPTER

CONCLUSIONS AND FUTURE WORK

A home is much more than a shelter; it is a world in which a person can create a material environment that embodies what he or she considers significant. In this sense the home becomes the most powerful sign of the self of the inhabitant who dwells within.

Mihaly Csikszentmihalyi and Eugene Halton
The Meaning of Things, 1981

The overarching purpose of this dissertation was to investigate how privacy and security has been transformed as homes have evolved into smart Internet-connected homes. In this final chapter, we conclude the dissertation by summarizing the research contributions. Finally, we identify future work based on these contributions.

6.1. Conclusions

The home has often been viewed as the quintessential place of privacy

and a space where one should feel secure from intrusion. In recent years, the home has been transformed into a smart connected home, comprising heterogenous IoT devices aiming to improve the efficiency and quality of life of their users. Nonetheless, connected technologies are increasingly cyber-physical and sensor-rich bringing forth numerous challenges when it comes to protecting privacy and security of the residents. Homes form a unique research domain, as they involve extremely personal and sensitive data, and simultaneously a variety of users with different expertise and expectations of privacy and security.

In this dissertation, we conducted research to understand how the nature of privacy and security has evolved in the context of the home. Specifically, in relation to the smart connected home, we: (i) investigated the state-of-the-art devices and data collected by such devices; (ii) identified and analyzed threats and threat agents; (iii) proposed a method to model and evaluate risks; and (iv) identified challenges and proposed risk mitigations. From a research outlook, the items (i)–(iv), resulted in the development of eight contributions, namely, in the form of design science artefacts, survey contributions, and empirical research contributions. These contributions are intended both for the scientific community but also for industry actors such as smart home developers and policymakers.

As a result of the research, we identify three main findings. First, we observe that most of the surveyed commercial devices are collecting instances of sensitive and personal data but are prone to critical vulnerabilities. As an example, we identified how smart connected cameras are transmitting sensitive data, e.g., device configuration data, to the Internet allowing such devices to be exploited, causing privacy and security threats to the residents. Moreover, by investigating privacy policies of smart home manufacturers, we found that all surveyed devices collect instances of personal data, specifically data associated with personal and account details. Second, we observe a shortage of scientific models that capture the complexity and heterogeneity of real-world smart home deployments. We observe that most of the available models for identifying threats and risks were created before the emergence of the smart connected home, and tend to prioritize security over privacy. Finally, we note that despite the increasing regulations and attention to privacy and security, e.g., through the EU GDPR and recently through the IoT Cybersecurity Improvement Act of 2020, there is still a lack of integrative approaches, including tools, intended to proactively safeguard the privacy and security of the residents. Overall, we contributed to addressing these shortcomings by developing a framework and models that enable early identification of threats, better planning for risk management scenarios, and mitigation of potential impacts caused by attacks before they reach the homes and compromise the lives of the residents.

The presented contributions are not meant to be definitive. They cannot be, because privacy and security, are evolving. In the future, new technologies and new ways of living will create new privacy and security concerns, and transform old ones. Therefore, we need a foundation to help deepen the understanding and reasoning about privacy and security concerns affecting smart connected homes. This dissertation is the beginning of what in the future hopefully will pave the way for improved regulations, leading to more coherent standards, and facilitating the release of more privacy-preserving and secure technologies to consumers, helping conserve the deep significance of the home.

6.2. Future Work

The work presented in this dissertation opens many interesting research directions for the future work. The focus of this dissertation was on rendering an understanding of the smart home with regards to the analysis and modeling of privacy and security risks. The proposed contributions can be further extended and there are several avenues that can be explored further to yield more privacy-preserving and secure smart homes. Some of the directions for future work are related to the following areas.

Embedding privacy and security-enhancing mechanisms directly into connected devices. Most of the smart home devices rely on the gateway or router as their primary protection mechanism. It would be beneficial to investigate how privacy- and security-enhancing mechanisms, such as encryption and data minimization, can be embedded directly into connected devices, including constrained devices. This would help maintain especially confidentiality if a compromised gateway or router exposes data.

Blockchain as a privacy and security-enhancing mechanism. The majority of smart home devices rely on centralized cloud-based servers for rendering their services. This may create a potential threat whereby different entities, e.g., cloud provider personnel, can gain unauthorized access to personal data of residents. It would be beneficial to explore blockchain as a technology to remove the requirement of a central authority, possibly helping increase privacy while also reducing the risk of a single point of security failure.

AI as a mechanism for automatically responding to threats. Homes may be subject to attacks that are difficult to detect using standard technologies such as firewalls and intrusion detection systems. It would be beneficial to explore AI, particularly machine learning

methods, to detect emerging risks targeting smart connected homes. A successful implementation may result in the smart connected home gaining a self-learning immunity that can automatically detect and potentially respond to attacks, including those previously unknown.

Automated and continuous risk assessment approaches. Most of the existing risk assessment approaches are not effective for dealing with the complexity, dynamicity, and automation aspects of IoT technologies. Thus, it would be beneficial to research approaches that can be used to extend the current risk assessment methodologies to support more automated and continuous risk assessment as needed in smart connected homes.

