UNIVERSITÄT
DES
SAARLANDES

On Privacy in Home Automation Systems

Dissertation zur Erlangung des Grades des Doktors der Naturwissenschaften (Dr. rer. nat.) der Fakultät für Mathematik und Informatik der Universität des Saarlandes

von Frederik Möllers

Saarbrücken, 2021

PhD Exam (Colloquium) Details

Dean: Prof. Dr. Thomas Schuster

Date of exam: Monday, 12 July 2021

Chair: Prof. Dr.-Ing. Thorsten Herfet

First Supervisor: Prof. Dr.-Ing. Christoph Sorge

Second Supervisor: Prof. Dr. Christian Rossow

Committee Member: Dr. Francis Dolière Somé

Abstract (English)

Home Automation Systems (HASs) are becoming increasingly popular in newly built as well as existing properties. While offering increased living comfort, resource saving features and other commodities, most current commercial systems do not protect sufficiently against passive attacks. In this thesis we investigate privacy aspects of Home Automation Systems. We analyse the threats of eavesdropping and traffic analysis attacks, demonstrating the risks of virtually undetectable privacy violations. By taking aspects of criminal and data protection law into account, we give an interdisciplinary overview of privacy risks and challenges in the context of HASs. We present the first framework to formally model privacy guarantees of Home Automation Systems and apply it to two different dummy traffic generation schemes. In a qualitative and quantitative study of these two algorithms, we show how provable privacy protection can be achieved and how privacy and energy efficiency are interdependent. This allows manufacturers to design and build secure Home Automation Systems which protect the users' privacy and which can be arbitrarily tuned to strike a compromise between privacy protection and energy efficiency.

Abstract (Deutsch)

Hausautomationssysteme (HAS) gewinnen sowohl im Bereich der Neubauten als auch bei Bestandsimmobilien stetig an Beliebtheit. Während sie den Wohnkomfort erhöhen, Einsparpotential für Strom und Wasser sowie weitere Vorzüge bieten, schützen aktuelle Systeme nicht ausreichend vor passiven Angriffen. In dieser Arbeit untersuchen wir Aspekte des Datenschutzes von Hausautomationssystemen. Wir betrachten die Gefahr des Abfangens von Daten sowie der Verkehrsanalyse und zeigen die Risiken auf, welche sich durch praktisch unsichtbare Angriffe für Nutzende ergeben. Die Betrachtung straf- und datenschutzrechtlicher Aspekte ermöglicht einen interdisziplinären Überblick über Datenschutzrisiken im Kontext von HAS. Wir stellen das erste Rahmenwerk zur formellen Modellierung von Datenschutzgarantien in Hausautomationssystemen vor und demonstrieren die Anwendung an zwei konkreten Verfahren zur Generierung von Dummy-Verkehr. In einer qualitativen und quantitativen Studie der zwei Algorithmen zeigen wir, wie Datenschutzgarantien erreicht werden können und wie sie mit der Energieeffizienz von HAS zusammenhängen. Dies erlaubt Herstellern die Konzeption und Umsetzung von Hausautomationssystemen, welche die Privatsphäre der Nutzenden schützen und die eine freie Parametrisierung ermöglichen, um einen Kompromiss zwischen Datenschutz und Energieeffizienz zu erreichen.

Acknowledgements

This thesis would not have been possible without the help and support of many people.

I would like to thank my supervisors Prof. Dr. Christoph Sorge and Prof. Dr. Christian Rossow for their support and patience, for always being available and for providing valuable feedback during every stage of this dissertation.

I would also like to thank all colleagues who I have worked with while writing this thesis. The work atmosphere has always been friendly and supportive. Without this, I couldn't have finished this work.

I would like to especially thank Raphaël Toledo for his brilliant ideas and his adamant support during our time in Tokyo, which provided the breakthrough that was needed to continue and finally finish this dissertation. Furthermore, I would like to extend my thanks to Prof. Dr. Isao Echizen for the support and the possibility of this research stay and to Associate Prof. Dr. Vincent Nozick and Dr. Thibault Julliand for making it unforgettable.

My gratitude also goes to Dr. Stephanie Vogelgesang and Dr. Jochen Krüger, who have been amazingly supportive and have made me push far beyond what I thought possible. The majority of my scientific work would not have been possible without their help and motivation.

Finally, I would like to thank my family for their constant and continuous support in any situation. The encouragement has been a driving force behind the decision to pursue a PhD in computer science and to finish this thesis.

Contents

List of Figures

Summary of Own Publications

The majority of the findings presented in this thesis have been published at conferences or in journals. This list gives an overview of all publications that have been incorporated into this work completely or in part.

Core Papers

The 2014 paper "Extrapolation and Prediction of User Behaviour from Wireless Home Automation Communication" [Möl+14] is a joint work with Andreas Hellmann, Sebastian Seitz and Christoph Sorge. It was published at the 7th ACM Conference on Security and Privacy in Wireless and Mobile Networks (ACM WiSec) and is based on the Master's Thesis "Exploration Und Evaluation von Sicherheitsaspekten Drahtloser Hausautomationssysteme" [Hel13] (2013) by Andreas Hellmann, who provided the experimental data and technical results recited in the paper. I am the main author of the paper and contributed the textual content. In this publication we demonstrate passive attacks on Home Automation System (HAS) communication and analyse two real-world installations with respect to information leakage. We show that an eavesdropping adversary can learn user habits such as sleep cycles or heating and ventilation behaviour by analysing captured traffic. The contents of the paper are incorporated into Section 2.1.

In the 2015 paper "Hausautomationssysteme im Datenschutzrecht" [MS15], a joint work with Christoph Sorge published at the 18th Legal Informatics Symposium IRIS, we investigate Home Automation Systems with respect to German data protection law. The paper was placed among the Top 10 at the Lexis Nexis Best Paper Award. The legal classification and deductions were contributed by Christoph Sorge whereas I supplied the technical context and problem description. We analyse whether data bein processed by and in the context of Home Automation Systems qualifies as

personal data as defined in the German data protection law. We conclude that HAS data must generally be considered personal data and can in some cases even qualify as special categories of personal data, which further increases the necessary levels of protection and imposes stronger requirements on the data processors. We also highlight which processes, precautions and other requirements are enforced for data processors by the national telecommunications and telemedia acts of Germany. The paper is incorporated into Section 2.4.2, but the classification has been adapted to the European General Data Protection Regulation (GDPR) which is effective since May 2018.

The 2016 paper "Deducing User Presence from Inter-Message Intervals in Home Automation Systems" [MS16] published at 31st IFIP TC 11 International Conference on ICT Systems Security and Privacy Protection (SEC) continues the technical, adversarial analysis and focuses on encrypted communication. The paper is a joint work with Christoph Sorge with me as the main author. It was shortlisted for the Yves Deswartes Best Student Paper Award. By disregarding all content data of the HAS communication, we show that a passive attacker can still deduce presence information about HAS users from the traffic volume alone. By running statistical tests on message timings, observers can identify periods of certain activity or inactivity given some limited a priori knowledge. While the amount of extractable information is far lower than for unencrypted communication, the experiments show that the privacy of the inhabitants can still be violated and that countermeasures beyond the encryption of communication content are necessary. The findings of the paper are described in Section 2.2.

A joint work with Stephanie Vogelgesang was published in the journal "Datenschutz und Datensicherheit – DuD" in 2016 under the title "Smart-Home-Systeme in Zeiten digitaler Kriminalität" [MV16]. The technical descriptions and context were supplied by me and Stephanie Vogelgesang contributed the legal analysis and dogmatic examination. In this paper we investigate attacks on Home Automation Systems with respect to the German Criminal Code. We demonstrate that established criminal offences such as trespassing, theft and property damage cannot be applied to digital settings such as HASs. We present the new laws which were introduced specifically for digital attacks and show their application as well as shortcomings of the current legal framework. The paper is summarized in Section 2.4.1.

In the 2018 paper "Modelling Traffic Analysis in Home Automation Systems" [Möl+18] we formalize the problem of privacy in Home Automation Systems and provide a framework for quantification and evaluation of countermeasures to traffic analysis. The paper was published at the 16th International Cryptology and Network Security Conference (CANS) and is a joint work with Stephanie Vogelgesang, Jochen Krüger, Isao Echizen and Christoph Sorge. I am the main author; Stephanie Vogelgesang and Jochen Krüger have contributed the legal context and international

overview in the second part of the publication. In this work we present a formal model for traffic analysis and countermeasures to it in the context of Home Automation Systems, most notably the generation of dummy traffic. The model is agnostic to concrete algorithms used for either traffic analysis or dummy traffic generation. Based on this model, we also formulate three distinct privacy goals that dummy traffic generation schemes or traffic shapers in general can aim to achieve. We further demonstrate their intuitive applicability by proving that none of the goals is reached if no traffic shaping is performed and that perfect privacy can only be achieved by forcing traffic into a constant rate. The publication is concluded with an overview over international attempts at capturing the problem in legal frameworks, with the focus being put on criminal law. We highlight the idea of the Budapest Convention, a European effort for standardisation and unification of the legal handling of digital attacks. The technical section of the paper forms the content of Chapter 3. The legal part is summarised in Section 2.4.1.

The paper "Energy-Efficient Dummy Traffic Generation for Home Automation Systems" [Möl20] demonstrates an approach towards solving the problem of privacy in Home Automation Systems. I am the sole author of this paper which was published in the journal "Proceedings on Privacy Enhancing Technologies" and presented at the corresponding symposium (PETS). In this publication we continue with the model previously introduced and further refine it by making slight amendments. We then proceed to investigate whether Constant-Rate Dummy Traffic (CRDT) can be feasibly implemented in wireless and battery-powered systems. For cases where this does not work, we present a compromise under the name Naive Exponential Dummies (NED). We prove that NED can achieve some of the privacy guarantees defined by our model and compare it to CRDT both in terms of privacy enhancement and power consumption overhead. We show that NED provides a possibility to arbitrarily tune one of the parameters while being able to predict the resulting other one. The paper is recited in Chapter 4.

Loosely Related Papers

During the work on this thesis, I have written and contributed to several other publications. The following list highlights those among them which are loosely related to the topic of this thesis. Despite being listed here, not all papers fit into the context and therefore only some are explicitly cited or summarized within this work. Papers unrelated to the matter are not listed.

In a joint work with Jochen Krüger from 2016 under the title "Metadaten – eine neue juristische Problemkategorie im Rahmen der elektronischen Aktenführung?" [KM16] we investigate the term *metadata* in a legal context. The paper was published

in the journal "MultiMedia und Recht". We show that legal frameworks fail to clearly regulate the relation between (primary) data and metadata. This introduces significant problems for data processing and surveillance. As traffic analysis attacks are launched on communication metadata rather than the communication contents, these problems apply to HASs as well.

The 2016 paper "Personenbezug bei dynamischen IP-Adressen – Anmerkung zur Entscheidung des EuGH vom 19.10.2016" [VM16] is a joint work with Stephanie Vogelgesang and was published in the journal "Datenschutz-Berater: Informationsdienst der Verlagsgruppe Handelsblatt". We present the European Court of Justice's ruling on whether dynamic IP addresses must be considered personal data in terms of data protection law. The case highlighted a central problem of the term "personal data", which appeared in previous data protection legislations as well as the current GDPR. Similar problems can appear in the context of internet-connected HASs.

In 2017 the paper "Auf der Jagd nach Schwachstellen – Eine strafrechtliche Bewertung von Portscans" [Vog+17] was published in the journal "Datenschutz und Datensicherheit – DuD". It is a joint work with Stephanie Vogelgesang, Stefan Hessel and Karin Potel. In this work we investigate port scans from a legal perspective with a focus on criminal law. We explain how port scans can be used for both benign and malicious activities and how criminal law handles this so-called *dual use*. We demonstrate that even in cases where malicious intent is present, a criminal prosecution is difficult as long as the subsequent actual attack has not been launched. These results can in part be transferred to the context of this thesis. Passive eavesdropping on encrypted communication and queries to (intentionally) public interfaces can, under certain circumstances, be exempt from criminal prosecution even if the eavesdropper's motivation malignant.

In the 2016 paper "Privacy Challenges in the Quantified Self Movement—An EU Perspective" [Lei+16] we investigate problems with privacy in several quantified self services. The publication is a joint work with Dominik Leibenger, Anna Petrlic, Ronald Petrlic and Christoph Sorge, was published in the journal "Proceedings on Privacy Enhancing Technologies" and presented at the corresponding symposium. As a central result, we find that even though users (of such services) state that privacy is important to them, actual behaviour indicates that there is little awareness for privacy problems and the danger of privacy violations. While this paradox has not been evaluated in the context of HASs specifically, there is reason to believe that customers of Smart Home products favor cheap and easy-to-use products over those that offer higher degrees of privacy.

The paper "Hardware-Keylogger: Die Tastatur in der Hand des Feindes" [VHM16] from 2016 deals with hardware-based keyloggers and the legal assessment of attacks involving them. It is a joint work with Stephanie Vogelgesang and Stefan Hessel and

was published in the journal "Datenschutz und Datensicherheit - DuD". We investigate whether passive attacks using these devices are considered criminal offences with respect to the German Criminal Code and associated legal and practical problems. While the target of the attacks is different (keyboard input vs. smart home communication), parallels can be drawn regarding the legal situation in particular and the possible detection of such attacks.

In a joint work with Stephanie Vogelgesang, Stefan Hessel and Lena Leffer titled "Mit Schirm, Charme und Kamera – Verbotene Sendeanlagen i.S.d. § 90 TKG" [Möl+17] we focus on Smart Toys—a category of products which can be considered part of Smart Home segment. The paper was published at the 20th International Legal Informatics Symposium IRIS. We demonstrate that these seemingly innoccuous items can be abused to serve as listening devices in order to spy on the users. We analyse the German legal frameworks and conclude that insecure and exploitable toys can be classified as forbidden equipment according to German telecommunication law.

Chapter 1

Introduction

In recent years, consumer electronics have picked up a new trend: Home Automation Systems (HASs) promise a variety of benefits. According to a market prognosis by the portal Statista [Blu20], the global Smart Home market is expected to grow by 16% per year in terms of revenue between 2019 and 2025.

As part of the "Smart Home", HASs relieve the inhabitants of everyday tasks: the comfort of living is increased by automating heating, ventilation and other controls; savings on energy and resource consumption due to intelligent power management; increased safety and security due to permanent observation capabilities and fully automated notification features. More and more real estate properties come with pre-installed Smart Home Systems and for those that do not, easily installable hardware is available in many electronics stores.

1.1 Home Automation Systems

A Home Automation System usually consists of several interconnected devices, each performing a specific tasks. Notable examples include, but are not limited to:

- A base station which controls other devices and usually performs scheduling and management functions. It keeps track of automation rules (such as "unlock the front door at 8:00[1]") and sends commands to the other devices.

[1] Times and dates in this thesis are written in ISO 8601 format (*YYYY-MM-DD* and *HH:MM[:SS]*), but with leading zeroes and the separation character *T* between date and time omitted for readability.

- Heating actuators which can adjust the room temperature either according to automation rules or they can be remotely controlled by the user (e.g. via a smartphone app).
- Electronic door locks which exchange mechanical keys for wireless ones or remote controls and can also be automated to unlock or lock doors at specific times.
- Temperature, humidity and air pressure sensors whose data can either be consulted by the user directly or used for automation rules (such as "turn off the heating when the outside temperature is $> 20°C$").
- Motion detectors which can be used for both comfort (e.g. turning on the lights when a person enters the room) and security (e.g. sounding an alarm when a person enters the room at night).

Home Automation Systems are produced by a variety of manufacturers and are based on disparate technologies. Sophisticated systems are often referred to as "Building Automation Systems" due to the fact that they are commonly found in public and company buildings. Communication is usually performed over wires, as devices are rarely added or repositioned during the building's life cycle. Cheaper and more consumer-oriented Home Automation Systems are often wireless to allow for an easy installation, extension and removal. They can be deployed in rented properties and transferred to other places with little effort.

It is uncommon for different Home Automation Systems to be interoperable. Some products are based on existing communication standards such as IEEE 802.11 [WiFi] or ZigBee [Zig], while others use dedicated or proprietary Home and Building Automation protocols such as KNX [KNX] or BACnet [BAC].

1.1.1 Topology

Most Home Automation Systems usually employ either of two possible network topologies: the star or the mesh topology. Examples of both are depicted in figure 1.1.

In a star network, a central node serves as the coordinator and all communication between two nodes at the "beams" of the star is relayed via the coordinator. Mostly used in wireless Home Automation Systems, the coordinator is commonly realised as a "Base Station". This device serves the user interface for configuration and schedules automation tasks (e.g. by sending commands to actuators at programmed times), but provides no immediate actuating or sensing functionality itself. Furthermore, (especially in the wireless case) Base Stations are mains-operated and do

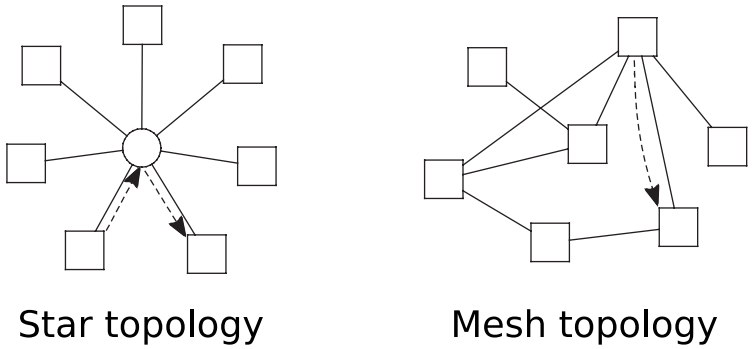

Figure 1.1: *Common network topologies in Home Automation Systems.*

not suffer from the negative impact of communication on battery lifetime. Another benefit is that communication can happen whenever the Base Station is in range. Individual nodes do not have to be in range of each other.

In wired systems, a star topology is not generally ideal. In order to be able to connect new devices more easily, a mesh topology can be used. Nodes in a mesh network can communicate directly and without the need of a central coordinator. This, however, does not mean that there cannot be a Base Station–links not involving the Base Station are merely possible. For example, a new device can be installed by drawing a cable to the nearest existing device, instead of drawing one all the way from the Base Station. In addition to this, nodes can relay other nodes' communication in order to compensate for signal attenuation (both in the wired and the wireless case) and increase the overall range of the system. While this setup can decrease communication overhead (if nodes are communicating directly and without a relay) and latency, it can also have a negative impact on the runtime of battery-powered devices and increase the complexity of the system. If message relaying is enabled for all devices, nodes have to store routing information. If it is disabled, a node has to be in range of every other node it is supposed to communicate with.

Bus networks can be considered a third topology, or they can be viewed as a special case of mesh networks where the network graph is fully connected. The latter is done herein. In bus networks, the bus (usually a cable) connects all devices and thus transports all communication. A popular example of a bus network is the KNX standard [KNX], which is widely used in Home and Building Automation Systems alike.

Star and mesh networks are not the only possible topologies for Home Automation Systems. A combination of both (e.g. two separate mesh networks which are connected by a single Base Station) is just as possible as hierarchical or other ones.

1.1.2 Summary

In summary, a Home Automation System can be defined as a network of interconnected nodes, performing household-related tasks (such as the examples given in Section 1.1) and communicating over a shared bus which is accessible to all devices in range. While technically, a Home Automation System can be set up using a unicast transmission medium, the vast majority of available products uses broadcast media such as bus systems (e.g. KNX) or wireless transmissions. We therefore incorporate broadcast transmissions into our definition. Note that this does not mean that all nodes can decrypt, let alone process all communication packets: The aforementioned topologies can all be realized using broadcast communication.

1.2 Problem Description

Naturally, Home Automation Systems are tied to the users' lives and process information about the inhabitants' private space. However, as most IoT devices, HASs have been developed with a strong focus on usability, energy efficiency and low cost. Security and privacy only play a minor role in the conception and development of Smart Home hardware.

This thesis aims to investigate privacy leaks in Home Automation Systems. As the operational area of a HAS is the user's most personal space, it is of paramount importance that the system does not negatively influence the user's privacy. There are different possible threats connected to the deployment of Home Automation Systems which need to be addressed thoroughly to ensure privacy.

1.2.1 Unauthorised Access

The most obvious threat to privacy which can be introduced by a Home Automation System is that of unauthorized access. If it is possible for an attacker to trick the system into providing access to the house or flat, the privacy as well as the property and health of the user are in danger.

Although there are cases of implementation errors which allow unauthorized persons to obtain access credentials or directly control appliances[2], this problem is mainly in the area of engineering and less an open research question. As detailed later in this work, there are established protocols and approaches to offer secure authentication and access control, preventing or significantly hampering unauthorized access. It is

[2] `https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2014-8868`, last accessed 2021-03-23.

therefore not covered in detail within this thesis. Striking a compromise between usable and secure systems is a separate area of research [Die+18; GL14], which must also be considered when developing systems. Usable *security* is however outside the scope of this thesis. Instead, we focus on usable *privacy* by considering usability issues while developing approaches for privacy enhancements.

1.2.2 Privacy Violation by Manufacturers or Service Providers

System manufacturers as well as providers of premium services (e.g. logging, analysis and provision of suggestions for energy consumption) by their very nature have access to the devices or to the data collected by a Home Automation System. This enables them to mount almost arbitrary attacks on the user, from the unsolicited collection of data to the installation of backdoors. It is virtually impossible to prevent this completely, short of not using the services or systems at all.

On one hand, laws and regulations – as detailed in Section 2.4 – can penalize offending behaviour and act as a deterrent. On the other hand, open standards, audits and open source systems can help establish trust in manufacturers and providers. For the provision of services, privacy-preserving protocols for data aggregation and analysis can be applied. Ultimately, however, this issue remains mostly a question of trust and supervision, so it is not exhaustively discussed herein.

1.2.3 Privacy Violation by Unauthorised Third Parties

Apart from manufacturers and service providers, independent third parties can also try to invade the users' privacy. Most available Home Automation Systems use some sort of broadcast medium for communication. Wired systems often use a bus, while in wireless systems the air is a broadcast medium per se. Consequently, if an attacker is able to obtain access to the medium, they are able to intercept all communication between the devices. In wireless systems, obtaining access is as easy as planting an antenna within reception range. For wired systems, it is not necessarily much more complicated, as we explain in Section 2.3.

To illustrate the problem, we consider a scenario where the majority of people living in a neighbourhood (or building) uses Home Automation Systems. An attacker who passively observes the communication of the systems can learn various information about the inhabitants, depending on the nature of the data. If, for example, lights and doors are automated (or at least supervised by the systems) and the communication is unencrypted, the attacker is able to learn how the inhabitants move through their houses or flats. If the communication is encrypted, the amount of communication might still tell them whether the inhabitants are at home and interacting with

the system. This information can then be used to plan a burglary when few or no neighbours are at home.

The main question being discussed in this thesis is

> How can we formally describe and measure privacy leakage in encrypted HAS communication?

and the follow-up question

> How can this privacy leakage be reliably prevented?

A significant problem with these questions is the lack of formalization in the field of Home Automation Privacy. One major goal of this thesis is therefore to establish such notions which can then in turn be used to design and evaluate new approaches.

In the remainder of this thesis, we start by exploring existing deficiencies in Home Automation Systems. We perform a quantitative analysis of existing HAS installations in Chapter 2 and show how attackers can leverage shortcomings in the design and implementation in order to learn sensitive information about its users. We also elaborate on the legal situation and analyse whether the current legal frameworks sufficiently protect the users from malicious actors. With the knowledge of possible attacks, we formalize notions of privacy and privacy violations in Home Automation Systems in Chapter 3 before developing approaches that aim to prevent these violations in Chapter 4.

1.3 Related Research

Passive attacks on communication are not entirely new and in different areas of research, various approaches have been developed to deal with this threat. However, the area of Home Automation Systems features a particular set of characteristics which make the direct application of these solutions to the problem at hand impossible. In this Section we address related research areas and projects which aim to solve similar problems. We compare these areas to our scenario and highlight the main differences which need to be considered when trying to transfer approaches into the HAS setting.

1.3.1 Smart Home Privacy

Little research has focused on the particular properties and distinct problems in Home Automation Systems with respect to privacy.

Mundt et al. present attacks on wired HAS communication using commercially available and cheap hardware. [MDG14] They deduct a similar analysis to the one presented in Section 2.1, albeit at smaller scale. In their work they demonstrate that wired systems are as susceptible to passive attacks as wireless ones and that an attacker can capture all traffic that is being exchanged by devices on the communication bus.

Copos et al. perform similar experiments to the ones presented in Section 2.2 on encrypted IEEE 802.11 (WiFi) traffic. [Cop+16] They analyse IEEE 802.11i (WPA) decrypted traffic from Nest thermostats and Nest smoke detectors, identifying recognisable traffic patterns (in terms of timing and amount of exchanged data) in transport-layer encrypted (TLS) and unencrypted (HTTP, NTP) connections to online services. By training on these patterns, they are able to classify traffic based on metadata such as IP addresses and packet sizes with high accuracy. While not explicitly testing against WPA-encrypted traffic, the authors argue that attackers can deduce information including presence and absence by performing the attack on MAC addresses and packet sizes alone, which are still accessible. Similar to our work, Copos et al. assume that the adversary places a wireless listening device in the vicinity of the victim's house or flat.

Both mentioned works complement our results presented in this thesis and help illustrate the extent of the problem at hand.

Apthorpe et al. perform HAS traffic analysis using a setup comparable to our experiments. [Apt+19] They monitor encrypted WiFi traffic of popular HA devices and identify interactions based on spikes in the traffic rate. Similar to our findings, Apthorpe et al. model user interaction as a stochastic process. Their work however differs from ours in several ways, including the following:

1. Their analysis is based on traffic from single devices whereas our analysis uses the output of a complete HAS system including interactions of different devices with each other. As we detail in Chapter 3, complex HASs with lots of individual devices can be more resilient to passive attacks, as the traffic approaches a constant "noise" and individual actions can no longer be recognized.

2. Their evaluation focuses on traffic overhead. As we show in Section 4.3.2, traffic overhead and energy overhead are related, but not proportional. Our evaluation takes this into account and provides an estimation of the energy

consumption overhead of dummy traffic in HASs.

3. The limitation regarding "long user activities" does not apply to our approach, because the differential-privacy-based model uses a more general notion of traffic patterns induced by user interaction. In fact, our findings apply to any system regardless of how user interaction is distributed.

Despite the differences, some ideas from Apthorpe et al.'s STP algorithm can be applied to our approach. Partitioning time into short intervals which are then viewed as constant-rate traffic sequences (either empty or padded to maximum rate) makes it harder for an adversary to identify traffic patterns belonging to individual devices. Furthermore, it reduces the temporal dimension from a continuous to a discrete one which simplifies the calculation of privacy guarantees.

Other researchers have focused on orthogonal aspects of security, for example:

- Abstract and societal risks of Home Automation Systems or IoT devices in general [JBC16; Kaa+17; Wei+15]
- Confidentiality of information and access controls [BDK01]
- Security, privacy and access control of the gateway between the HAS and the internet [Jun+12; Chi+19]
- Risks of privacy violations against inhabitants by the Home or Building Automation System operator [Mey+16]

1.3.2 Wireless Sensor Networks

Wireless Sensor Networks (WSNs) share many aspects with HASs. Devices are usually far less powerful than an average PC and communication is costly in terms of power consumption. Research on privacy in WSNs thus focuses on minimal communication overhead and low computational effort—points which are just as important in HASs. However, most WSN topologies rely on multi-hop routing for the delivery of messages. In contrast, devices in HASs usually communicate with each other directly. Moreover, research on privacy in WSNs mainly focuses on location privacy—i.e. hiding the origin of a message. [Mat+08; CP03; CWC13] Few researchers have looked into the issue of hiding the existence of traffic in general.

Yang et al. propose a scheme employing dummy traffic generation, which can be applied to systems not using multi-hop routing. [Yan+08] However, they require all nodes in the network to generate traffic constantly and delay the transmission of real messages if necessary to not introduce anomalies into the traffic patterns. While

there are devices in HASs which generate traffic at a constant rate (e.g. temperature and humidity sensors), most devices only become active once an event occurs. Modifying their behaviour to fit the scheme would significantly impair their battery lifetime and diminish the acceptance of this technology by consumers. Furthermore, messages in HASs cannot always be delayed for a (humanly) perceptible amount of time. Systems which only react to user interactions with a significant delay will most likely not be economically successful in practice. We highlight this particular matter in Section 4.3.

The authors propose a scheme for random generation of dummy traffic as well. [Sha+08] The approach offers the possibility to balance event notification delay and source location privacy (which in their scenario corresponds to event unobservability for the attacker), requiring much less overall traffic. In HASs however, delays are to be avoided whenever possible and the balance needs to be found between event unobservability and energy consumption (which is related to the amount of dummy traffic generated) rather than message delays. Furthermore, the behaviour of the inhabitants is unforeseeable. This means that an exponential distribution of message intervals might not be the optimal choice as opposed to other distributions. Ideally, the approach should be independent of a specific random distribution function and adapt to the actual behaviour.

In Low-Power Wide Area Networks (LPWANs), Leu et al. have formalized information leakage and cover traffic. [Leu+18] Their work significantly overlaps with our findings, but targets a different scenario. The system and attacker model are quite similar: Essentially, both models try to capture the confidence of an attacker guessing the genuine events in a captured traffic sample. However, we establish a dedicated model that offers verifiable privacy goals which match intuitive, desirable properties of a HAS. We then base our further work on this model to offer quantifiable privacy guarantees for HAS manufacturers.

1.3.3 MIX networks

The idea of using dummy traffic to hide communication in networks has been around for a long time. [PPW91] The first approaches use constant-rate dummy traffic. We evaluate this approach for HASs in Section 4.3 and prove that it is infeasible in some scenarios where the idle power consumption of the system is relatively low. More recently, there has been significant research on traffic analysis in low-latency MIX networks. [Lev+04]

Systems such as Tor [DMS04] or Loopix [Pio+17] provide anonymity for their users. Similar to their attacker models, we have to assume a global adversary in HASs. We demonstrate in Section 2.1 that using readily available hardware it is easy to

capture any and all traffic from a single HAS.

Contrary to MIX networks, routing in HASs is often not performed at all and messages are being broadcast (most systems are either wireless or use a bus network). Furthermore, the attacker's goal is fundamentally different in HASs: The adversary tries to identify user interaction or the absence thereof, i.e. the existence and/or pattern of genuine communication. In models for MIX networks, the adversary's goal is usually to link the sender and receiver or to estimate the average sending or receiving behaviour of users. In Chapter 3, we therefore develop and use metrics that are more suitable for the HAS scenario.

Previous research has shown limits of dummy traffic generation in MIX networks. [OTP14; Das+18] However, HASs exhibit incompatible differences: Das et al. explicitly exclude protocols where information is contained in the absence of messages. [Das+18] Oya et al. assume an attacker who tries to estimate the generic sender or receiver profile of users in contrast to examining specific, fixed timing patterns. [OTP14] Furthermore, our model is agnostic to sender behaviour, notably to changes in the genuine message rate.

Despite the differences, some approaches from MIX networks like Constant-Rate Dummy Traffic [PPW91] can be adapted and applied. Shmatikov and Wang have developed an approach which uses adaptive padding to offer privacy at a lower communication overhead. [SW06] The scheme presented in Section 4.4 is similar, but does not build on pre-sampled traffic patterns and is tailored towards the characteristics HASs.

Loopix [Pio+17] also offers a property called *Sender online unobservability* which corresponds to our goal of hiding the existence of user interaction. It does so by having the users send data through the MIX network back to themselves. In HASs, we cannot leverage this route of partially trusted MIXes, as the system does not route messages at all. However, our approach builds on some of the same principles: Inter-message timings are modelled as random variable. This allows for an intuitive and well manageable model.

1.3.4 Website Fingerprinting

A large body of literature is available on the topic of website traffic fingerprinting and recognition. [Cai+12; Pan+11; Kir+08] The general idea is that web browsers exhibit a unique traffic pattern when accessing a single website. These patterns can be learned and later recognized to match users to websites even if the traffic is encrypted and possibly routed through a MIX network.

Some ideas from this field can be adapted and used in HAS settings as well. For

example, traffic fingerprinting algorithms could be used to recognize known communication patterns between particular pairs of devices. However, in our particular use case the attacker has little to no a priori information about the system, the devices or the inhabitants. The attacks presented in Section 2.2 leak information about the users without requiring large amounts of sample data and our model in Chapter 3 is abstract. In general, the models of website fingerprinting and HAS traffic analysis differ in several regards:

1. In website fingerprinting attacks, the initiator of the communication is known and the attacker tries to match the counterpart against a set of known and publicly reachable entities. In our scenario, only the HAS to which the communicating parties belong is known. The attacker tries to determine whether some particular category of communication (e.g. genuine user interaction) is happening.

2. Countermeasures against website fingerprinting attacks generally aim to be applied at the user's node and possibly at nodes along the way to the web server. Unrelated third parties and the website itself are to be protected from negative side-effects of the countermeasure. In HASs, all nodes are under the user's control. Except for regulatory thresholds, no third parties are involved and have to be considered.

3. In computer networks, a large traffic overhead degrades the performance of the system. In HASs, this directly affects battery lifetime and can lead to system unresponsiveness if regulatory thresholds for communication bandwidth are exceeded.

4. Routing or at least direction information is available in computer networks [Pan+11]. In broadcast HASs where the destination of a packet is sent unencrypted, only this information is available to an observer. If the destination address is encrypted as well by a mechanism such as SlyFy [Gre+08], no routing or direction information is available at all.

1.3.5 Differential Privacy

Definitions of Differential Privacy are used to model problems very similar to the one at hand. They are used for the development of techniques which provide unobservability of events or user data. As we show in Chapter 3 however, the two models are not quite compatible.

Approaches from the field of Differential Privacy are tailored to specific computations on input data. The definition of the original work [Dwo06] states that *a randomized*

function K gives ε-Differential Privacy if for all data sets D_1 and D_2 differing on at most one element, and all $S \subseteq Range(K)$,

$$Pr[K(D_1) \in S] \leq e^{\varepsilon} \cdot Pr[K(D_2) \in S] \tag{1.1}$$

This already indicates that ε-Differential Privacy is a property of a specific function K. However, we cannot know in advance which computations an attacker will perform on captured HAS traffic to learn information about the user. Thus, a solution to the problem of traffic analysis in HASs must be independent of K or at least apply to a wide range of different possible computations K.

As an example, Dwork et. al. have developed an approach to continuously monitor an event source and count its events with the counter guaranteeing ε-Differential Privacy even when its internal state is visible to the attacker at some point in time. [Dwo+10] However, they describe a concrete implementation of the counting function which satisfies the requirements. The guarantees do not necessarily apply to other functions.

Closely related to Differential Privacy is the area of Private Information Retrieval. [TDG16] Our model presented in Chapter 3 takes ideas from this field and applies them to a communication network.

In his dissertation [Cop17], Copos defines a model for capturing side channel information in arbitrary systems. Contrary to the first intuition, the approach is effectively very similar Toledo's idea [TDG16] and tries to capture the same question: What is the probability of the (hidden) system performing a particular sequence of steps, given some adversary-observed output? Or, reversed: What is the probability of observing a particular output given a system that performs a particular set of operations?

1.3.6 Steganography and Covert Channels

The goal of steganography is to hide the very existence of data. At first glance, this makes it an ideal mechanism for hiding communication in HASs. However, steganography requires some cover data (also called a *carrier*) to embed the actual data in. [Kes04] Since there usually is no such cover data available in HAS communication, many steganography approaches cannot be applied to this setting.

Hiding the traffic among noise of a wireless channel has been investigated by Bash et. al. [Bas+15] They prove that a passive adversary is unable to decode transmissions if the sender abides by certain rules (i.e. limiting their transmission power and using suitable spread spectrum techniques). The application of their approach to HASs faces three issues:

1. The authors highlight fundamental limits of this approach. It is unclear (and a driving force behind this work) whether approaches on higher layers of the network stack can either exceed these limitations or achieve the same results at lower (manufacturing or communication) costs.

2. The approach was developed for wireless networks. While in theory one could apply the same techniques in a wired network, it would likely require the introduction of artificial noise generators (due to the much lower amount of "natural" noise) and would require transceivers to be more complex.

3. Due to the fact that the approach works on the physical network layer, it needs to be incorporated in the transceivers of appliances. Depending on the spread spectrum technique used, this implies changes to the hardware and/or firmware. Solutions on higher layers might be easier to implement for HAS manufacturers, because they only require changes to the software written by them.

1.4 Outline

This thesis is structured as follows.

In Chapter 2 we investigate the problem of privacy in Home Automation Systems from a practical perspective. We demonstrate the first published passive attacks on the communication of a complete HAS and illustrate that adversaries can form highly detailed behavioural profiles by listening to unencrypted traffic. We show that information such as sleep cycles, heating and ventilation habits can leak to interested parties. We then proceed to analyze how well the application of encryption can protect against this kind of attacks. We expose that passive attackers can still deduce information about user presence by applying statistical tests to traffic metadata. Lastly we delve into the legal situation regarding privacy in HASs. We highlight the past and current situation regarding both criminal and data protection law. Furthermore, we sketch shortcomings of current legal frameworks and how legislators aim to define boundaries for permitted usage of communication equipment.

In Chapter 3 we present a formal model to describe, quantify and work on privacy in Home Automation Systems. We show how notions of privacy can be used to quantify information leakage and how protection mechanisms can be evaluated against a universal model. To the best of our knowledge this is the first general model which can be applied to any network topology, communication protocol and privacy-enhancing technology in the context of HASs, all while capturing the system as a whole. Our model is uniquely suited for the application in this particular setting and intuitive properties of protection mechanisms can be easily formulated

using the notions established here. We demonstrate the application of the model by taking two exemplary cases—no dummy traffic and Constant-Rate Dummy Traffic (CRDT)—and evaluating them against the notions of our model.