BIBLIOGRAPHY

1. Mark Weiser. The Computer for the 21st Century. *Scientific american*, 265(3):94–105, 1991.
2. Kevin Ashton et al. That 'Internet of Things' Thing. *RFID Journal*, 22(7):97–114, 2009.
3. Andrew Whitmore, Anurag Agarwal, and Li Da Xu. The Internet of Things – a survey of Topics and Trends. *Information Systems Frontiers*, 17(2):261–274, 2015.
4. Manuel Silverio-Fernández, Suresh Renukappa, and Subashini Suresh. What is a Smart Device? – a Conceptualisation within the Paradigm of the Internet of Things. *Visualization in Engineering*, 6(1):3, 2018.
5. Hammad Iqbal, Jamie Ma, Qing Mu, Venkatesh Ramaswamy, Gabby Raymond, Daniel Vivanco, and John Zuena. Augmenting Security of Internet-of-Things using Programmable Network-Centric Approaches: a Position Paper. In *26th International Conference on Computer Communication and Networks*, page 1–6. IEEE, 2017.
6. Mussab Alaa, Aws Alaa Zaidan, Bilal Bahaa Zaidan, Mohammed Talal, and Miss Laiha Mat Kiah. A Review of Smart Home Applications based on Internet of Things. *Journal of Network and Computer Applications*, 97:48–65, 2017.
7. Vincent Ricquebourg, David Menga, David Durand, Bruno Marhic, Laurent Delahoche, and Christophe Loge. The Smart Home Concept: our Immediate Future. In *1st IEEE International Conference on E-learning in Industrial Electronics*, page 23–28. IEEE, 2006.
8. Mu-Yen Chen, Edwin Lughofer, and Ken Sakamura. Information Fusion in Smart Living Technology Innovations. *Information Fusion*, (21):1–2, 2015.
9. Statista Research Department. IoT: Number of Connected Devices Worldwide 2012-2025. `https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/`, 2016. Accessed: December 31, 2020.

10. James Manyika and Michael Chui. By 2025, Internet of Things Applications could have $11 Trillion Impact. https://www.mckinsey.com/mgi/overview/in-the-news/by-2025-internet-of-things-applications-could-have-11-trillion-impact, 2015. Accessed: December 31, 2020.
11. International Data Corporation. The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things. https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm, 2014. Accessed: December 31, 2020.
12. Zion Market Research. Smart Home Market Size & Share will Hit $53.45 Billion by 2022. https://www.globenewswire.com/news-release/2017/04/12/959610/0/en/Smart-Home-Market-Size-Share-will-hit-53-45-Billion-by-2022.html, 2017. Accessed: December 31, 2020.
13. Daniel J Solove. *Understanding Privacy*. Harvard University Press, 2008.
14. W Keith Edwards and Rebecca E Grinter. At Home with Ubiquitous Computing: Seven Challenges. In *International conference on Ubiquitous Computing*, page 256–272. Springer, 2001.
15. Sarah Spiekermann, Jana Korunovska, and Marc Langheinrich. Inside the Organization: Why Privacy and Security Engineering is a Challenge for Engineers. *Proceedings of the IEEE*, 107(3):600–615, 2018.
16. Harald Sundmaeker, Patrick Guillemin, Peter Friess, and Sylvie Woelfflé. Vision and challenges for realising the internet of things. *Cluster of European Research Projects on the Internet of Things, European Commission*, 3(3):34–36, 2010.
17. Imtiaz Ahmad, Rosta Farzan, Apu Kapadia, and Adam J Lee. Tangible Privacy: Towards User-Centric Sensor Designs for Bystander Privacy. *ACM on Human-Computer Interaction*, 4(CSCW2):1–28, 2020.
18. Orlando Arias, Jacob Wurm, Khoa Hoang, and Yier Jin. Privacy and Security in Internet of Things and Wearable Devices. *IEEE Transactions on Multi–Scale Computing Systems*, 1(2):99–109, 2015.
19. Jorge Granjal, Edmundo Monteiro, and Jorge Sá Silva. Security for the Internet of Things: a Survey of Existing Protocols and Open Research issues. *IEEE Communications Surveys & Tutorials*, 17(3):1294–1312, 2015.
20. Sabrina Sicari, Alessandra Rizzardi, Luigi Alfredo Grieco, and Alberto Coen-Porisini. Security, Privacy and Trust in Internet of Things: the Road Ahead. *Computer networks*, 76:146–164, 2015.
21. Constantinos Kolias, Georgios Kambourakis, Angelos Stavrou,

and Jeffrey Voas. DDoS in the IoT: Mirai and other Botnets. *Computer*, 50(7):80–84, 2017.
22. Jingjing Ren, Daniel J Dubois, David Choffnes, Anna Maria Mandalari, Roman Kolcun, and Hamed Haddadi. Information Exposure from Consumer IoT Devices: a Multidimensional, Network-informed Measurement Approach. In *Internet Measurement Conference*, page 267–279. ACM, 2019.
23. Noah Apthorpe, Dillon Reisman, and Nick Feamster. Closing the Blinds: Four Strategies for Protecting Smart Home Privacy from Network Observers. *arXiv preprint arXiv:1705.06809*, 2017.
24. George Danezis, Josep Domingo-Ferrer, Marit Hansen, Jaap–Henk Hoepman, Daniel Le Metayer, Rodica Tirtea, and Stefan Schiffner. Privacy and Data Protection by Design – from Policy to Engineering. *arXiv preprint arXiv:1501.03726*, 2015.
25. UN General Assembly. Universal Declaration of Human Rights. *UN General Assembly*, 302(2), 1948.
26. Nadine Guhr, Oliver Werth, Philip Peter Hermann Blacha, and Michael H Breitner. Privacy Concerns in the Smart Home Context. *SN Applied Sciences*, 2(2):247, 2020.
27. H Jeff Smith, Tamara Dinev, and Heng Xu. Information Privacy Research: an Interdisciplinary Review. *MIS Quarterly*, page 989–1015, 2011.
28. The MITRE Corporation. System of Systems. `https://www.mitre.org/publications/systems-engineering-guide/enterprise-engineering/systems-of-systems`. Accessed: December 31, 2020.
29. Roger Clarke. Introduction to Dataveillance and Information Privacy and Definitions of Terms. `http://www.rogerclarke.com/DV/Intro.html`, 1997. Accessed: December 31, 2020.
30. Dieter Gollman. *Computer Security*. John Wiley & Sons, 2013.
31. Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. Design Science in Information Systems Research. *MIS Q.*, 28(1):75–105, 2004.
32. Marie Chan, Eric Campo, Daniel Estève, and Jean-Yves Fourniols. Smart Homes – Current Features and Future Perspectives. *Maturitas*, 64(2):90–97, 2009.
33. Nicola King. Smart Home – a Definition. *Intertek Research and Testing Center*, page 1–6, 2003.
34. Sam Solaimani, Wally Keijzer-Broers, and Harry Bouwman. What we do – and don't – know about the Smart Home: an analysis of the Smart Home Literature. *Indoor and Built Environment*, 24(3):370–383, 2015.
35. Richard Harper. Inside the Smart Home: Ideas, Possibilities and

Methods. In *Harper R. (eds) Inside the Smart Home*, page 1–13. Springer, 2003.
36. Drew Hendricks. The History of Smart Homes. `https://www.iotevolutionworld.com/m2m/articles/376816-history-smart-homes.htm`, 2014. Accessed: December 31, 2020.
37. Kevin Gotkin. When Computers were Amateur. *IEEE Annals of the History of Computing*, 36(2):4–14, 2014.
38. Janet Abbate. Getting Small: a Short History of the Personal Computer. *Proceedings of the IEEE*, 87(9):1695–1698, 1999.
39. Terence KL Hui, R Simon Sherratt, and Daniel Díaz Sánchez. Major Requirements for Building Smart Homes in Smart Cities based on Internet of Things technologies. *Future Generation Computer Systems*, 76:358–369, 2017.
40. Sharda R Katre and Dinesh V Rojatkar. Home Automation: Past, Present and Future. *International Research Journal of Engineering and Technology*, 4(10):343–346, 2017.
41. Tatsuya Yamazaki. Beyond the Smart Home. In *2006 International Conference on Hybrid Information Technology*, volume 2, page 350–355. IEEE, 2006.
42. Ke Xu, Xiaoliang Wang, Wei Wei, Houbing Song, and Bo Mao. Toward Software Defined Smart Home. *IEEE Communications Magazine*, 54(5):116–122, 2016.
43. Alex Herceg. Defusing the Hype in the Smart Home Space. *Renewable Energy Focus*, 17(3):102–104, 2016.
44. Cory D Kidd, Robert Orr, Gregory D Abowd, Christopher G Atkeson, Irfan A Essa, Blair MacIntyre, Elizabeth Mynatt, Thad E Starner, and Wendy Newstetter. The Aware Home: a Living Laboratory for Ubiquitous Computing Research. In *International Workshop on Cooperative Buildings*, page 191–198. Springer, 1999.
45. Diane J Cook, Michael Youngblood, Edwin O Heierman, Karthik Gopalratnam, Sira Rao, Andrey Litvin, and Farhan Khawaja. MavHome: An Agent-based Smart Home. In *1st IEEE International Conference on Pervasive Computing and Communications*, page 521–524. IEEE, 2003.
46. Sumi Helal, William Mann, Hicham El-Zabadani, Jeffrey King, Youssef Kaddoura, and Erwin Jansen. The Gator Tech Smart House: a Programmable Pervasive Space. *Computer*, 38(3):50–60, 2005.
47. Stephen S Intille. Designing a Home of the Future. *IEEE Pervasive Computing*, 1(2):76–82, 2002.
48. Stephen S Intille, Kent Larson, Emmanuel Munguia Tapia, Jennifer S Beaudin, Pallavi Kaushik, Jason Nawyn, and Randy Rockinson. Using a Live-in Laboratory for Ubiquitous Computing Research. In *International Conference on Pervasive Comput-*

ing, page 349–365. Springer, 2006.
49. Gregory Mone. Intelligent Living. *Commun. ACM*, 57(12):15–16, 2014.
50. Tiago DP Mendes, Radu Godina, Eduardo MG Rodrigues, João CO Matias, and João PS Catalão. Smart Home Communication Technologies and Applications: Wireless Protocol Assessment for Home Area Network Resources. *Energies*, 8(7):7279–7311, 2015.
51. Andrew Guthrie Ferguson. The Smart Fourth Amendment. *Cornell L. Rev.*, 102:547, 2016.
52. Cédric Lévy-Bencheton, Eleni Darra, Guillaume Tétu, Guillaume Dufay, and Mouhannad Alattar. Security and Resilience of Smart Home Environments Good Practices and Recommendations. *The European Union Agency for Network and Information Security (ENISA)*, 2015.
53. Carsten Bormann, Mehmet Ersue, and Ari Keranen. Terminology for Constrained–Node Networks. *Internet Engineering Task Force*, 2014.
54. Meredydd Williams, Jason RC Nurse, and Sadie Creese. The Perfect Storm: The Privacy Paradox and the Internet-of-Things. In *11th International Conference on Availability, Reliability and Security*, page 644–652. IEEE, 2016.
55. Judith DeCew. Privacy. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2018 edition, 2018.
56. Samuel D Warren and Louis D Brandeis. The Right to Privacy. *Harvard Law Review*, page 193–220, 1890.
57. Alan F Westin. Privacy and Freedom. *Washington and Lee Law Review*, 25(1):166, 1968.
58. Irwin Altman. *The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding*. Brooks-Cole, 1975.
59. Helen Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press, 2009.
60. Secretary's Advisory Committee on Automated Personal Data Systems (SACAPDS). Records, Computers, and the Rights of Citizens: Report. US Department of Health, Education & Welfare, 1973.
61. Paul M Schwartz and Daniel J Solove. The PII problem: Privacy and a New Concept of Personally Identifiable Information. *NYUL rev.*, 86:1814, 2011.
62. European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016. Technical report, 2016.
63. European Commission. Proposal for a Regulation on Privacy and