Chapter 4 focuses on finding a compromise between privacy and energy-efficiency for battery-powered and bandwidth-limited HASs. We investigate the effect of Constant-Rate Dummy traffic on data rates and power consumption using different models. As the first quantitative study in this area, we show the relation between traffic volume and energy consumption when applying Constant-Rate Dummy Traffic in a Home Automation System. Contrary to the first intuition, we demonstrate that in certain settings, CRDT is a suitable and feasible method to provide optimal privacy at moderate cost. For scenarios where the power consumption of CRDT cannot be satisfied, we present an algorithm which offers certain privacy guarantees while providing arbitrary tuning capabilities. This allows manufacturers and users to strike a compromise between privacy and energy efficiency.

The thesis is complemented by a conclusion in Chapter 5, which highlights the results of our findings and indicates possibilities for continued research.

Chapter 2

Attacks on Current Commercial Systems

In this chapter we investigate attacks on the privacy of Home Automation System users. The goal of this chapter is to develop an understanding of how the privacy of the inhabitants can be violated by a passive adversary monitoring the HAS traffic. This understanding is necessary to develop a formal model and to come up with approaches that guarantee privacy-preserving interaction with such systems.

While more recently, manufacturers have started to encrypt Home Automation System communication or based their system on standards which offer this feature, this has not always been the case and is still neglected by some. We start by analysing the privacy leakage in unencrypted systems and then proceed to investigate the use of encryption, showing that this does not fully protect the users' privacy. Afterwards we take a look at how the legal situation has developed up to now.

The following analyses focus on data from different Home Automation System. One part of the data stems from two HomeMatic installations – a commercially available mid-budget HAS that offers a wide variety of applications. These properties (availability, affordability and range of use cases) make it an ideal setup for our investigations. The system may gain a high popularity due to the first two properties and the results are representative for HASs due to the third one. Details on the setups are elaborated on in Sections 2.1.2 and 2.1.4. Another part of the data was taken from a series of news articles. The system was assembled by the tenants themselves, using various devices from different manufacturers and building on an IEEE 802.11 (Wi-Fi) network. It serves as a suitable example for a different architecture and demonstrates that the findings are not unique to HomeMatic systems. Details

on the data are found in Section 2.2.

2.1 Unencrypted Communication

Research Question: *What information can a passive attacker learn about the user of a Home Automation System if communication is unencrypted?*

Even without a detailed analysis, it is obvious that unencrypted HAS communication can leak private information to eavesdroppers. However, in order to show the extent of this privacy violation, we performed a study using two distinct real-world installations of Home Automation Systems. The study was initiated as part of a Master's Thesis [Hel13] which lays the groundwork for this work and the results have been published [Möl+14]. They are presented in the following.

The goal of the study is to estimate the level of detail that can be learned about a victim's habits and accommodation situation from the traffic data of their HAS. Specifically we investigate whether an attacker can identify recurring habits such as regular times of going to sleep and waking up, going to and coming from work as well as heating and ventilation behaviour.

2.1.1 Attacker Model and Attack Methodology

For the analysis we assume the role of an attacker with a limited, but realistic set of abilities:

- The attacker is global and thus able to capture all messages which are transmitted by the system. It is possible to achieve this by using cheap and readily available hardware.
- The attacker knows the geographical location of the victim's property. Since the attacker will usually place a receiving device in the vicinity of the Home Automation System, they know the location. This is not a necessary assumption for the complete attack, but the knowledge of the location was used to find out even more about the victim's habits.

After getting consent from the participants (cf. Section 2.4.2), we placed a capturing device (using a CC1101 USB Lite (CUL) stick) in the vicinity of two properties which had HomeMatic Home Automation Systems installed. We captured the traffic of the systems over a period of several weeks and subsequently analysed it using different methods described hereafter.

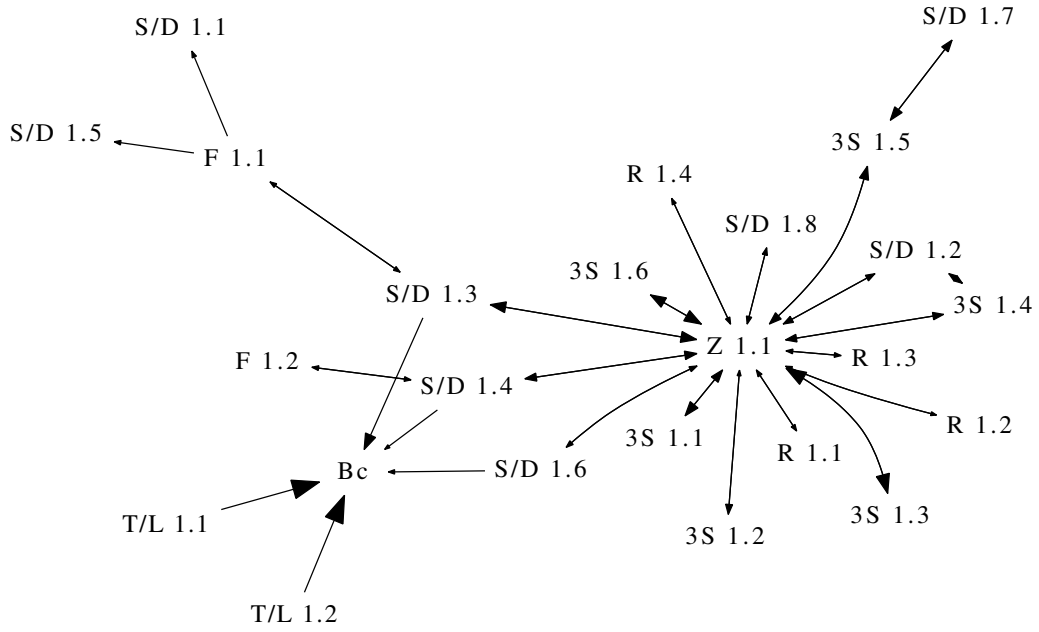

Figure 2.1: *Communication Graph of a HomeMatic Home Automation System. The abbreviations for device types are listed in Table 2.1.*

2.1.2 Analysis Procedure

First, the different devices making up the HAS have to be identified. Using the addressing information from the captured packets as a start and applying regular expressions as well as plausibility checks to the message packet contents lead to a mostly error-free identification of all devices in use. The plausibility checks (e.g. temperatures in Germany are unlikely to reach values below $-20°C$ or above $50°C$) in particular help solve ambiguities in the message payloads. Slight manual adjustments based on plausibility of communication links (e.g. a remote control is more likely to communicate with a lock than with a thermometer) further improve the classification.

After identifying the individual devices, communication graphs similar to those in Figure 1.1 can be generated to visualise the relationships between them. One example of such a graph is shown in Figure 2.1.

Message contents can be interpreted depending on the sender's and receiver's device types. For example, messages from a thermometer to the base station contain numeric values (i.e. the temperature) and messages from the base station or a remote

Abbr.	Device
3S	Tri-state Sensor
Bc	Broadcast Address
F	Remote Control
KF	KeyMatic Remote Control
KS	KeyMatic Lock
R	Smoke Detector
S/D	Switch / Dimmer
ST	Heating Actuator / Thermostat
T/L	Temperature / Humidity Sensor
Z	Central Unit

Table 2.1: *Abbreviations for the different devices.*

control to a lock contain a desired binary state (locked or unlocked). Using these interpretations, we are able to identify correlations between events which indicate user habits and automation rules.

Table 2.1 shows the abbreviations used for devices in the following summary of the results.

After plotting the general communication links between devices, we project the messages to and from a single device in relation to the time onto a 2-dimensional graph. We call this step the *Manual Examination of Message Graphs*. This graph type helps identify temporal structures and periodic events.

In addition to the manual identification of correlated events, we perform an automated *correlation analysis* using a sliding window approach. We define an event as a 4-tuple of sender address, receiver address, message type and message content. For each event e, we examined other events e^* that occured in a time frame after e. We then pair e with each of these other events e^* and for each pair (e, e^*) calculate the number of occurences over the whole observation time. 3 parameters allow filtering out events: The minimum total number of occurences of the event e, the minimum chance of e being followed by e^* and the length of the time frame in which e^* has to follow e in order to be counted.

With a similar approach we *filter out programmed automation rules*. We assume that automated events occur at a fixed time which differs only marginally. For each event we collect all occurences over the observation period. We then strip the date so only the time of day remains. As a last step, we sort the occurences in chronological order. The sorted list allows for an easy identification of events that often occured at similar times during a day.

2.1.3 Analysis Results: System 1

The Home Automation System was installed in a single home. We recorded 45,679 messages over a period of 3,111,908 s (36.02 days). The individual message timestamps were saved with a precision of 1 second. During the data capture, the listening device experienced two outages during which no message was captured. These appear as gaps larger than 4000 s in the data. These gaps are ignored, i.e. the time during which no message was captured is removed from the data, for the remainder of this thesis. Subtracting these gaps from the overall timespan leaves it at 3,066,575 s (35.49 days).

Only a few devices cannot be identified with acceptable certainty. As we found out after a debriefing with the owner, this was the case for the smoke detectors that only send heartbeat messages to the central unit as well as one of the tri-state sensors which did not report any state changes during the observation period.

The analysis follows the aforementioned procedure. We first identify the installed devices and infer information about the accommodation situation from the system architecture. We then proceed to examine individual device's and device groups' communication patterns in order to identify habits and automation rules.

Communication Overview

Figure 2.1 provides a graphical overview of the communication. Expectedly, the central unit *Z 1.1* communicates with most sensors and actuators so it can be easily identified. The graph also allows us to identify which components are directly paired with each other and do not exclusively communicate over the central unit. This information might be useful, for example for later active attacks against the system, and might also be an indicator for manual interaction.

Manual Examination of Message Graphs

Figure 2.2 shows the temperature status messages of two temperature/humidity sensors *T/L 1.1* and *T/L 1.2*. The obvious, significant difference in the temperature ranges leads to the conclusion that *T/L 1.1* is located outside the house whereas *T/L 1.2* is located on the inside. We confirmed this by comparing the recorded values with weather reports from the area.

Values of the in-house sensor *T/L 1.2* consistently lie in the range between $20°C$ and $25°C$. The not perfectly regular rise and fall suggests that the heating is controlled manually and indicates a user habit. Furthermore we deduce from the low outside

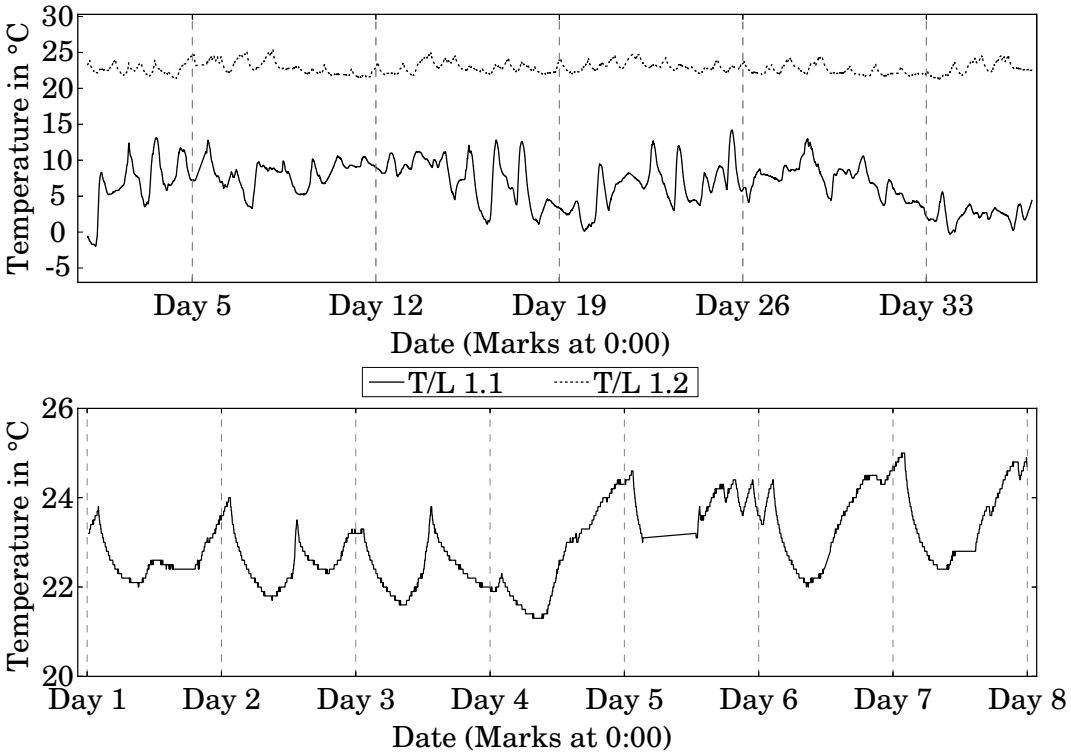

Figure 2.2: *Temperature values of* T/L 1.1 *and* T/L 1.2 *over the course of 6 weeks (upper) and temperature values of* T/L 1.2 *over the course of 7 days (lower).*

temperature and the slow temperature drop inside that the room is rarely ventilated by widely opening the windows for at least 10 minutes.

The tri-state sensors[1] can be coarsely divided into two groups. The first group consists of two sensors, *3S 1.2* and *3S 1.4*. Both send only very few messages with usually the same content. No conclusions can be drawn about their role. The second group consists of the remaining tri-state sensors whose traffic mainly consists of `open` and *close* state announcements.

Examining the protocol data for *3S 1.3* and *3S 1.6* which is visualised in Figure 2.3, reveals that they frequently switch the state. The `open` state is never held for more than 1.5 minutes and usually lies in the order of seconds, suggesting that the sensors are placed on doors rather than windows. The activity over longer time spans shows gaps during the nights and early mornings.

[1]The family of tri-state sensors includes different devices: Window sensors that distinguish between `open`, `closed` and `tilted` and door sensors which only distinguish between `open` and `closed`. Technically, they are the same kind of device.

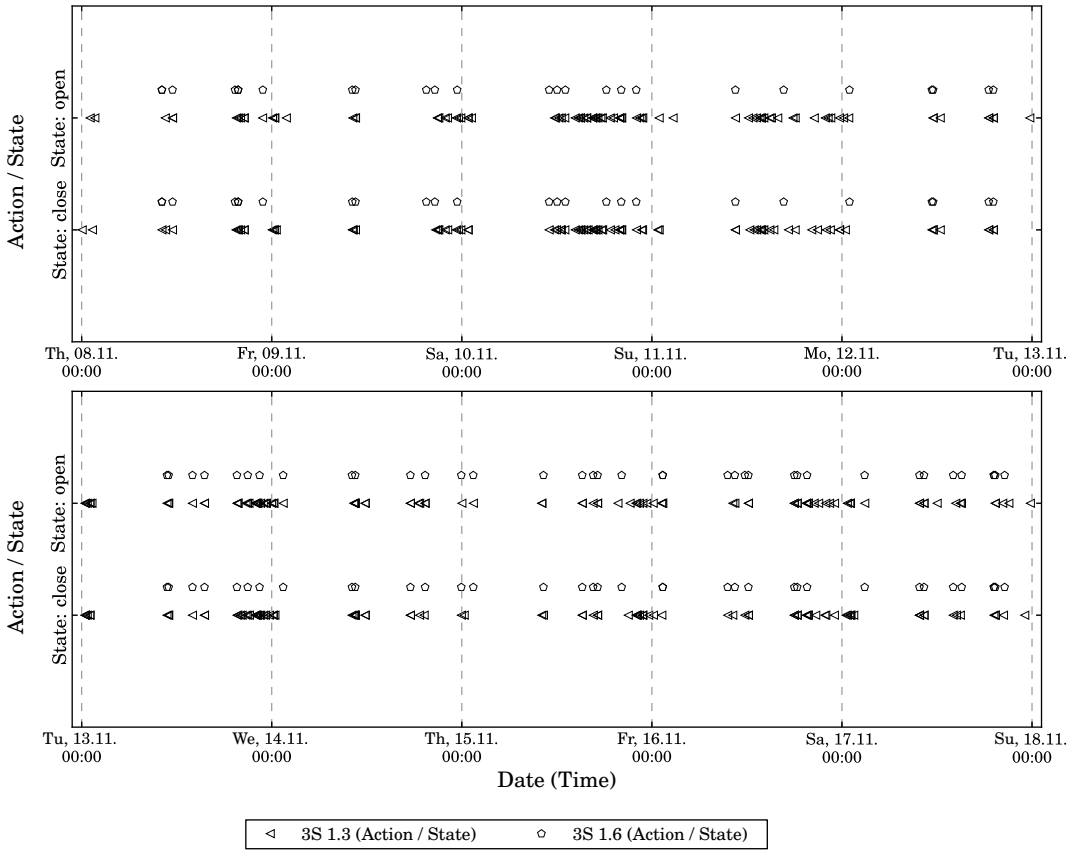

Figure 2.3: *Sensor values of* 3S 1.3 *and* 3S 1.6 *over the course of 10 days. Since they are attached to doors, they only report two distinct states instead of three, as the name suggests.*

Since tri-state sensors notify about state changes usually caused by user interaction, these gaps are good indicators for the inhabitants' sleep cycles. *3S 1.6* changes its state to `closed` some time before the gaps, which supports the assumption that it is installed on the front door. This is a major discovery for an attacker, because they can tell when the first inhabitant leaves the property in the morning. If there is only one inhabitant, this knowledge is already enough to plan a burglary during the user's absence.

Similar to the tri-state sensors, the switches and dimmers can be divided into two groups. *S/D 1.2*, *S/D 1.6* and *S/D 1.8* exhibited very little activity over the observation period. *S/D 1.3* and *S/D 1.4* regularly alternate between *on* and *off* states. The activities of *S/D 1.4*—shown in Figure 2.4—reveal a strong regularity in the afternoon between 16:30 and 17:00 when the actuactor is switched *on* and at 1:00

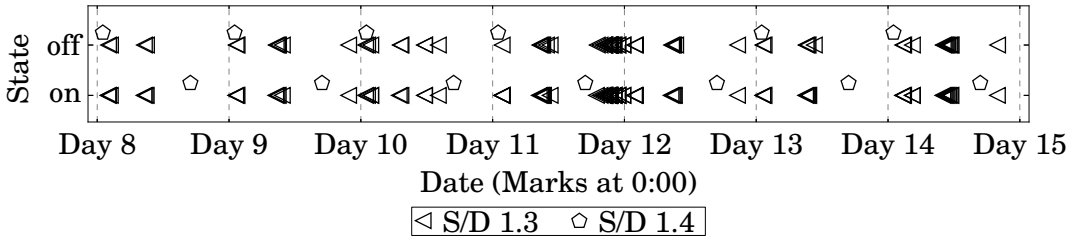

Figure 2.4: *Data sent to and from switches/dimmers* S/D 1.3 *and* S/D 1.4.

when it is switched *off* again. Each day the former action is performed 1.5 minutes earlier than the day before. This is a very strong indicator for an automation rule which compensates for the sunset times. This assumption is supported by the fact that the respective commands come directly from the base station rather than a remote control.