Electronic Communications (ePrivacy Regulation) (2017).
64. California Senate Judiciary Committee et al. California Consumer Privacy Act: Ab 375 legislative history. 2018.
65. Robin L. Kelly. H.R.1668–116th Congress (2019–2020): IoT Cybersecurity Improvement Act of 2020, 2020.
66. James Michael Stewart, Mike Chapple, and Darril Gibson. *CISSP: Certified Information Systems Security Professional Official Study*. John Wiley & Sons, 2015.
67. Eric Kenneally. Privacy and Security. *IEEE Internet of Things Magazine*, 1(1):8–10, 2018.
68. Jan Henrik Ziegeldorf, Oscar Garcia Morchon, and Klaus Wehrle. Privacy in the Internet of Things: Threats and Challenges. *Security and Communication Networks*, 7(12):2728–2742, 2014.
69. Mina Deng, Kim Wuyts, Riccardo Scandariato, Bart Preneel, and Wouter Joosen. A Privacy Threat Analysis Framework: Supporting the Elicitation and Fulfillment of Privacy Requirements. *Requirements Engineering*, 16(1):3–32, 2011.
70. Jesus Luna, Neeraj Suri, and Ioannis Krontiris. Privacy-by-Design based on Quantitative Threat Modeling. In *7th International Conference on Risks and Security of Internet and Systems*, page 1–8. IEEE, 2012.
71. Abdullahi Arabo, Ian Brown, and Fadi El-Moussa. Privacy in the Age of Mobility and Smart Devices in Smart Homes. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing*, page 819–826. IEEE, 2012.
72. Andreas Jacobsson, Martin Boldt, and Bengt Carlsson. A Risk Analysis of a Smart Home Automation System. *Future Generation Computer Systems*, 56:719–733, 2016.
73. Noah Apthorpe, Dillon Reisman, Srikanth Sundaresan, Arvind Narayanan, and Nick Feamster. Spying on the Smart Home: Privacy Attacks and Defenses on Encrypted IoT Traffic. *arXiv preprint arXiv:1708.05044*, 2017.
74. Serena Zheng, Noah Apthorpe, Marshini Chetty, and Nick Feamster. User Perceptions of Smart Home IoT Privacy. *ACM Human-Computer Interaction*, 2(CSCW), 2018.
75. Eric Zeng, Shrirang Mare, and Franziska Roesner. End user security and privacy concerns with smart homes. In *13th Symposium on Usable Privacy and Security*, page 65–80, 2017.
76. Verena Zimmermann, Merve Bennighof, Miriam Edel, Oliver Hofmann, Judith Jung, and Melina von Wick. 'Home, Smart Home' – Exploring End Users' Mental Models of Smart Homes. In *Mensch Computer Workshopband*. Gesellschaft für Informatik eV, 2018.

77. Mike Moody and Aaron Hunter. Exploiting Known Vulnerabilities of a Smart Thermostat. In *14th Annual Conference on Privacy, Security and Trust*, page 50–53. IEEE, 2016.
78. Nathan Malkin, Julia Bernd, Maritza Johnson, and Serge Egelman. "What Can't Data Be Used For?" Privacy Expectations about Smart TVs in the US. In *European Workshop on Usable Security*, 2018.
79. Yue Huang, Borke Obada-Obieh, and Konstantin Beznosov. Amazon vs. My Brother: How Users of Shared Smart Speakers Perceive and Cope with Privacy Risks. In *CHI Conference on Human Factors in Computing Systems*, page 1–13, 2020.
80. Simon Moncrieff, Svetha Venkatesh, and Geoff West. Dynamic Privacy in a Smart House Environment. In *2007 IEEE international conference on Multimedia and Expo*, page 2034–2037. IEEE, 2007.
81. Jason RC Nurse, Ahmad Atamli, and Andrew Martin. Towards a Usable Framework for Modelling Security and Privacy Risks in the Smart Home. In *International Conference on Human Aspects of Information Security, Privacy, and Trust*, page 255–267. Springer, 2016.
82. Simone Fischer-Hübner and Stefan Berthold. Privacy-Enhancing Technologies. In *Computer and Information Security Handbook*, page 759–778. Elsevier, 2017.
83. Marit Hansen. Marrying Transparency Tools with User-Controlled Identity Management. In *IFIP International Summer School on the Future of Identity in the Information Society*, page 199–220. Springer, 2007.
84. Ann Cavoukian. Privacy by Design: The 7 Foundational Principles. Implementation and Mapping of Fair Information Practices. *Information and Privacy Commissioner of Ontario, Canada*, 5, 2009.
85. Marc Langheinrich. Privacy by Design – Principles of Privacy-Aware Ubiquitous Systems. In *International Conference on Ubiquitous Computing*, page 273–291. Springer, 2001.
86. Herman T Tavani. *Ethics and Technology*. Wiley, 2013.
87. Jerome H Saltzer and Michael D Schroeder. The Protection of Information in Computer Systems. *Proceedings of the IEEE*, 63(9):1278–1308, 1975.
88. Bishop Matt. *Computer security: art and science*. Addison-Wesley, 2002.
89. Arsalan Mosenia and Niraj K Jha. A Comprehensive Study of Security of Internet–of–Things. *IEEE Transactions on Emerging Topics in Computing*, 5(4):586–602, 2016.
90. William Stallings and Lawrie Brown. *Computer Security: Prin-*

ciples and Practice. Pearson Education, 2015.

91. A Reference Model of Information Assurance & Security, author=Cherdantseva, Yulia and Hilton, Jeremy, booktitle=*2013 international conference on availability, reliability and security*, pages=546–555, year=2013, organization=IEEE.
92. Yassine Maleh, Mohammad Shojafar, Ashraf Darwish, and Abdelkrim Haqiq. *Cybersecurity and Privacy in Cyber Physical Systems*. CRC Press, 2019.
93. Ahmad Amini, Norziana Jamil, Abdul Rahim Ahmad, and Muhammad Reza Zaba. Threat Modeling Approaches for Securing Cloud Computing. *Journal of Applied Sciences*, 15(7):953–967, 2015.
94. Goncalo Martins, Sajal Bhatia, Xenofon Koutsoukos, Keith Stouffer, CheeYee Tang, and Richard Candell. Towards a Systematic Threat Modeling Approach for Cyber-Physical Systems. In *2015 Resilience Week*, page 1–6. IEEE, 2015.
95. Adam Shostack. *Threat modeling: Designing for security*. John Wiley & Sons, 2014.
96. Matthew Rosenquist. Prioritizing Information Security Risks with Threat Agent Risk Assessment. *Intel Corporation White Paper*, 2009.
97. Douglas Gray, Brian Wisniewski, Julia Allen, Constantine Cois, Anne Connell, Erik Ebel, William Gulley, Michael Riley, Robert Stoddard, and Marie Vaughn. *Improving Federal Cybersecurity Governance Through Data-Driven Decision Making and Execution*. Technical Report CMU/SEI-2015-TR-011, Software Engineering Institute, Carnegie Mellon University, 2015.
98. Yonglei Tao and Chenho Kung. Formal Definition and Verification of Data Flow Diagrams. *Journal of Systems and Software*, 16(1):29–36, 1991.
99. Fulvio Corno and Muhammad Sanaullah. Design-Time Formal Verification for Smart Environments: an Exploratory Perspective. *Journal of Ambient Intelligence and Humanized Computing*, 5(4):581–599, 2014.
100. OWASP. OWASP IoT Top 10 Vulnerabilities. `https://www.owasp.org/images/1/1c/OWASP-IoT-Top-10-2018-final.pdf`, 2018. Accessed: December 31, 2020.
101. Bela Genge and Calin Enachescu. ShoVAT: Shodan-based Vulnerability Assessment Tool for Internet–Facing Services. *Security and communication networks*, 9(15):2696–2714, 2016.
102. Bruce Schneier. Attack trees. *Dr. Dobb's journal*, 24(12):21–29, 1999.
103. Naomi Lefkovitz and Kaitlin Boeckl. *NIST Privacy Framework: An Overview*. National Institute of Standards and Technology,

2020.
104. Rodrigo Roman, Pablo Najera, and Javier Lopez. Securing the Internet of Things. *Computer*, 44(9):51–58, 2011.
105. Tamara Denning, Tadayoshi Kohno, and Henry M Levy. Computer Security and the Modern Home. *Communications of the ACM*, 56(1):94–103, 2013.
106. Daniel Bastos, Mark Shackleton, and Fadiali El-Moussa. Internet of Things: a Survey of Technologies and Security Risks in Smart Home and City Environments. In *Living in the Internet of Things: Cybersecurity of the IoT – 2018*. IET, 2018.
107. Jason RC Nurse, Sadie Creese, and David De Roure. Security Risk Assessment in Internet of Things Systems. *IT Professional*, 19(5):20–26, 2017.
108. Mujahid Mohsin, Muhammad Usama Sardar, Osman Hasan, and Zahid Anwar. IoTRiskAnalyzer: a Probabilistic Model Checking based Framework for Formal Risk Analytics of the Internet of Things. *IEEE Access*, 5:5494–5505, 2017.
109. Briony J Oates. *Researching Information Systems and Computing*. Sage, 2005.
110. John W Creswell and J David Creswell. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage publications, 2017.
111. Jan Recker. *Scientific Research in Information Systems: a Beginner's Guide*. Springer Science & Business Media, 2012.
112. Heather R Hall and Linda A Roussel. *Evidence-based practice: an Integrative Approach to Research, Administration, and Practice*. Jones & Bartlett Learning, 2017.
113. Philipp Mayring. *Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution*. Klagenfurt, 2014.
114. Robert B Johnson, Anthony J Onwuegbuzie, and Lisa A Turner. Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research*, (2):112–133, 2007.
115. David Haynes. *Metadata for Information Management and Retrieval: Understanding Metadata and its Use*. Facet, 2018.
116. Valerie J Easton and John H McColl. Statistics glossary v1.1. `http://www.stats.gla.ac.uk/steps/glossary/`, 1997. Accessed: December 31, 2020.
117. Salvatore T March and Gerald F Smith. Design and Natural Science Research on Information Technology. *Decision support systems*, 15(4):251–266, 1995.
118. Robert K Yin. *Case Study Research and Applications: Design and Methods*. Sage publications, 2017.

119. Christian Debes, Andreas Merentitis, Sergey Sukhanov, Maria Niessen, Nikolaos Frangiadakis, and Alexander Bauer. Monitoring Activities of Daily Living in Smart Homes: Understanding Human Behavior. *IEEE Signal Processing Magazine*, 33(2):81–94, 2016.
120. Dorottya Papp, Zhendong Ma, and Levente Buttyan. Embedded Systems Security: Threats, Vulnerabilities, and Attack Taxonomy. In *13th Annual Conference on Privacy, Security and Trust*, page 145–152. IEEE, 2015.
121. Erik Hofstee. Constructing a Good Dissertation. *Sandton: EPE*, 2006.