Regarding *S/D 1.3*, we found regular activity on weekdays between 1:00 and 2:00 as well as between 8:30 and 9:30. The slight variations support the conclusion that this indicates a user habit rather than an automation rule. The payloads of the recorded packets revealed that on weekday mornings, the base station would send timer commands to the switch between 8:00 and 9:15. These commands would turn the switch on for an hour after which it would turn itself off again. We attributed this behaviour to either a habit of the user after waking up or to an alarm function actually waking the user by e.g. turning on the lights.

Another regularity is the absence of activity of *S/D 1.3* between 13:00 and 17:30. The fact that this coincides with a lack of activity of *S/D 1.7* (12:00 to 18:30) lead us to the conclusion that the user is absent (e.g. at work) during this time of day.

Correlation Analysis

During the correlation analysis we tested different possible parameter values. Since we have no prior knowledge about the systems, the only viable approach is to manually determine suitable threshold values.

The results of the correlation analysis largely support the previous findings which could already be observed in the graphical analysis. However, we can make some additional findings.

In 72.5% of all cases where sensor *3S 1.6* was turned on, it would be turned off again within 10 seconds. A similar behaviour was observed for sensor *3S 1.4*, which was closed within the 10-second interval in 58% of all cases. In accordance with our

reasoning above, we conclude that the sensors are installed on doors rather than windows.

When selectively analysing the behaviour of *3S 1.3* and *3S 1.5*, we found them to act very much alike. In most cases the state *open* did not hold for longer than 90 seconds. In the case of *3S 1.3*, this was especially interesting when considering the timer commands sent to it by the base station in the mornings. The commands would turn on the switch for 300 seconds, but in 96% of these cases, the switch would be manually turned off within the first 90 seconds after reception. This supports our theory that the switch is part of an alarm function.

Filtering Automated Events

In order to filter out automated events, we initially started the analysis with very strict parameters: The minimum number of occurrences of an event were set to 120, the maximum overall deviation of events possibly originating from the same automation rule was set to 60 seconds and the maximum deviation of two consecutive events from the same rule was set to 30 seconds. The only event to match at first was a command from the base station which turns off *S/D 1.4* at precisely 1:00, confirming our assumptions. We then proceeded to loosen the parameters to search for other rules. The command coming from the base station and turning on *S/D 1.4* in the afternoon between 16:25 and 17:10 came out next. Although the maximum distance between the different occurrences is 38 minutes, we concluded that this event indicates the presence of an automation rule. The distance between two consecutive occurrences is about 90 seconds and each event occurred later than the one on the day before. Rather than a user habit, we attributed this regularity to an automation rule that incorporates sunset times.

2.1.4 Analysis Results: System 2

The second installation was split up in two parts which are interconnected via a VPN. One part was the user's private flat and the other part was his office. For this reason, we performed the data collection in two parts. We first installed the sniffer at the office, then moved it to the user's home. During these two periods we recorded 33,708 packets over a timespan of 720,112 s (8.33 days) and 999 packets over a timespan of 1,161,565 s (13.44 days), respectively. Similar to System 1, the message timestamps were measured with a precision of 1 s. The total number of distinct devices is 20.

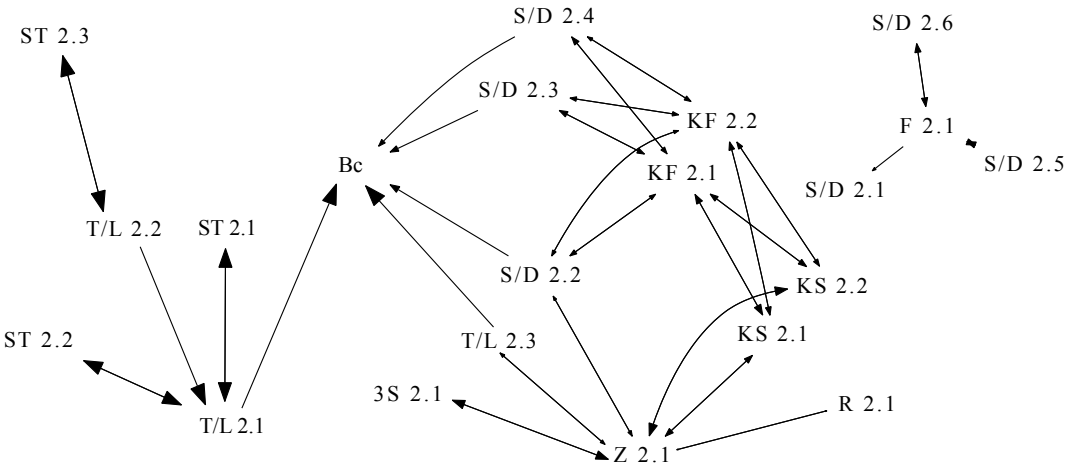

Figure 2.5: *Directed communication graph for the second installation. The abbreviations are again those from Table 2.1*

Communication Overview

Figure 2.5 shows the communication graph for the second installation. It can easily be seen that there is no single center of communication as opposed to the first system. Many devices are paired directly with each other and only 6 of 19 available peripheral devices communicate with *Z 2.1*. Furthermore, neither the remote control *F 2.1* nor any of the three switches paired with it communicate with *Z 2.1*. Thus, we can almost certainly rule out any automation rules for these devices, which gives us more insight into the habits of the inhabitants. Nevertheless, the segmentation of this installation and the VPN connection between the segments make it difficult to derive information about the physical presence of the inhabitants from the automated events alone. The KeyMatic remote controls *KF 2.1* and *KF 2.2* are paired with many actuators in addition to the KeyMatic door locks and both remote controls are paired with both door locks.

Manual Examination of Message Graphs

The temperature values given by the sensors *T/L 2.1* (shown in Figure 2.6), *T/L 2.2* and *T/L 2.3* as well as the corresponding actuators *ST 2.1*, *ST 2.2* and *ST 2.3* are strong indicators for an automated heating concept. Over the weekends, the temperatures drop gradually and then rise again sharply at the start of the week. The different temperature and humidity curves and the temperature differences of up to 10°C between the sensors lead us to the assumption that they are installed in

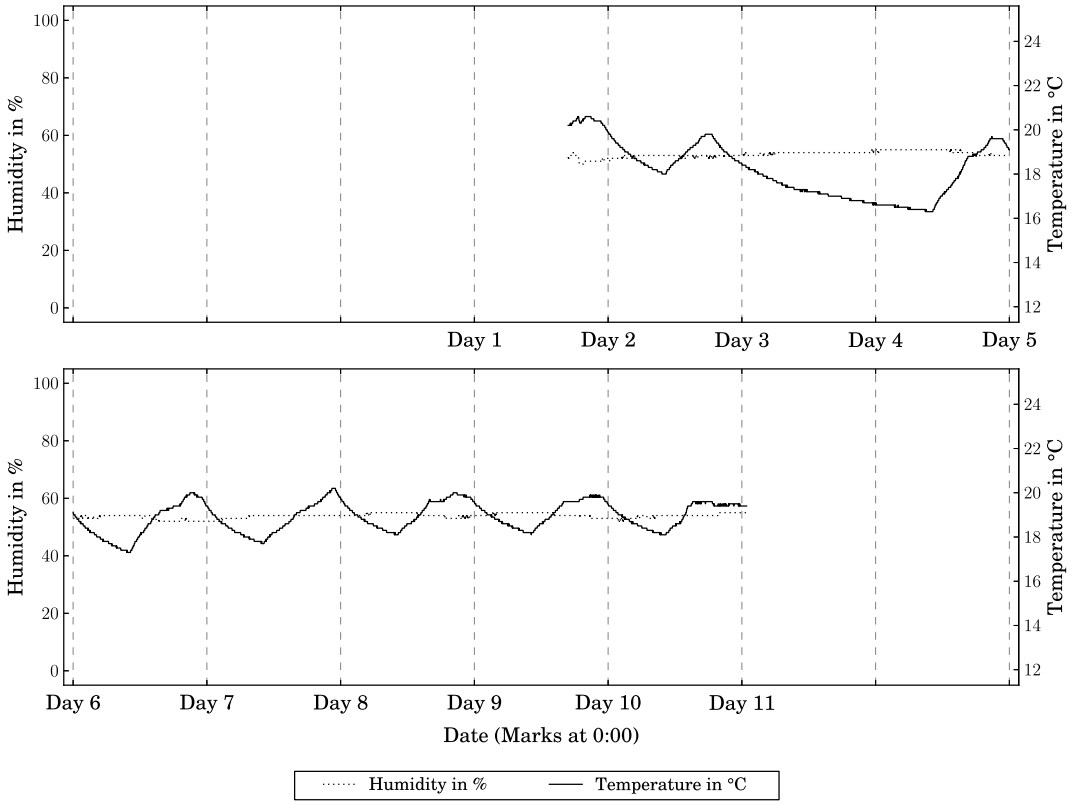

Figure 2.6: *Temperature and Humidity Values from Sensor* T/L 2.1.

2-3 different rooms.

When examining the activity of the remote control *F 2.1* we find events only in the first part of the observation period. This means that the user only uses the remote control within the office itself and does not control any devices at home. The observed activity is thus a very reliable indicator of when the user is definitely present at the office and thus, not at home.

Keymatic Door Lock System

The most interesting data for the second installation came from the automatic door lock system. Every day at about 9:15 as well as between 20:00 and 22:00 the door locks report their status to the central unit Z 2.1. We observe that there are always two messages shortly after one another, first *3S 2.1* sends the state `open` and after a maximum of 60 seconds the state *close*. Correlating the states of *3S 2.1* and

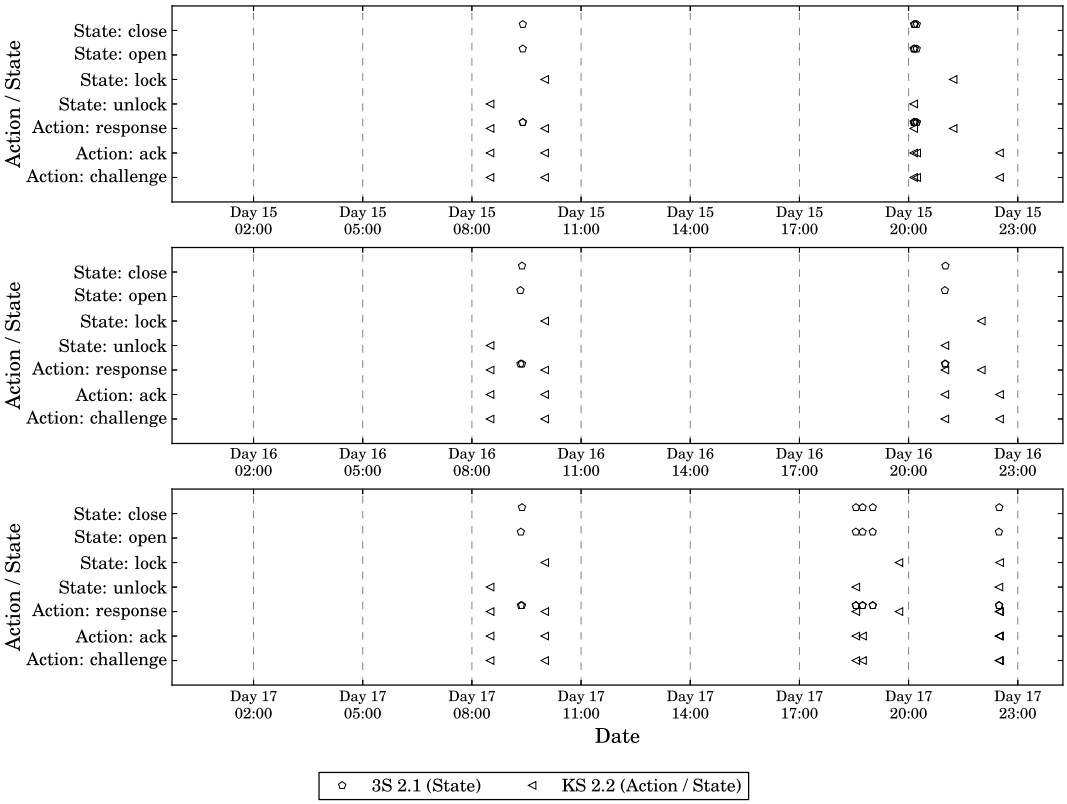

Figure 2.7: *Sensor data from* 3S 2.1 *correlated with commands received and states reported by* KS 2.2*.*

KS 2.2 as shown in Figure 2.7, we conclude that *3S 2.1* is installed on the same door as *KS 2.2*. Due to this combination, the presence of the inhabitant can be easily predicted. Usually there is nobody at home between 9:30 and 20:30, except for Mondays, where the time of absence lies between 13:00 and 21:00 (assuming the user lives alone).

Correlation Analysis

Our correlation analysis found strong correlations between the thermostats and the heating actuators as well as between the remote control *F 2.1* and the actuators *S/D 2.1*, *S/D 2.5* and *S/D 2.6*. The thermostats show a consecutive acknowledgement of the heating actuators' new positions in 98.8% of all cases. The switch and dimmmer actuators even send their status as a reaction to a previous command from the remote control in 100% of all cases.

We see similar clear results for the reactions of the Keymatic door lock systems and the switches/dimmers *S/D 2.2*, *S/D 2.3* and *S/D 2.4* to the Keymatic remote controls *KF 2.1* and *KF 2.2*, where we recorded a reaction in 90% of all cases. In addition to what we already knew, the correlation analysis reveals that the tri-state sensor *3S 2.1* seems to have a relation to *KS 2.2*, since over 60% of all status changes of *KS 2.2* result in a status change of *3S 2.1*.

Filtering Automated Events

To filter automated events we used the same approach as for the first system. We generally found the results to support our findings from the manual analysis.

Using the automatic filtering method we can confirm our assumption that the unlocking command sent from the base station *Z 2.1* to *KS 2.2* at 8:30 in the morning does belong to an automation rule. The same holds for the command that locks *KS 2.2* again at 22:30.

Furthermore we found that the the heating actuators are automatically turned off at night. They regularly receive a `Pos.: 0%` command from the temperature/humidity sensors.

2.1.5 Confirmation of Results

After our experiments, we interviewed both system owners and discussed our findings with them. We were able to confirm our conclusions about the locations and purposes of the different devices. The owners also confirmed our assumptions about automation rules and user habits.

2.1.6 Conclusion of the Analysis

As can be seen from the results of our analysis, unencrypted communication in Home Automation Systems poses a significant threat to the privacy of the users. In general, these systems leak a large amount of information to any observer keen enough to look for it. No prior knowledge about the installation or the victim is necessary to perform this kind of attack. Criminals can indeed resort to a form of *wardriving*, where they move around populated areas carrying a receiver and searching for systems using a supported protocol.

As we show later on, even encrypted systems do not protect the user's privacy appropriately. Knowledge gained from the analysis of traffic of an unencrypted

system such as common communication links and communication patterns can be used to partly remove the confidentiality that encryption is supposed to offer.

While we performed many tasks and checks manually during our experiments, most of them can be automated with the knowledge from our findings. Our parameters for finding automated events proved useful, as long as the automation rules did not change times themselves. Possible communication links, the usual frequencies with which the different devices send status messages and the number of messages being exchanged for each action of one device can be used to program heuristics which can then in turn identify devices in a system even if packet payloads are encrypted.

In conclusion, the analyses have demonstrated which habits and automation rules can definitely be extracted from unencrypted HAS traffic. While it may be possible to extract even more details, the goal of this section was to gain insights into the danger of passive attacks and to understand how these attacks can be performed. This knowledge is used later to derive a model for attacks on privacy in HASs.

2.2 Encrypted Communication

The most obvious countermeasure to the passive attacks detailed in the previous sections is the encryption of all traffic. However, as sketched in Section 1.2.3, even metadata can disclose sensitive information to an attacker. In order to quantify this leakage, we perform a second study on a dataset overlapping with that of Section 2.1. We assume the same attacker model with two slight modifications:

As a restriction, the attacker is only able to observe message timestamps, i.e. they are unable to see packet contents, lengths or metadata which allows the identification of individual devices (e.g. addressing information). This is a strong limitation of the attacker's abilities, but serves an important purpose: We aim to show that even with this highly limited visibility, passive adversaries are still able to infer presence information about the HAS users. As detailed in Section 3.1, the detail of information assumed to be visible to the attacker can be varied without invalidating our findings and the established model. While the first part of this chapter aims to show how much attackers can learn when given a high amount of detail, this section deals with the privacy violations that are possible even when the information is highly limited.

As a second modification to the attacker mode, we assume that the adversary has captured all communication packets during one hour of HAS operation and knows whether the inhabitants were at home during this time. System 1 provided us with the necessary information for their dataset. For System 2, we could not obtain the necessary data. We therefore used a different dataset (denoted as System 3) which

we obtained from a series of news articles[2].

The question we are trying to answer by this study is:

Research Question: *If an attacker has captured traffic from a user's HAS over a given time frame and knows whether the user was present at that time, can the attacker infer the user's state (presence/absence) during a second timeframe from that (second) timeframe's metadata?*

To answer this question, we assume the role of an attacker and collect pairs of traffic samples. For one sample we look up the user state and for the other sample we try to infer it from traffic metadata. We then compare the guessed state to the actual state of the second sample.

A base for the following study was built and used to obtain first results. These have been published.[MS16] After publication, we have re-run the simulations with slightly different parameters, resulting in a higher number of samples and, consequently, a longer computation time. The idea behind this reiteration is to reinforce the results from the first run, leveraging access to new computational equipment. The different parameter choice—the step size for sampling message groups was decreased—is elaborated in the following section.

2.2.1 Attack Methodology

As specified in the attacker model, we only use message timestamps for the analysis. Using the captured communication packets from two different time frames of one hour each, we try to find similarities in the statistical distribution of inter-message intervals (the time difference between two subsequent messages).

As a first step, we annotate each message with the user state: Present and Absent are chosen based on the available data. A third state, Asleep is introduced to handle the fact that during the night, users are usually asleep and thus the activity of the system is reduced. Since we do not know the exact sleeping times and habits of the users, we annotate all samples between 22:00 and 08:00 with the state Asleep and exclude them from further analysis. We only investigate messages whose state is either Present or Absent.

Analysing each system by itself, we construct intervals of 1 hour each during which the user state did not change. The intervals overlap and are sampled in steps of 1

[2] `https://www.spiegel.de/netzwelt/gadgets/sensorenresidenz-was-ein-smart-home-ueber-seine-bewohner-verraet-a-1065421.html`, accessed 2021-03-23. The data is no longer available from the articles, but is archived on a separate page: `https://opendatacity.github.io/sensorenresidenz/`, accessed 2021-03-23

minute. This accounts for the fact that the attacker might start their capture at any point in time. We have performed the experiment with larger step sizes and different interval lengths and have obtained similar results. The results are compared in Sections 2.2.2 and 2.2.3.

The first iteration of the simulation [MS16] was run using intervals that were sampled in steps of 1 hour–a compromise between sample data size and simulation runtime. Furthermore, similar samples (intervals with the same message inter-arrival times) were only considered once. Leveraging new equipment with higher computational power, we have decreased the step size to one minute and have modified the simulation to treat every sample separately, no matter if a similar one has been encountered before. As a result, the number of samples is roughly 60 times as high as during the first iteration and the number of pairs consequently is about 3600 times as high.