Bibliography

[1] Omar Alrawi, Chaz Lever, Manos Antonakakis, and Fabian Monrose. Sok: Security evaluation of home-based iot deployments. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 1362–1380. IEEE, 2019.

[2] Orlando Arias, Kelvin Ly, and Yier Jin. Security and privacy in iot era. In *Smart Sensors at the IoT Frontier*, pages 351–378. Springer, 2017.

[3] Florian Arnold, Wolter Pieters, and Mariëlle Stoelinga. Quantitative penetration testing with item response theory. In *2013 9th International Conference on Information Assurance and Security (IAS)*, pages 49–54. IEEE, 2013.

[4] Junaid Arshad, Muhammad Ajmal Azad, Roohi Amad, Khaled Salah, Mamoun Alazab, and Razi Iqbal. A review of performance, energy and privacy of intrusion detection systems for iot. *Electronics*, 9(4):629, 2020.

[5] Waqar Asif, Indranil Ghosh Ray, and Muttukrishnan Rajarajan. An attack tree based risk evaluation approach for the internet of things. In *Proceedings of the 8th International Conference on the Internet of Things*, pages 1–8, 2018.

[6] Audit Analytics. Trends in cybersecurity breach disclosures, 2020.

[7] D Barnard-Wills et al. Enisa threat landscape and good practice guide for smart home and converged media. *ENISA (The European Network and Information Security Agency)*, 2014.

[8] Meriem Bettayeb, Omnia Abu Waraga, Manar Abu Talib, Qassim Nasir, and Omar Einea. Iot testbed security: Smart socket and smart thermostat. In *2019 IEEE Conference on Application, Information and Network Security (AINS)*,

pages 18–23. IEEE, 2019.

[9] Ben Buchanan. A national security research agenda for cybersecurity and artificial intelligence. 2020.

[10] Joseph Bugeja, Paul Davidsson, and Andreas Jacobsson. Functional classification and quantitative analysis of smart connected home devices. In *2018 Global Internet of Things Summit (GIoTS)*, pages 1–6. IEEE, 2018.

[11] Joseph Bugeja, Andreas Jacobsson, and Paul Davidsson. An analysis of malicious threat agents for the smart connected home. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 557–562. IEEE, 2017.

[12] Joseph Bugeja, Désirée Jönsson, and Andreas Jacobsson. An investigation of vulnerabilities in smart connected cameras. In *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 537–542. IEEE, 2018.

[13] Yunus Çadirci. Callstranger cve-2020-12695, 2020.

[14] Commission Nationale de l'informatique et des Liberte (CNIL). Privacy impact assessment (pia) methodology—how to carry out a pia. 2015.

[15] Mina Deng, Kim Wuyts, Riccardo Scandariato, Bart Preneel, and Wouter Joosen. A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. *Requirements Engineering*, 16(1):3–32, 2011.

[16] Tamara Denning, Tadayoshi Kohno, and Henry M Levy. Computer security and the modern home. *Communications of the ACM*, 56(1):94–103, 2013.

[17] Department of Health Care Services. List of hipaa identifiers, 2019.

[18] Nitesh Dhanjani. *Abusing the internet of things: blackouts, freakouts, and stakeouts*. 'O'Reilly Media, Inc.', 2015.

[19] Manoj R Dhobale, Rekha Y Biradar, Raju R Pawar, and Sharad A Awatade. Smart home security system using iot, face recognition and raspberry pi. *International Journal of Computer Applications*, 176:1–6, 2020.

[20] E Hacking News. Fake applications are replicating 'tracetogether,' a singapore covid-19 contact tracing application, 2020.

[21] European Commission. Guidelines on data protection impact assessment (dpia) (wp248rev.01) - european commission, 2017.

[22] FIRST. Cvss v3.1 user guide, 2019.

[23] Unabhängiges Landeszentrum für Datenschutz. The standard data protection model: A concept for inspection and consultation on the basis of unified protection goals, 2017.

[24] Eoghan Furey and Juanita Blue. She knows too much–voice command devices and privacy. In *2018 29th Irish Signals and Systems Conference (ISSC)*, pages 1–6. IEEE, 2018.

[25] Mengmeng Ge, Jin B Hong, Walter Guttmann, and Dong Seong Kim. A framework for automating security analysis of the internet of things. *Journal of Network and Computer Applications*, 83:12–27, 2017.

[26] Dimitris Geneiatakis, Ioannis Kounelis, Ricardo Neisse, Igor Nai-Fovino, Gary Steri, and Gianmarco Baldini. Security and privacy issues for an iot based smart home. In *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pages 1292–1297. IEEE, 2017.

[27] Douglas Gray, Julia Allen, Constantine Cois, Anne Connell, Erik Ebel, William Gulley, Michael Riley, Robert Stoddard, Marie Vaughan, and Brian D Wisniewski. Improving federal cybersecurity governance through data-driven decision making and execution. Technical report, Carnegie-mellon Univ Pittsburgh Pa Pittsburgh United States, 2015.

[28] Foad Hamidi, Kellie Poneres, Aaron Massey, and Amy Hurst. Who should have access to my pointing data? privacy tradeoffs of adaptive assistive technologies. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 203–216, 2018.

[29] Yue Huang, Borke Obada-Obieh, and Konstantin Beznosov. Amazon vs. my brother: How users of shared smart speakers perceive and cope with privacy risks. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[30] Daniel Jackson. *Software Abstractions: logic, language, and analysis*. MIT press, 2012.

[31] Andreas Jacobsson, Martin Boldt, and Bengt Carlsson. A

risk analysis of a smart home automation system. *Future Generation Computer Systems*, 56:719–733, 2016.

[32] Anshul Jain, Tanya Singh, and Satyendra K Sharma. Threats paradigmin iot ecosystem. In *2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*, pages 1–7. IEEE, 2018.

[33] Christos Kalloniatis, Evangelia Kavakli, and Stefanos Gritzalis. Addressing privacy requirements in system design: the pris method. *Requirements Engineering*, 13(3):241–255, 2008.

[34] Kamalanathan Kandasamy, Sethuraman Srinivas, Krishnashree Achuthan, and Venkat P Rangan. Iot cyber risk: a holistic analysis of cyber risk assessment frameworks, risk vectors, and risk ranking process. *EURASIP Journal on Information Security*, 2020:1–18, 2020.

[35] Richard A. Kemmerer and Phillip A. Porras. Covert flow trees: A visual approach to analyzing covert storage channels. *IEEE Transactions on Software Engineering*, 17(11):1166, 1991.

[36] E. Kenneally. Privacy and security. *IEEE Internet of Things Magazine*, 1(1):8–10, 2018.

[37] Tom Kirkham, Django Armstrong, Karim Djemame, and Ming Jiang. Risk driven smart home resource management using cloud services. *Future Generation Computer Systems*, 38:13–22, 2014.

[38] Aleksandr Lenin, Jan Willemson, and Dyan Permata Sari. Attacker profiling in quantitative security assessment based on attack trees. In *Nordic Conference on Secure IT Systems*, pages 199–212. Springer, 2014.

[39] Chao Li and Balaji Palanisamy. Privacy in internet of things: from principles to technologies. *IEEE Internet of Things Journal*, 6(1):488–505, 2018.

[40] Zhen Ling, Kaizheng Liu, Yiling Xu, Chao Gao, Yier Jin, Cliff Zou, Xinwen Fu, and Wei Zhao. Iot security: An end-to-end view and case study. *arXiv preprint arXiv:1805.05853*, 2018.

[41] Zhen Ling, Junzhou Luo, Yiling Xu, Chao Gao, Kui Wu, and Xinwen Fu. Security vulnerabilities of internet of things:

A case study of the smart plug system. *IEEE Internet of Things Journal*, 4(6):1899–1909, 2017.

[42] Javier Lopez, Ruben Rios, Feng Bao, and Guilin Wang. Evolving privacy: From sensors to the internet of things. *Future Generation Computer Systems*, 75:46–57, 2017.

[43] Jesus Luna, Neeraj Suri, and Ioannis Krontiris. Privacy-by-design based on quantitative threat modeling. In *2012 7th International Conference on Risks and Security of Internet and Systems (CRiSIS)*, pages 1–8. IEEE, 2012.

[44] Astor Maggie. Your roomba may be mapping your home, collecting data that could be shared, 2017.

[45] Alessandro Mantelero. Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection. *Computer law & security review*, 32(2):238–255, 2016.

[46] Kirsten Martin and Helen Nissenbaum. Measuring privacy: an empirical test using context to expose confounding variables. *Colum. Sci. & Tech. L. Rev.*, 18:176, 2016.

[47] Sergio Mascetti, Nadia Metoui, Andrea Lanzi, and Claudio Bettini. Epic: a methodology for evaluating privacy violation risk in cybersecurity systems. *Transactions on Data Privacy*, 11(3):239–277, 2018.

[48] Mujahid Mohsin, Muhammad Usama Sardar, Osman Hasan, and Zahid Anwar. IoTRiskAnalyzer: a probabilistic model checking based framework for formal risk analytics of the Internet of Things. *IEEE Access*, 5:5494–5505, 2017.

[49] Arsalan Mosenia and Niraj K Jha. A comprehensive study of security of internet-of-things. *IEEE Transactions on Emerging Topics in Computing*, 5(4):586–602, 2016.

[50] N. Apthorpe et al. Discovering smart home internet of things privacy norms using contextual integrity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(2):59, 2018.

[51] Ben Nassi, Yaron Pirutin, Adi Shamir, Yuval Elovici, and Boris Zadov. Lamphone: Real-time passive sound recovery from light bulb vibrations. 2020.

[52] National Institute of Standards and Technology. Nist privacy framework: A tool for improving privacy through enterprise risk management, version 1.0. 2020.

[53] Jan-Peter Nicklas, Michel Mamrot, Petra Winzer, Daniel Lichte, Stefan Marchlewitz, and Kai-Dietrich Wolf. Use case based approach for an integrated consideration of safety and security aspects for smart home applications. In *2016 11th System of Systems Engineering Conference (SoSE)*, pages 1–6. IEEE, 2016.

[54] Helen Nissenbaum. Privacy as contextual integrity. *Wash. L. Rev.*, 79:119, 2004.

[55] NIST. National vulnerability database, 2020.

[56] Sukhvir Notra, Muhammad Siddiqi, Hassan Habibi Gharakheili, Vijay Sivaraman, and Roksana Boreli. An experimental study of security and privacy risks with emerging household appliances. In *2014 IEEE conference on communications and network security*, pages 79–84. IEEE, 2014.

[57] Jason RC Nurse, Ahmad Atamli, and Andrew Martin. Towards a usable framework for modelling security and privacy risks in the smart home. In *International Conference on Human Aspects of Information Security, Privacy, and Trust*, pages 255–267. Springer, 2016.

[58] Jason RC Nurse, Sadie Creese, and David De Roure. Security risk assessment in internet of things systems. *IT professional*, 19(5):20–26, 2017.

[59] OWASP. Top internet of things vulnerabilities, 2018.

[60] Nisha Panwar, Shantanu Sharma, Sharad Mehrotra, Łukasz Krzywiecki, and Nalini Venkatasubramanian. Smart home survey on security and privacy. *arXiv preprint arXiv:1904.05476*, 2019.