By using overlapping intervals, we also account for the fact that actual attackers are able to arbitrarily choose the time of their attack. Overlapping message groups thus include more possible attack settings for the given data and the confidence of our conclusions is thus higher. Given that the results from the larger step size do not differ significantly from those with a smaller one, we conclude that experiments with smaller step size are unlikely to reveal new information. Due to the quadratic increase in computation time and storage requirements, we have refrained from further decreasing the step size.

For each interval, we gather the messages sent during this time into a *Message Group*. Each Message Group is thus identifiable by its system and the timestamp of the start of the capture period. Also, as per the construction of intervals described above, each group has a fixed user state.

For System 1, we obtain 19,056 Message Groups, 10,836 with state PRESENT and 8220 with state ABSENT. For System 3, we obtain 23,696 Message Groups, 15,122 with state PRESENT and 8574 with state ABSENT.

For all combinations of 2 different Message Groups (only considering those with states PRESENT and ABSENT)—181,556,040 in total for System 1 and 280,738,360 for System 2—we perform 3 statistical tests described hereafter. We then visualize the results in box plots, both overall per system and individually for each combination of user states. In a second step, we test different thresholds for all tests and plot the true and false positive rates in ROC diagrams.

In order to compare the samples, we use the Kolmogorow-Smirnow Two-Sample Test [Kol33], the Chi-Square Test of Independence [Pea92] and the "Message Counts Test". The former two are well-known and thoroughly tested approaches to compare samples from random distributions. The third test was developed specifically to complement the other two tests by using a robust and intuitive metric.

These statistical tests tackle the null hypothesis that *the two samples have the same underlying distribution function.* We assume that the traffic patterns of the HAS follow an unknown random distribution and that traffic samples of intervals with the same user state follow the same distribution with similar parameters. Instead of rejecting the null hypothesis with a certain confidence at a threshold (which in turn depends on the desired confidence), we analyse the computed test statistics and try to determine suitable thresholds ourselves. The reason behind this is twofold: On the one hand, we do not have any a priori knowledge about the underlying distribution functions. In fact, we cannot even be sure that the traffic can be modelled using a random distribution at all. On the other hand, we want to determine whether the difference in the distributions between two samples with different user states is high enough to allow a distinction based on the calculated test statistics. If this is the case, we can subsequently calculate thresholds and resulting confidence values for HASs.

The Kolmogorow-Smirnow Test

The Kolmogorow-Smirnow Test for homogeneity [Kol33] is based on the empirical cumulative distribution functions of the two input samples. Informally speaking, it measures the maximum vertical distance between the two curves. Formally, given two samples $X = [x_1, x_2, \ldots, x_n]$ and $Y = [y_1, y_2, \ldots, y_m]$ with respective empirical cumulative distribution functions F_X and F_Y, it computes the value

$$D = \sup_a |F_X(a) - F_Y(a)| \tag{2.1}$$

Due to the test statistic being the difference between two distribution functions, D takes values within $[0, 1]$. A value of 1 shows that all values in one sample lie below a certain threshold while all values in the other sample lie above this same threshold. The test statistic is 0 if the cumulative distribution functions match exactly.

If the test statistic D is high, the null hypothesis is rejected.

We use the `SciPy`[3] implementation of the KS 2-sample test from `SciPy` version 0.19.1 and apply it to the inter-message time intervals. In addition to the KS statistic D (sometimes referred to as $d_{\max}$ or $D_{a,b}$ in literature), the implementation computes a *p-value* as a function of D and the sample sizes. This accounts for the fact that large samples with the same underlying distribution are expected to show fewer differences than smaller samples (as per the law of large numbers). We examine both the value of D and the *p-value*.

[3] `http://www.scipy.org`, accessed 2021-03-23

The Chi-Square Test

Pearson's Chi-Square Test [Pea92] follows a similar approach as the KS test, but calculates the sum of squared differences between the actually measured frequencies and the expected ones. The test can be performed in two scenarios: One possibility is to test measured frequencies against a (suspected) random distribution function. The other possibility is to test two samples of measured frequencies against each other.

In the two-sample form used here, the expected frequencies are estimated by taking the average frequencies of the two samples. Formally, the test expects categories and respective frequencies as inputs. Given two samples x, y and m categories, these frequencies can be written as $X = [x_1, x_2, \dots x_m]$ and $Y = [y_1, y_2, \dots, y_m]$, where x_i is the number of elements in sample x which fall into the i-th category. Using the intermediate definitions

$$n = n_x + n_y = \sum_{i=1}^{m} x_i + \sum_{i=1}^{m} y_i \tag{2.2}$$

$$\forall z \in \{x, y\} : E_{z,i} = \frac{n_z \times (x_i + y_i)}{n} \tag{2.3}$$

the test statistic is then defined as

$$\chi^2 = \sum_{i=1}^{n} \frac{(x_i - E_{x,i})^2}{(E_{x,i})} + \sum_{i=1}^{n} \frac{(y_i - E_{y,i})^2}{E_{y,i}} \tag{2.4}$$

There is no intuitive interpretation of the test statistic. It can take values between 0 and $\max(|x|, |y|)$, where x and y are the two samples. More generally, the value is within $[0, +\infty[$.

If the value of χ^2 is high, the null hypothesis ("The two samples have the same underlying distribution function.") is rejected. Due to the dependency on the input samples, there is no standardized threshold for rejecting the null hypothesis. Sensible values have to be found separately for every application scenario.

For the Chi-Square Test, we use a custom implementation. Similar to the Kolmogorow-Smirnow Test, it is applied to the inter-message time intervals.

As the test expects the two samples to be categorized into bins, we need to do this before calculating the actual test statistic. Literature suggests choosing bin sizes so that no bin contains less than 5 elements for any sample [FY63]. Thus, we adaptively choose bins of varying size. The lower bound for the first bin is the lowest value in any of the two input samples. The upper bound for a bin (which is also the lower bound for the next bin) is chosen as the smallest number which results in at least

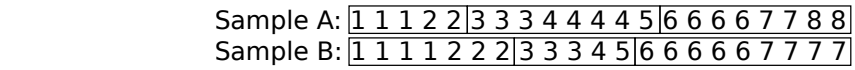

Figure 2.8: *Example of the approach used for binning using a minimum bin size of 5. The bounds are chosen so that at least 5 elements of each sample fall into one bin.*

5 elements of each sample falling into this bin. We thus guarantee that at least 5 values are in each bin for each sample. An example for the binning approach is depicted in Figure 2.8. For the Chi-Square Test we calculate and examine the test statistic.

The Message Counts Test

Our own "Message Counts Test" is a custom test developed for this particular use case. One is added to the number of messages in each sample so that neither number is zero. Then, the higher number is divided by the lower and one is subtracted, resulting in a value within $[0, \max(|x|+1, |y|+1)]$, where x and y are the two samples. More generally, it is within $[0, +\infty[$. Higher values indicate larger differences in the amounts of messages, just as higher results in the other tests indicate different distributions. The idea behind it is that if the sheer amount of activity in the system is very different to that during the reference time frame, the user state is likely to be different. For example, if the reference capture was taken while Alice is present and the capture in question shows significantly lower activity, Alice is likely to be absent.

Formally, given two samples $X = [x_1, x_2, \ldots, x_n]$ and $Y = [y_1, y_2, \ldots, y_m]$, the test statistic is defined as

$$C = \frac{\max(n, m) + 1}{\min(n, m) + 1} - 1 \tag{2.5}$$

Note that this definition differs slightly from the publication [MS16] in order to also work for cases where one sample is empty—i.e. contains at most one message and thus no inter-arrival times.

Similar to the Chi-Square test, we calculate and examine the test statistic.

2.2.2 Test Suitability in the General Case

At first, we plot all test results by system and test and only distinguish between the two cases whether the samples have different user states. This section gives a general and quick overview over the suitability of the tests for our purposes. If the plots of the two cases differ significantly, the test results carry a high amount of information and if they are largely the same, the information immediately available from the test result is limited. The plots are visualized in Figures 2.9 and 2.10. The box plots do not show any immediately obvious peculiarities. For both systems and all tests, the boxes overlap and thus suggest that the tests cannot be used as a universal oracle telling an attacker whether the 2 compared samples have been taken with the same user state. However, the plots do not always fully overlap. We discuss this in the following sections.

Interestingly, both the Message Counts Test for System 1 and the Chi-Square Test for System 3 show counter-intuitive results. The test statistic produces higher maximum values for samples with the same user state than for samples with different user states. Intuitively, the test statistics should be smaller for same-state samples, as they are assumed to follow a similar distribution.

System 1

For System 1, the *Chi-Square Test* values are broadly spread. Comparing samples with the same user state yields values from 0 to 79.64, samples from different states lead to values from 0 to 79.45. This suggests that in general, the Chi-Square Test is unsuitable for determining whether two samples have the same or a different state if the state of both samples is unknown. Intuitively, the test statistic would have lower values for same-state samples, as the samples' underlying distribution functions are expected to be similar or the same. The arithmetic mean, median and quartiles support this assumption, but the maximum values prove that it is not generally the case.

The *Kolmogorow-Smirnow Test* statistic D yields values in the full range $[0, 1]$ for samples with the same state. For samples with different states, however, the minimum value is 0.02, which suggests that there may be a lower bound.

The *Kolmogorow-Smirnow Test p-values* provide similar information. For the same state, the values range from 1.42×10^{-14} to 1. For different states they range from 0 to 1. The null hypothesis ("The two samples originate from the same distribution [=the same state].") is rejected for p-values lower than a threshold. The lower minimum value for different states shows that there might be a bound.

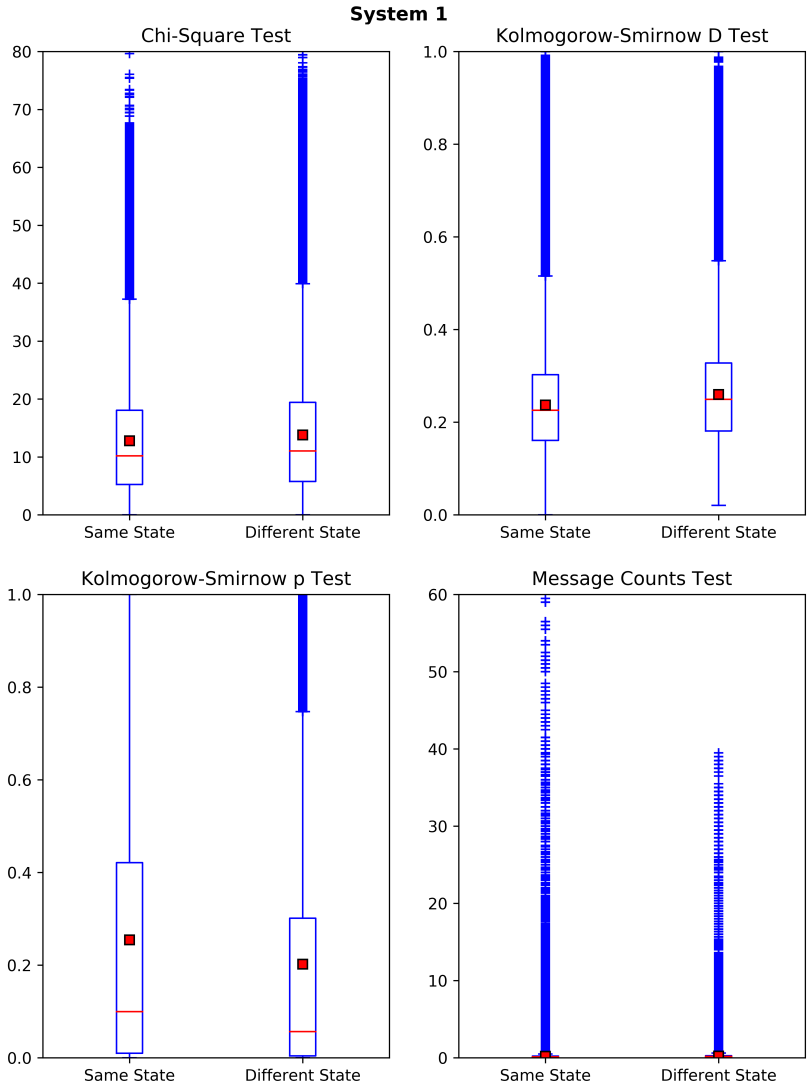

Figure 2.9: *General Test Results for System 1. The boxes extend from the first to the third quartile. The whiskers extend up to* $1.5 \times IQR$ *past the boxes, where* IQR *is the interquartile range. If* $IQR = 0$ *(not the case here), the whiskers extend up to the minimum and maximum values. Red lines mark the medians while red squares mark the arithmetic means. Blue plus signs show outliers beyond the whiskers. Due to the large number of samples and outliers, the plus signs seem to merge into lines and some boxes seem to disappear.*

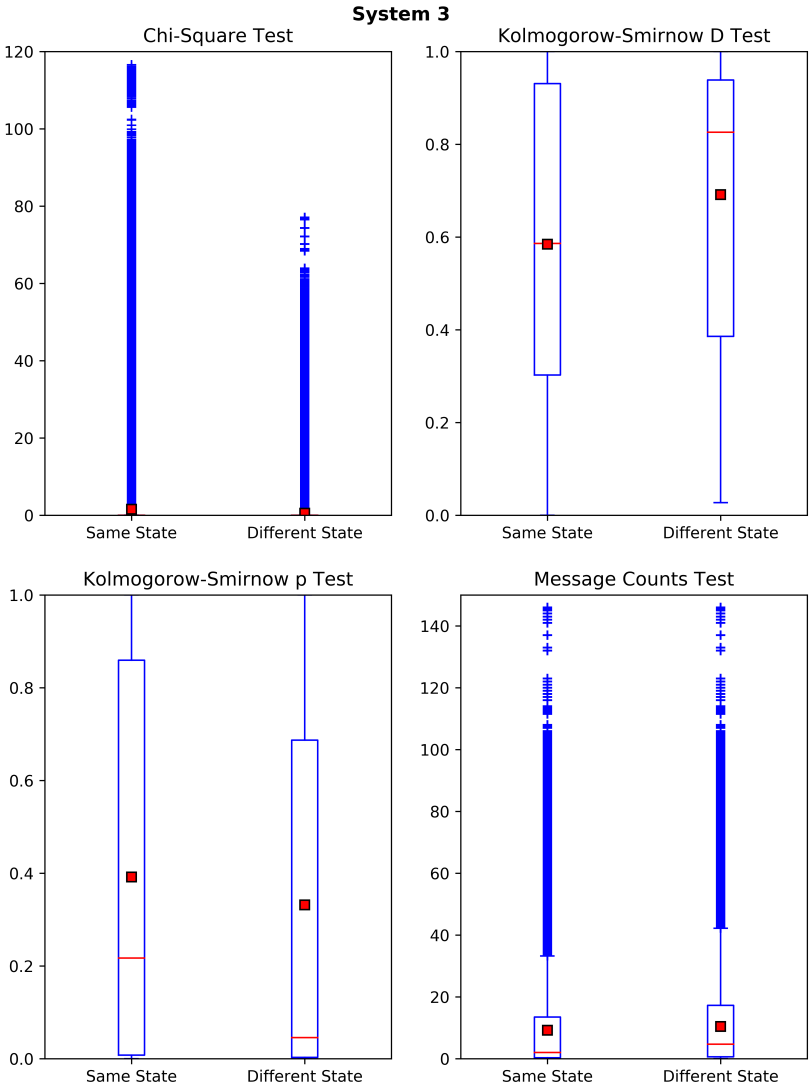

Figure 2.10: *General Test Results for System 3. The plot parameters are the same as for Figure 2.9.*

The *Message Counts Test* provides the least useful results. The values are in fact misleading: While they range from 0 to 59.5 for samples with the same state and the minimum is the same for different states, the maximum value in the latter case is only 39.5. This shows that while the user state does not change, the number of messages being generated in a given time frame can differ significantly.

System 3

The results for System 3 offer much less information than those for System 1. The *Chi-Square Test* values range from 0 to 116.59 for samples with the same state and from 0 to 77.06 for samples with different states. As shown in Figure 2.10, 75% (the lower three quartiles) of the tests with different states had the result 0. These values are misleading if interpreted in the same way as those of System 1. Intuitively, the values should be higher for different states (and they are on average for System 1). We conclude that either the test's usefulness depends on the type of the HAS or that the previous results were not representative.

The *Kolmogorow-Smirnow Test* statistic D yields values in the full range $[0, 1]$ for samples with the same state. Similar to System 1, the minimum value for samples with different states is slightly higher at 0.03. This supports the theory of a threshold indicating different user states in compared samples.

The *Kolmogorow-Smirnow Test p-values* are inconclusive: They range from 0 to 1 for both cases.

The *Message Counts Test* surprisingly yields the exact same minimum and maximum values for both cases: The results range from 0 to 146, whether the two samples have the same user state or not.

2.2.3 Test Suitability Per State Pair

In the next step we take a closer look at the different combinations of user states. Our hypothesis is that the tests may give useful results for certain combinations of states and less useful results for others. This section deals with the performance of the tests for a given pair of user states. Figures 2.11 and 2.12 summarize the results for both systems.

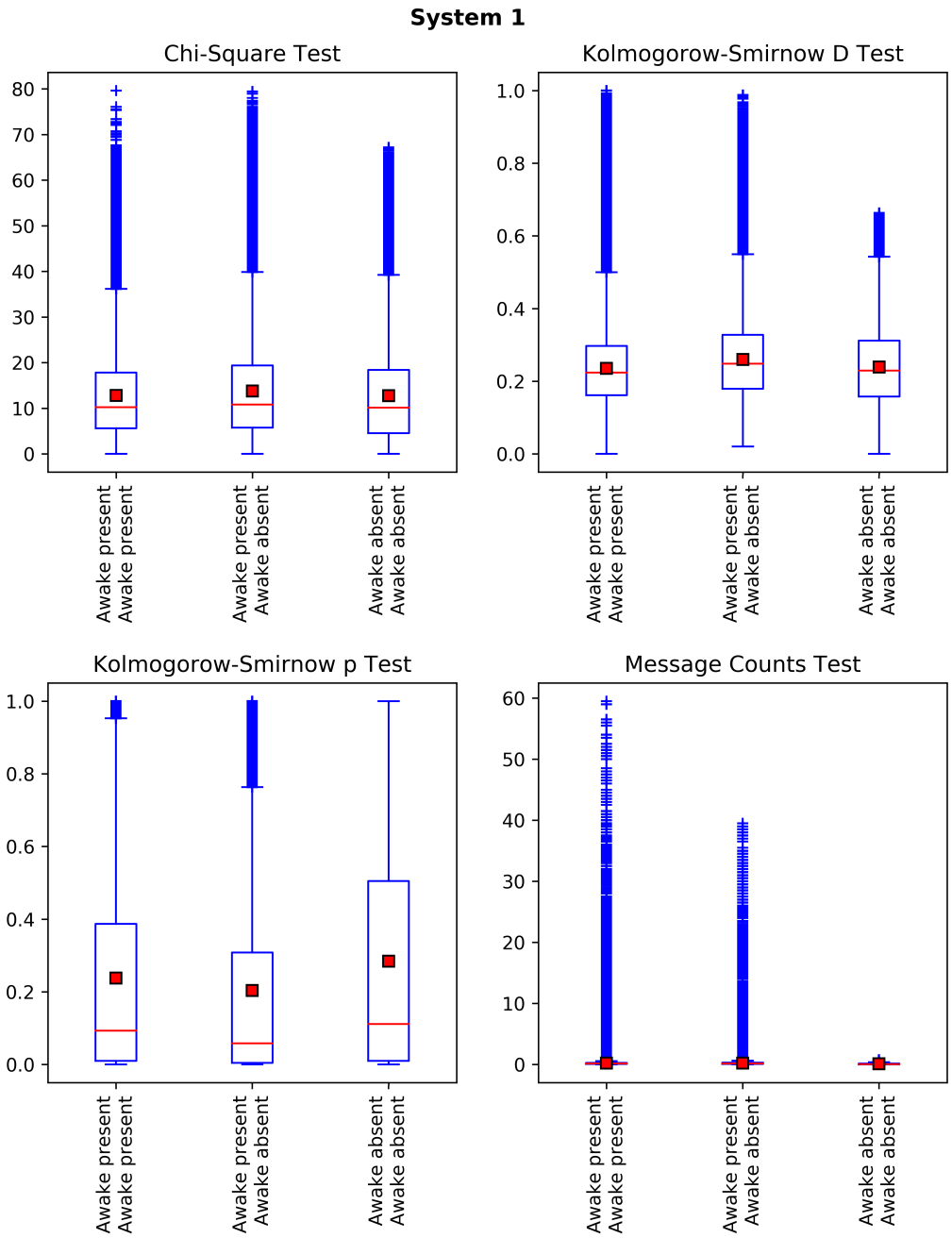

Figure 2.11: *Per state pair test results for System 1. The plot parameters are the same as for Figure 2.9. The combination* Present-Absent *does not appear due to the symmetry of all tests:* $T(a, b) = T(b, a)$.

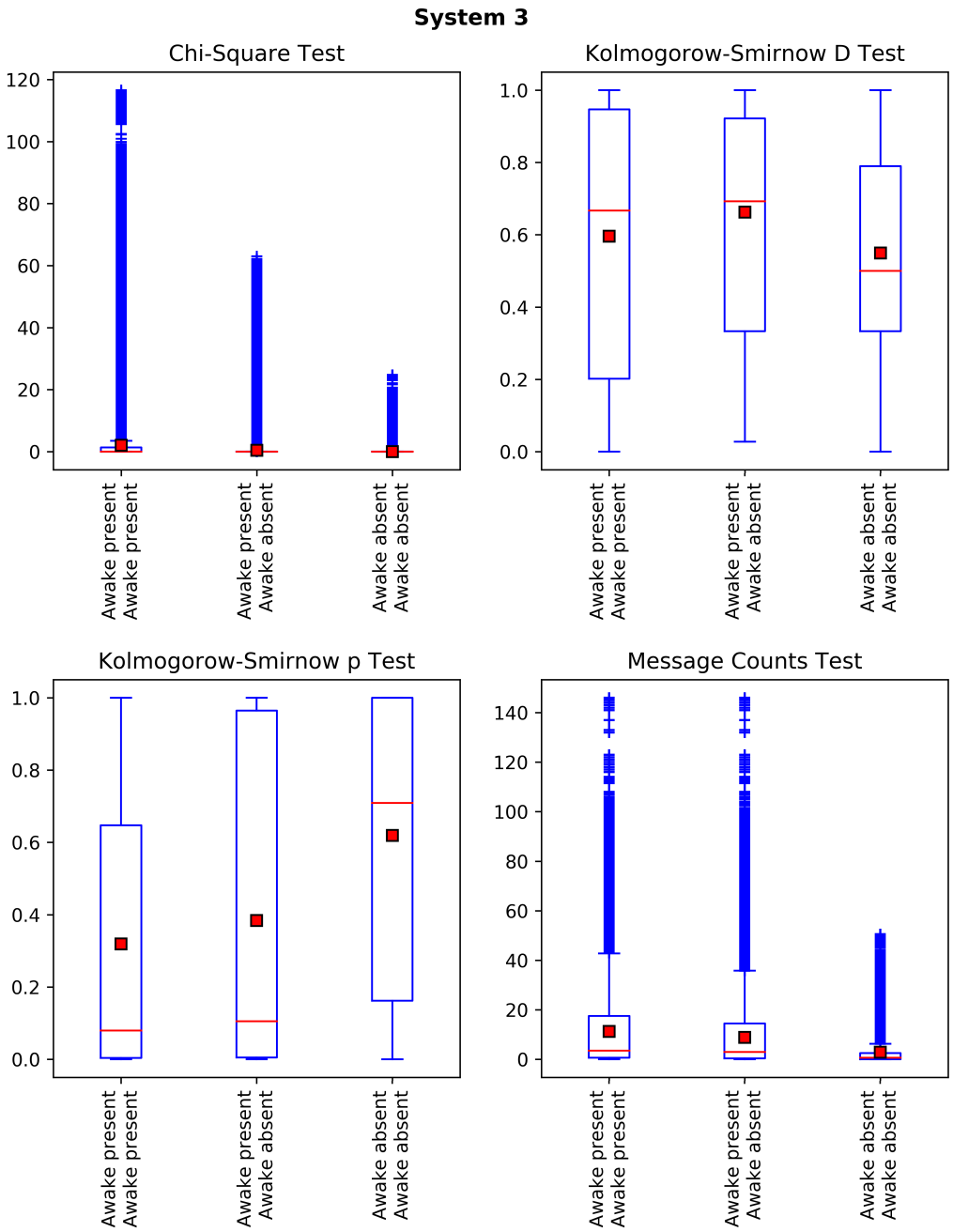

Figure 2.12: *Per state pair test results for System 3. The plot parameters are the same as for Figure 2.9. The combination* Present-Absent *does not appear due to the symmetry of all tests:* $T(a, b) = T(b, a)$.

System 1

The detailed view on the combinations of states provides new insights: 3 of the 4 test statistics exhibit thresholds for specific combinations of user states. This is most prominent for the Message Counts Test. If both samples have the state ABSENT, the values do not go above 0.77. This means that if an attacker obtains a sample known to have the state ABSENT, and gets higher value when comparing it to a second (unknown) sample, they can be sure that the user was present during the time frame of the second sample. However, this is only the case for 1.52 % of the tested ABSENT-PRESENT sample pairs.

Additionally, given a PRESENT source sample, the attacker can be certain that a second sample has the state PRESENT as well if the test statistic is above a certain threshold.

Deductions about a second ABSENT sample cannot be made with absolute certainty using the Chi-Square or Message Counts Test, regardless of the first sample's state. However, the Kolmogorow-Smirnow test statistic D suggests that there is a lower bound to the comparison of samples with different states. This allows for certain identification of ABSENT samples.

System 3

Compared to System 1, the boxplot of the *Chi-Square Test* for System 3 exhibits larger differences between the state pairs. If one of the samples has the state ABSENT, 92.10 % of the tests evaluate to 0. Similarly to the Message Counts Test for System 1, the plots show that there is a threshold above which the attacker can be sure that the user is PRESENT if their first (known) sample has the state ABSENT. This threshold is at 24.78 and 0.17 % of the ABSENT-PRESENT pairs reach a higher value. Similar to System 1, some conclusions can be drawn given a PRESENT source sample as well, but an ABSENT second sample can never be identified with absolute certainty.

The *Kolmogorow-Smirnow Test* exhibits a lower bound for different-state comparisons similar to System 1

The *Message Counts Test* confirms the observation from the Chi-Square Test and yields another threshold. The threshold value is 50.51 and 1.37 % (1,778,053 out of 129,656,028) of the tests with different states result in higher values. Surprisingly, though, none of these 1,778,053 Message Group pairs give a result above the threshold for the Chi-Square Test. In fact, some pairs even evaluate to 0 in the Chi-Square Test. This proves that a combination of different tests with the same input data can provide significantly more information than one test alone.

2.2.4 The Effect of Different Thresholds on Classification Rates

As shown in the previous section, some tests exhibit maximum values for certain state combinations, and knowing such values may enable an attacker to infer the user's state at a given time with absolute confidence. Below these, however, statements about presence and absence are more difficult to make. In this section we examine the effect of different chosen threshold values on the classification rates.

We compute True and False Positive Rates TPR and FPR for all possible threshold levels using the data from the tests previously conducted. In our case, the rates are defined as follows:

If $s(a)$ is the state of a sample a, $T(a, b)$ is the test result of the pair (a, b), t is the threshold value below which sample pairs are classified as having the same state and $N_{a,b}(cond)$ is the number of sample pairs a, b which satisfy a condition $cond$, then

$$TPR = \frac{N_{a,b}(s(a) = s(b) \wedge T(a, b) < t)}{N_{a,b}(s(a) = s(b))} \tag{2.6}$$

$$FPR = \frac{N_{a,b}(s(a) \neq s(b) \wedge T(a, b) < t)}{N_{a,b}(s(a) \neq s(b))} \tag{2.7}$$

TPR is the number of correctly classified same-state pairs divided by the total number of same-state pairs and FPR is the number of different-state-pairs which were incorrectly classified as having the same state divided by the total number of different-state pairs. TPR is a measure for how well the test can identify samples with the same state as the source and FPR is a measure for how often the test falsely reports two samples for having the same state.

In order to visualise the rates, we plot ROC (Receiver Operating Characteristics) curves and calculate the AUC (Area Under Curve) for all of them. ROC curves illustrate how fast the test performance drops (i.e. how fast the False Positive Rate increases) when raising the threshold to get a higher True Positive Rate. The AUC is a numerical measure for this quality: In the ideal case (the test has a TPR of 1.0 and a FPR of 0.0) the value is 1 and in the worst case (the test does not perform better than randomly guessing), the value is 0.5. Values below 0.5 are similar to values above, since the test result interpretation can be inverted to invert the ROC curve (i.e. values *above* the threshold are interpreted as indicators for a same-state pair).

The ROC curves are depicted in Figures 2.13 through 2.16 for System 1 and Figures 2.17 through 2.20 for System 3. Some tests (most notably the Chi-Square Test for System 3) yield high values for both rates with the lowest possible threshold, which is why the curves do not start at the origin $[0, 0]$. To calculate the AUC for

these cases, we use the *line of no-discrimination*—the values obtained by randomly guessing—up to the FPR of the lowest threshold (the X coordinate). From there on, we proceed with the regular estimation and calculate the area below the straight line between two subsequent data points.

Most curves do not exhibit large deviations from the mean line. For System 1, both the Chi-Square Test and the two Kolmogorow-Smirnow Tests yield an AUC between 0.52 and 0.57. Only the Message Counts Test performs slightly better, the AUC is 0.525 for a source sample with state PRESENT and 0.688 for an ABSENT source sample (shown in Figure 2.16).

Overall, the results for System 1 suggest that statistical tests are only of limited use in deducing user states from inter-message intervals.

System 3 mostly confirms this observation, although the performance of the different tests varies drastically.

The Chi-Square Test performs badly: For a PRESENT source sample, the minimum obtainable False Positive Rate is 91.6 % at a True Positive Rate of 61.3 % (the threshold value in this case is 0). For an ABSENT source sample, the minimum False Positive Rate is consequently the same, but the minimum True Positive Rate is 98.0%. The Kolmogorow-Smirnow Test and the Message Counts Test perform much better. Their AUC values are relatively high and significant True Positive Rates can be obtained while keeping the False Positive Rates below 50%.

From the analysis of the ROC curves we draw two conclusions. Firstly some tests exhibit a significant deviation from the line of no-discrimination. Combining multiple tests could further improve the results and yield more information. Secondly we can confirm our previous observation that extreme threshold values lead to absolute certainty in the classification.

2.2.5 Feasibility of Detection in Practice

The statistical tests do not yield clear results in all cases we examined. However, upper or lower bounds can be determined in some cases, which then allow an attacker to make statements with absolute confidence. The requirements for these thresholds to be useful for the attacker are not hard to meet: They need a source sample which—when tested in conjunction with samples of a different state—yields values above or below the thresholds.

To verify the practicability of this attack we divide our traffic data into a training set and a test set. For training, we use the first 70% of our data (13,338 Message Groups from System 1, 16,586 Message Groups from System 3).

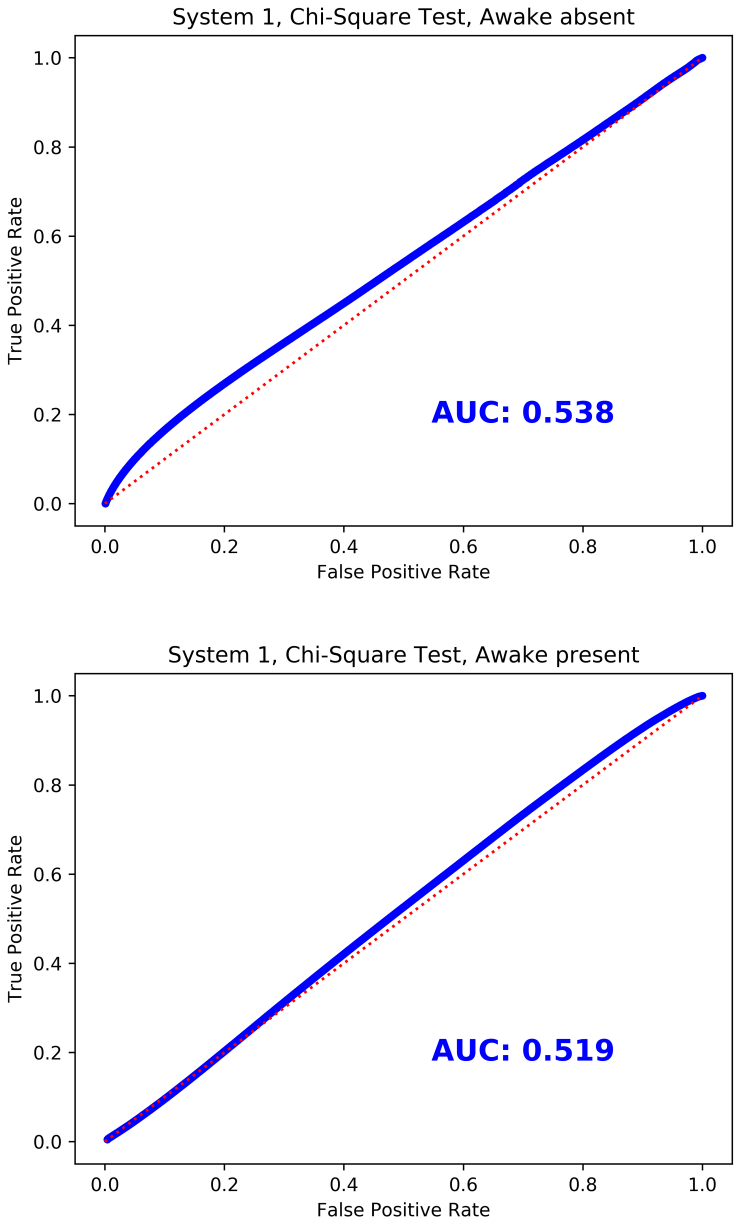

Figure 2.13: *ROC curves for the Chi-Square Test on System 1. Blue points show the actual values, dotted red* lines of no-discrimination *show linear ascension from* $[0, 0]$ *to* $[1, 1]$*—the values obtained by randomly guessing.*

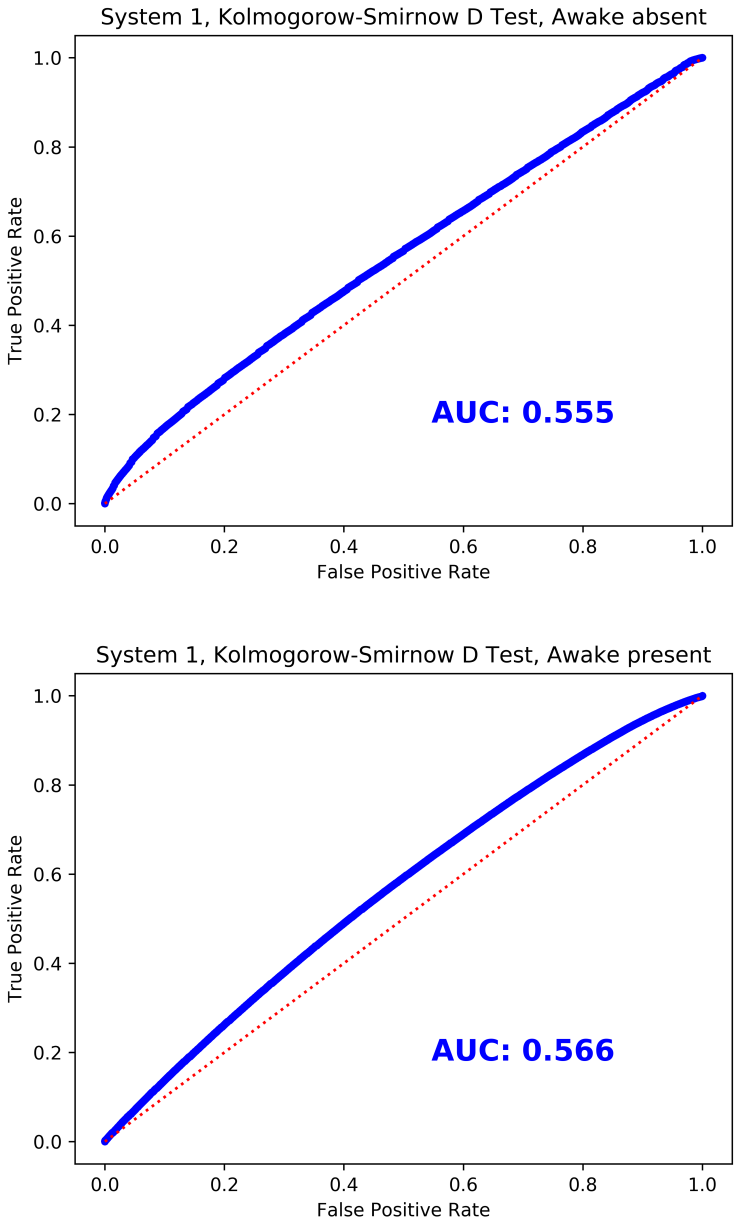

Figure 2.14: *ROC curves for the Kolmogorow-Smirnow Test statistic D on System 1. The plot parameters are the same as for Figure 2.13.*