[61] Nicolas Papernot, Patrick McDaniel, Arunesh Sinha, and Michael P Wellman. Sok: Security and privacy in machine learning. In *2018 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 399–414. IEEE, 2018.

[62] Mookyu Park, Haengrok Oh, and Kyungho Lee. Security risk measurement for information leakage in iot-based smart homes from a situational awareness perspective. *Sensors*, 19(9):2148, 2019.

[63] Wolter Pieters, Dina Hadziosmanovic, Aleksandr Lenin, Lorena Montoya, and Jan Willemson. Trespass: Plug-and-play attacker profiles for security risk analysis (poster). In *35th IEEE Symposium on Security and Privacy 2014*. IEEE

Computer Society, 2014.

[64] Ismini Psychoula, Liming Chen, and Feng Chen. Privacy modelling and management for assisted living within smart homes. In *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pages 1–6. IEEE, 2017.

[65] Youyang Qu, Shui Yu, Wanlei Zhou, Sancheng Peng, Guojun Wang, and Ke Xiao. Privacy of things: Emerging challenges and opportunities in wireless internet of things. *IEEE Wireless Communications*, 25(6):91–97, 2018.

[66] Jingjing Ren, Daniel J Dubois, David Choffnes, Anna Maria Mandalari, Roman Kolcun, and Hamed Haddadi. Information exposure from consumer iot devices: A multidimensional, network-informed measurement approach. In *Proceedings of the Internet Measurement Conference*, pages 267–279, 2019.

[67] ResearchGate. What is the difference between a framework and a model in educational research?, 2015.

[68] Marco Rocchetto and Nils Ole Tippenhauer. Cpdy: extending the dolev-yao attacker with physical-layer interactions. In *International Conference on Formal Engineering Methods*, pages 175–192. Springer, 2016.

[69] Matthew Rosenquist. Prioritizing information security risks with threat agent risk assessment. *Intel Corporation White Paper*, 2009.

[70] Samsung. Samsung ballie at ces 2020 - samsung us newsroom, 2020.

[71] Reza Shokri, Marco Stronati, Congzheng Song, and Vitaly Shmatikov. Membership inference attacks against machine learning models. In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 3–18. IEEE, 2017.

[72] Adam Shostack. *Threat modeling: Designing for security.* John Wiley & Sons, 2014.

[73] Sabrina Sicari, Alessandra Rizzardi, Daniele Miorandi, and Alberto Coen-Porisini. A risk assessment methodology for the internet of things. *Computer Communications*, 129:67–79, 2018.

[74] Vijay Sivaraman, Dominic Chan, Dylan Earl, and Roksana Boreli. Smart-phones attacking smart-homes. In *Proceed-*

ings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks, pages 195–200, 2016.

[75] W Snyder and F Swiderski. Threat modeling. *Microsoft Press*, 35:36–37, 2004.

[76] Daniel J Solove. Understanding privacy. 2008.

[77] Sarah Spiekermann and Lorrie Faith Cranor. Engineering privacy. *IEEE Transactions on software engineering*, 35(1):67–82, 2008.

[78] William Stallings, Lawrie Brown, Michael D Bauer, and Arup Kumar Bhattacharjee. *Computer security: principles and practice*. Pearson Education Upper Saddle River, NJ, USA, 2012.

[79] Statista Research Department. Internet of things (iot) connected devices installed base worldwide from 2015 to 2025, 2016.

[80] J. Sturgess, J. R. C. Nurse, and J. Zhao. A capability-oriented approach to assessing privacy risk in smart home ecosystems. In *Living in the Internet of Things: Cybersecurity of the IoT - 2018*, pages 1–8, 2018.

[81] Frank Swiderski. *Threat modeling*. Microsoft Press, 2004.

[82] The MITRE Corporation. Cve - cve-2020-12695, 2020.

[83] R. Thorburn, A. Margheri, and F. Paci. Towards an integrated privacy protection framework for iot: Contextualising regulatory requirements with industry best practices. In *Living in the Internet of Things (IoT 2019)*, pages 1–6, 2019.

[84] Eran Toch, Claudio Bettini, Erez Shmueli, Laura Radaelli, Andrea Lanzi, Daniele Riboni, and Bruno Lepri. The privacy implications of cyber security systems: A technological survey. *ACM Computing Surveys (CSUR)*, 51(2):1–27, 2018.

[85] Gurkan Tuna, Dimitrios G Kogias, V Cagri Gungor, Cengiz Gezer, Erhan Taşkın, and Erman Ayday. A survey on information security threats and solutions for machine to machine (m2m) communications. *Journal of Parallel and Distributed Computing*, 109:142–154, 2017.

[86] Wikipedia. Extended backus–naur form, 2020. http://tiny.cc/rj11tz.

[87] Mingmin Zhao, Tianhong Li, Mohammad Abu Alsheikh, Yonglong Tian, Hang Zhao, Antonio Torralba, and Dina

Katabi. Through-wall human pose estimation using radio signals. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7356–7365, 2018.

[88] Jan Henrik Ziegeldorf, Oscar Garcia Morchon, and Klaus Wehrle. Privacy in the internet of things: threats and challenges. *Security and Communication Networks*, 7(12):2728–2742, 2014.

[89] Zion Market Research. Global smart home market worth usd 53.45 billion by 2022, 2020.

ISBN 978-91-7877-163-9 (print)
ISBN 978-91-7877-164-6 (pdf)
DOI 10.24834/978917877164-6

MALMÖ UNIVERSITY
205 06 MALMÖ, SWEDEN
WWW.MAU.SE