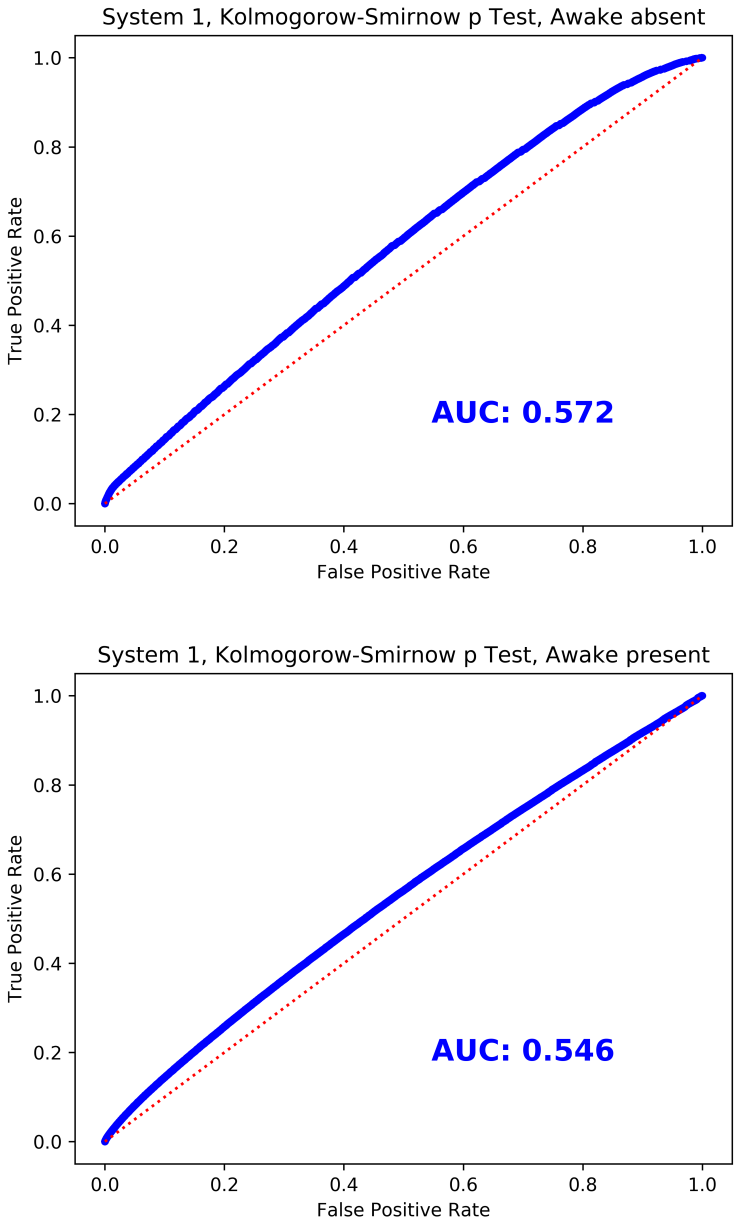

Figure 2.15: *ROC curves for the Kolmogorow-Smirnow Test p-value on System 1. The plot parameters are the same as for Figure 2.13.*

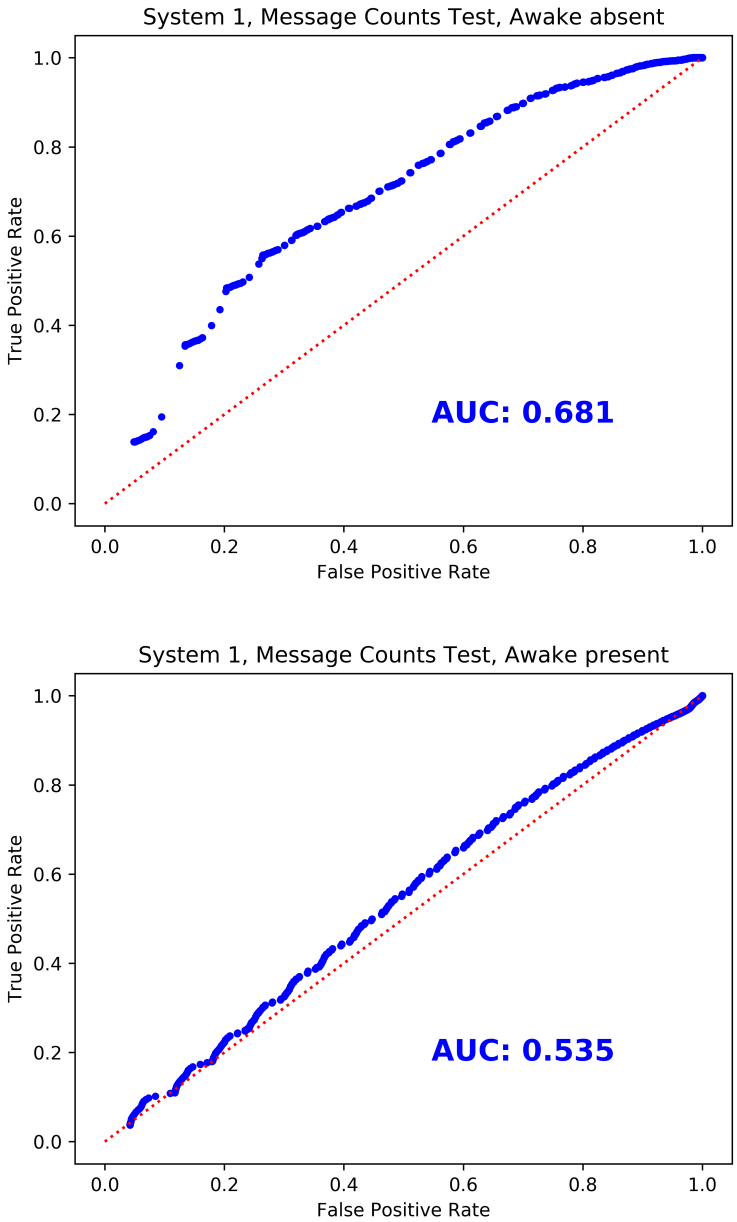

Figure 2.16: *ROC curves for the Message Counts Test on System 1. The plot parameters are the same as for Figure 2.13.*

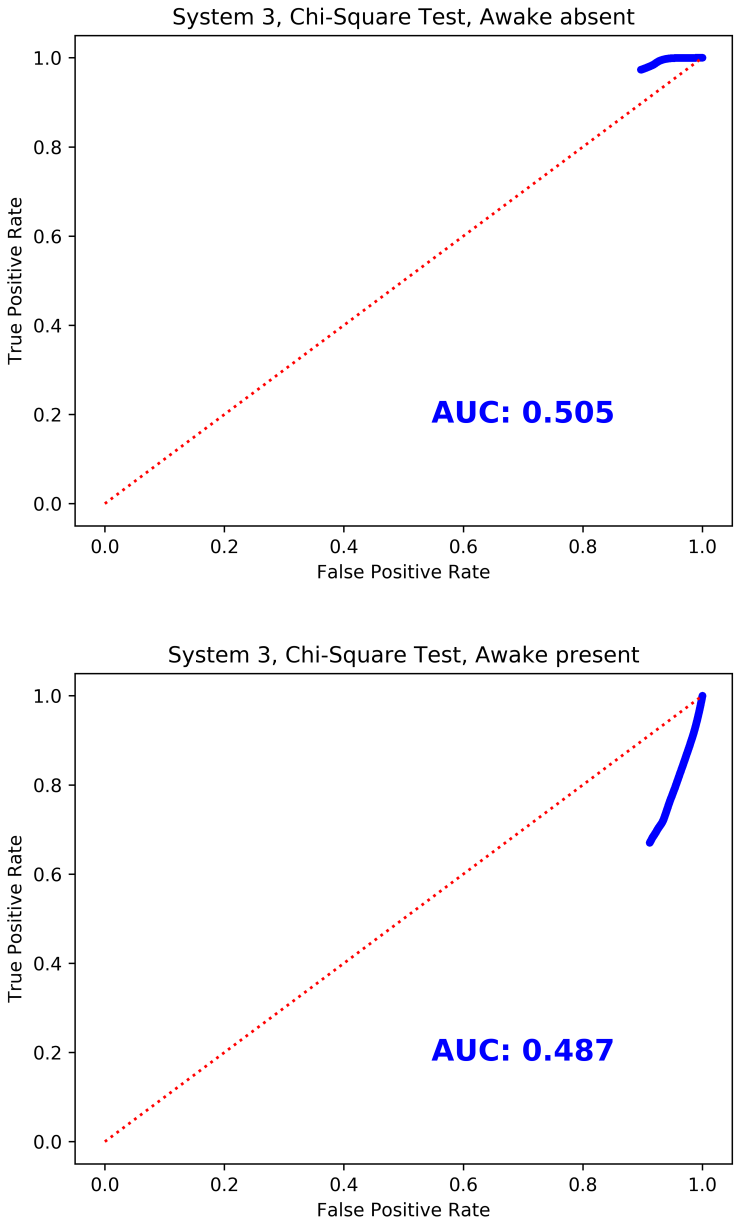

Figure 2.17: *ROC curves for the Chi-Square Test on System 3. The plot parameters are the same as for Figure 2.13. As noted in Section 2.2.2, the test produces counter-intuitive results for a* PRESENT *source sample.*

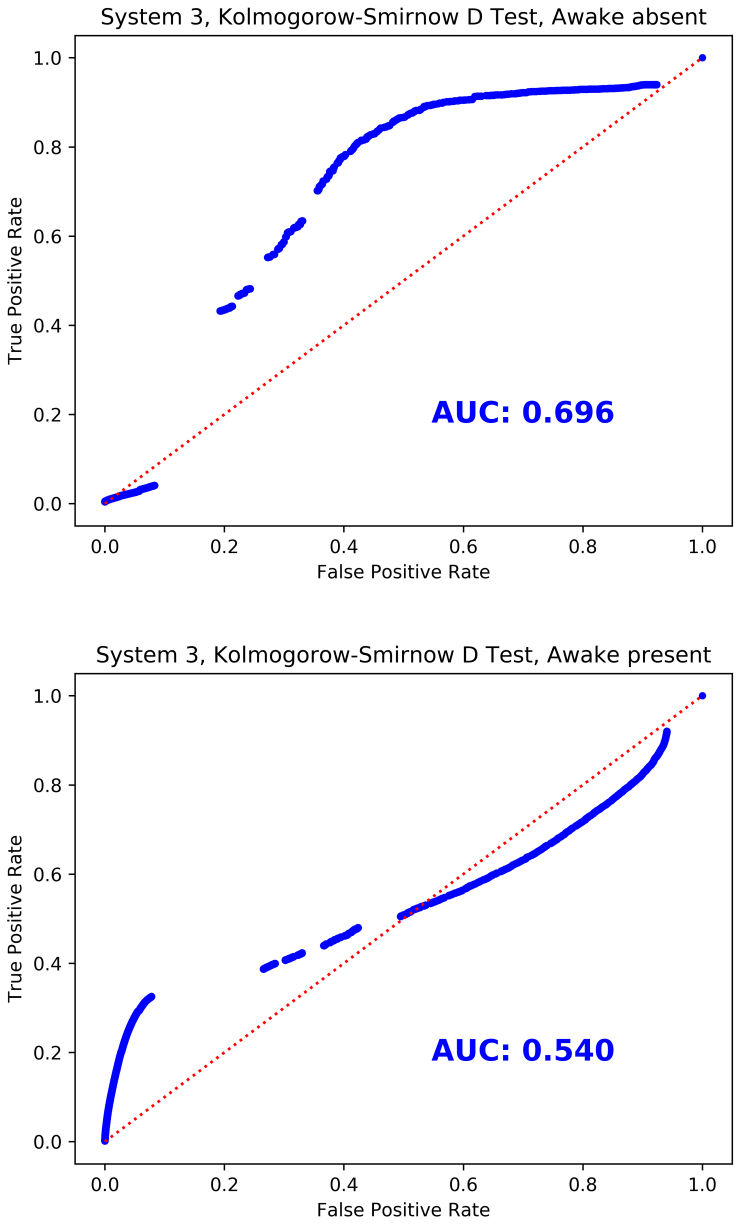

Figure 2.18: *ROC curves for the Kolmogorow-Smirnow Test statistic D on System 3. The plot parameters are the same as for Figure 2.13.*

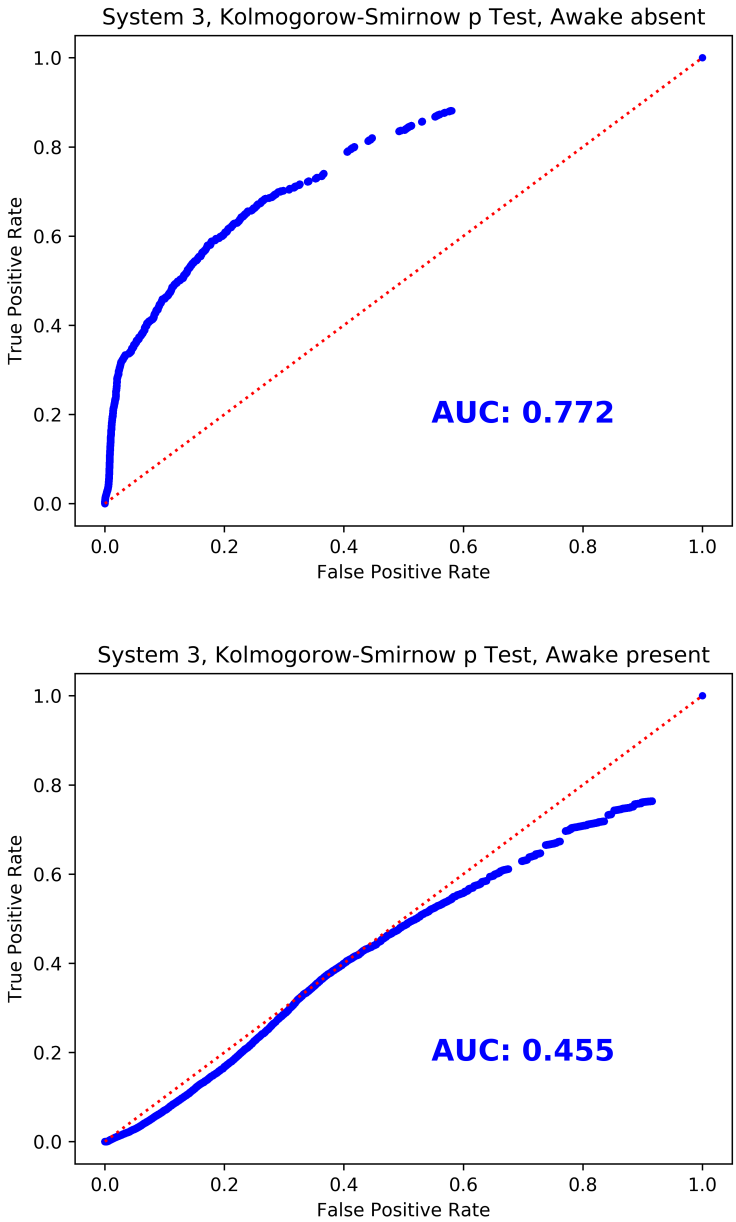

Figure 2.19: *ROC curves for the Kolmogorow-Smirnow Test p-value on System 3. The plot parameters are the same as for Figure 2.13.*

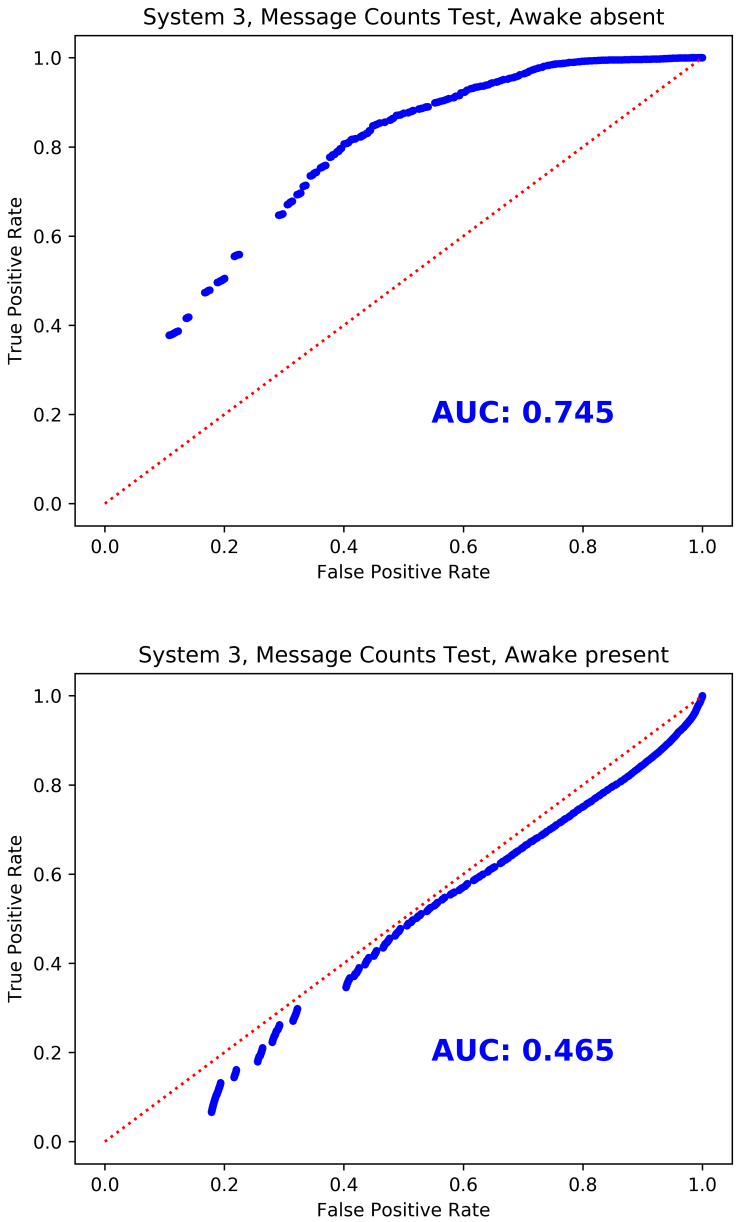

Figure 2.20: *ROC curves for the Message Counts Test on System 3. The plot parameters are the same as for Figure 2.13.*

We perform all aforementioned tests on the training data and calculate thresholds for Message Group pairs with the same state. Using these thresholds, we choose one Message Group for every system and state where the amount of correct classifications among the training data is maximized—i.e. the Group with the highest TPR among the training data. We then check each of these Groups against the test data and calculate True and False Positive Rates using the thresholds calculated from the training data before.

For System 1 using an ABSENT source sample, we reach a TPR of 5.70 % and a FPR of 0.65 %. This suggests that the attack is not useful in practice. No PRESENT source sample reaches a TPR above 0. For System 3, the best ABSENT source sample achieves a TPR of 19.02 % while the FPR is at a low 0.32 %. Similar to System 1, no suitable PRESENT source sample exists. The tests do not yield thresholds which allow for an unanimous classification.

This particular attack is not likely to be encountered in reality: An attacker would have to manually observe the user's home for several hours or even days, annotating the captured traffic with the user states for every one-hour sample. However, the experiment shows that under the right circumstances, unanimous classification is possible. The experiment supports the theory that system-wide thresholds exist which allow for a classification of states with absolute certainty. The follow-up question is whether such thresholds exist for a manufacturer or production series. This could not be confirmed due to a lack of sample data.

2.2.6 Classification Using Machine Learning

As shown in the previous sections, the combination of different tests can significantly improve the results when trying to determine the user state of an unknown sample. In order to take a further step past the individual sample selection and evaluation from the previous section, we use the calculated test statistics as input for machine learning classifiers. This section extends beyond the scope of the previous publication [MS16]. The results have not been published as they do not provide significant new insights to the problem.

The previous tests can only output individual values or thresholds and thus provide a binary classification. Machine learning algorithms aim to improve this by being able to model non-binary and possibly non-linear classifications. They are therefore suited for improving the results gained by applying the statistical tests.

We use the *Linear Support Vector Classification* as well as the *Linear Classifier based on Support Vector Machines (SVM) using Stochastic Gradient Descent (SGD) training* from the `scikit-learn` library [Ped+11] with their default parameters.

Both are applicable to this problem class and are suited for the amount of data available.

Both classifiers are trained on the same input. Again, we use the first 70% of our data as training data. Each training sample consists of a set of features as well as the target class. As features we use a numerical representation of the user state of the first sample as well as all 4 test statistics from the aforementioned tests (Chi-Square Test statistic, Kolmogorow-Smirnow Test statistic D, Kolmogorow-Smirnow Test p-value, Message Counts Test statistic). The target is the user state of the second sample. This setup corresponds to our scenario where the attacker wants to learn the user state of an unknown sample based on the comparison with a known sample.

After training the classifiers, they are applied to the test data. Due to the fact that both classifiers work with distances of points to hyperplanes, they can output a confidence value for each tested sample. Based on these, we generate ROC curves which are depicted in Figures 2.21 and 2.22.

For the data from System 1, both classifiers fail to provide significant improvements over the individual test classifications from Section 2.2.4. They only slightly deviate from the line of no-discrimination and the AUC values are below 0.55. They do, however, allow for a finer tuning than the individual tests alone as the curves do not exhibit larger gaps.

For the data from System 3, both classifiers provide reasonable error rates with AUCs just short of 0.7, but fall behind the individual tests in several spots. Most notably, both classifiers fail to reach the extremely high TPR of the Chi-Square Test and Message Counts Test. Furthermore, even the classifiers exhibit larger gaps—a significant difference to the results for System 1. It is unclear why the classifiers fail to reach similar error rates as individual tests even though they combine all test statistics. One possible explanation is that the different test might provide contradicting results for two samples and thus combining them results in higher error rates than using only one test. However, detailed research into the classifiers and their applicability is beyond the scope of this thesis.

The difference between the machine learning classifier performance for the two analysed systems may be attributed to the traffic patterns. The individual tests already perform differently on the two systems. This indicates that samples with the same user state in System 3 are more coherent in terms of inter-arrival times than in System 1.

The machine learning approach to the deduction of user presence is not immediately applicable to the real world. An adversary needs to have a trained classifier which they can then use on the captured known and unknown samples. Whether a single

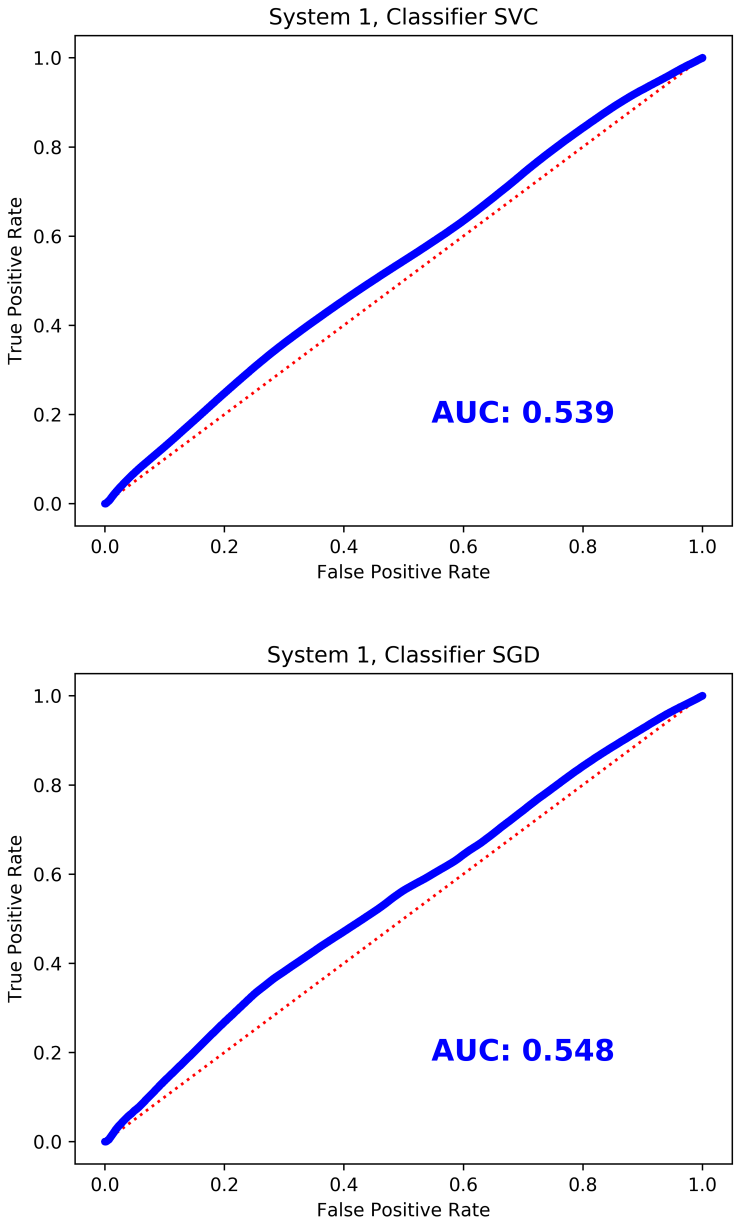

Figure 2.21: *ROC curves for linear classifiers on System 1. The plot parameters are the same as for Figure 2.13. The classifiers do not perform well for this System 1*

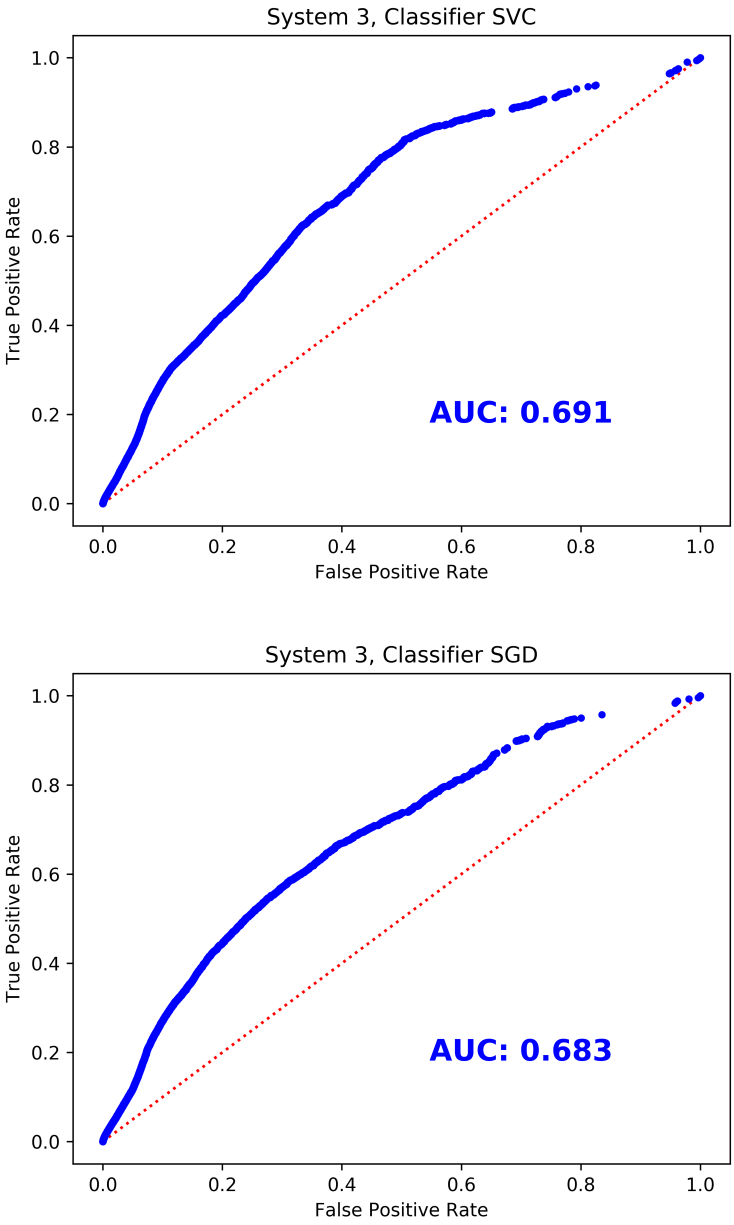

Figure 2.22: *ROC curves for linear classifiers on System 3. The plot parameters are the same as for Figure 2.13. Both classifiers provide reasonable results for this System.*

(trained) classifier can be used for different HASs is unclear. The results however show that a classifier can be applied to the same HAS it was trained on and in some cases provide predictions with reasonable confidence.

2.2.7 Conclusion of Statistical Tests

In this section we have performed an analysis of inter-message intervals in Home Automation using statistical goodness of fit tests as well as machine learning techniques. We have used sample data from two real world installations to measure the ability of an attacker in deducing user states. In particular, we tried to answer the question:

Research Question: *If an attacker has captured 1 hour of traffic from a user's HAS and knows whether the user was present at that time, can the attacker deduce the user's state by capturing another hour of traffic?*

Comparing and combining various tests, we were able to identify conditions under which the question above could be confidently answered with *yes*.

The Chi-Square Test provides little information with regard to the question. However, the Message Counts Test and, in some cases, the Kolmogorow-Smirnow Test reveal identifiable discrepancies between samples with different states. A combination of all three tests allow an attacker to mount a practical attack on the system and infer the user state by passively listening after obtaining a suitable source sample.

Machine learning classifiers do not perform better than simple combinations of individual statistical tests. However, they may in some cases offer more possibilities for parameter tuning.

We note that the demonstrated results show the *minimum* of what attackers are able to learn about HAS users. Using other tests not shown here, even more information might leak from the system. However, the information shown to be leaked by our attack is already sensitive and can pose a danger to the HAS user. Further research into attack vectors is thus outside the scope of this thesis.

2.2.8 Using Encryption

From the attacks on both unencrypted and (virtually) encrypted communication, we can draw two fundamental conclusions:

1. **Encryption alone does not prevent information leakage to passive adversaries.**

We have shown that using statistical tests, an eavesdropper is able to deduce user presence or absence using little to no a priori information.

2. **Encryption does make attacks on the user's privacy significantly harder.**
   The results of the statistical tests show that in some cases, information does leak from the system to the attacker. However, no test and no combination of tests was able to provide the desired information with absolute certainty in all cases. It is clearly visible that the attacks on unencrypted traffic are—if at all—not as easy to perform on a system using encryption.

The consequences of these conclusions are simple: On the one hand, encryption shall be used by all Home Automation Systems to decrease the leak of information. Due to the fact that encryption algorithms and hardware is readily available, it is possible to include it in a HAS with little impact on the user in terms of pricing, battery lifetime and user friendliness. On the other hand, the need for obfuscation techniques still exists. We therefore continue to model traffic analysis and develop an approach for dummy traffic generation in the following chapters.

2.3 Wired Systems

The analyses from the previous sections focused on wireless Home Automation Systems. Due to the broadcast nature of radio transmissions, these systems are especially susceptible to eavesdropping attacks. At first glance, wired systems are safe from sniffing and traffic analysis.

However, Mundt et al. have proven this assumption to be false. [MDG14] Data transmissions in wired systems can be intercepted almost as easily as in wireless systems. Instead of placing an antenna in the vicinity of the system, the authors used the coil of a computer loudspeaker to capture data transmission from a cable by means of magnetic induction. This way they were able to intercept packets from a KNX cable which was located about 10cm into a wall. The monetary cost of the attack is below 200€.

The experiment shows that wired systems are just as susceptible to eavesdropping and traffic analysis attacks as wireless HASs. The monetary costs are about the same and the effort to install a listening device are comparable.

In conclusion, the attacks sketched in the sections above are possible for both wireless and wired systems. Consequently, countermeasures have to be applied in both settings. While the situation at hand is similar, different properties of wired and

wireless HASs might lead to different countermeasures being useful in one case but infeasible in the other. In Chapter 4 show which properties affect the applicability and usefulness of individual countermeasures and how the differences can be leveraged to implement the most effective and efficient countermeasure for each scenario.

2.4 Legal Situation

The research area of privacy and data protection encompasses not only the technical side with its products and problems, but also the legal situation with its own set of issues. Technical measures and legal frameworks are intertwined and both have to be taken into account when developing approaches for the protection of privacy. On the one hand, the law serves to protect people's privacy (or, more generally, their human rights, which include the right to self-determination both in the real and the digital world) by deterring adversaries from violations. On the other hand, technical solutions can to some extent prevent adversaries from violating the users' privacy even if they are willing to break the law.

In this section, we summarize the evolution of the legal situation with regard to both criminal and data protection law. We sketch changes that have been made to keep up to date with technical advancements and to further secure people's privacy. We also highlight problem areas which have either not been dealt with yet legislatively or have appeared as a side effect of the changes made to the legal frameworks.

While we focus on German[4] criminal law in connection with EU regulations and German as well as EU data protection law, problems similar to the ones highlighted here might appear in other legal frameworks as well. Law usually "chases" technological advancements in the sense that legal changes affecting certain technology are generally made after this technology has been invented. Due to this discrepancy, it is common for legal frameworks to not cover all problematic aspects of current technology. The issues covered in this section illustrate this and the legal reforms show how legislators can (and do) cope with advancing technology.

2.4.1 Criminal Law

Research Question: *Does criminal law successfully capture the attacks described here? Or are amendments to the law necessary in order to provide legal protection for users?*

[4]Within the context of this thesis we use the English terminology and recite translated versions of German laws. These translations are not official, so in doubt we refer the reader to the original German texts.

Digital attacks on Home Automation Systems invade the users' privacy at home. In contrast, many legal frameworks aim to protect people's homes as safe spaces not only in terms of privacy. For example, both the German Constitution ("Grundgesetz", Article 13) and the U.S. Bill of Rights (the first ten amendments to the Constitution, specifically the fourth amendment) explicitly state that homes are to be protected from unlawful searches.

The following remarks summarize basic aspects of digital attacks on Home Automation Systems with regard to the German Criminal Code ("Strafgesetzbuch"). They are consolidated from different publications [MV16; KM16; Möl+18] which focus on different aspects of the legal situation regarding Home Automation Systems.

Problems of the Past

As an enforceable law and punishable offense, the German Criminal Code sanctions illegal trespassing in Section 123. However, the prevailing interpretation in legal literature is that this requires the offender to physically enter a spatially delimited area. [Fis16, § 123, Recital 15] Gaining access to or capturing traffic from a Home Automation System thus does not qualify as trespassing according to German criminal law.

Another offense that comes into mind when considering illegal interception of communication traffic is that of theft. The classical definition of theft from Section 242 of the German Criminal Code however also fails to capture this, as it requires a *chattel*—a movable, physical object—to be taken (as per the definition in Section 90 of the German Civil Code). [Vog16] The same holds for damaging or destruction of property as penalised by Section 303 of the German Criminal Code.

While in German civil law, analogies can be used as a tool to cope with missing legal coverage, this is not possible in criminal law. The German criminal law is fundamentally influenced by the *principle of legal certainty*, which is derived from Article 103, Section 2 of the German Constitution in connection with Section 1 of the German Criminal Code. This principle effectively forbids the use of analogies in criminal law. [Fis16, § 1, Recital 2]

Legal Reforms

In order to close these loopholes, the German legislator started to adapt the legal frameworks to the digital age. In 1986, Section 202a (Data Espionage), Section 263a (Computer Fraud), Section 303a (Data Tampering) and Section 303b of the German Criminal Code were introduced and have formed the basis of the so-called "Computer

Criminal Law" ("Computerstrafrecht"). As described in [Möl+18], some laws have been passed or modified to implement European Guidelines—more specifically the Convention of Cybercrime of the Council of Europe, which is also known as the *Budapest Convention*. The Budapest Convention has been opened for signature in 2001 and has since been signed and/or ratified by over 50—European and non-European—countries. It is thus of dogmatic importance for questions regarding computer-related crime and has ramifications on a global scale rather than a national one. This is especially true in the following two areas.

- **Prosecution:** Attacks involving computers ("Cyber Attacks") often reach across borders. National solo efforts regarding the combat, solution and sanctioning of these acts are rarely promising. Instead, a coordinated concept supported by as many countries as possible is necessary.
- **Technical and Legal Terms:** One part of this coordinated concept is a collection of common terms describing specific topics or problem areas. This will be further addressed later in this section.

Along the implementation of the Budapest Convention, Section 202a of the German Criminal Code has been updated and Section 303b of the German Criminal Code has been extended to include Denial-of-Service attacks. Additionally, Section 202b (Data Interception) and Section 202c (Preparation of Data Interception) of the German Criminal Code have been introduced. Section 238, which has been passed in 2007 as well, also features a telecommunication-specific alternative (Paragraph 1, Number 2) and thus targets digital crime.

As mentioned above, the Budapest Convention aims to establish a collection of common terms for specific topics related to digital crime. One unsolved issue is however the definition of *metadata* (called "traffic data" in the Budapest Convention). In Article 2 d), it states that

> "traffic data" means any computer data relating to a communication by means of a computer system, generated by a computer system that formed a part in the chain of communication, indicating the communication's origin, destination, route, time, date, size, duration, or type of underlying service.

Article 33 governs mutual judicial assistance regarding traffic data and Article 34 governs mutual judicial assistance regarding content data.

However, neither the Budapest Convention nor the German Criminal Code clearly state whether content and traffic data (or data and metadata) are to be handled

equally in a legal sense. For example, if an attacker observes and saves metadata of a communication between two people (e.g. the date, time and length of the communication), this does not qualify as Data Interception according to German law. Problems related to the distinction between content and metadata have been encountered in various legal areas and are a topic of current research as well as legislative discussion.[KM16]

Criminal Law and Data Protection

The development of the German criminal law as well as the globally important Budapest Convention reveal a central idea: Criminal law can be considered part of a comprehensive data protection concept which aims to optimise both technical and legal aspects. For this reason, criminal law is and needs to be adapted to current and future developments in technology.

2.4.2 Data Protection Law

Research Question: *Is Home Automation System communication protected by data protection law? If yes, which implications does this have for users and providers?*

Concerning the connection between Home Automation Systems and data protection law, the main issue is whether the law affects the handling (or interception) of data from a Home Automation System by a third party. There are multiple possible scenarios where this question is relevant:

- A provider (possibly, but not necessarily the manufacturer of the HAS) might offer additional services to the user. For this to work, the system needs to submit data to the provider who in turn sends information to the user or directly controls the system to optimise performance and/or efficiency.
- A researcher might want to capture traffic from wireless HASs in order to develop approaches for dummy traffic generation. In order to save time and effort, the capturing shall be performed in secret and without consent.

The following discussion is based on a publication [MS15] which has been written before the Regulation (EU) 2016/679 (General Data Protection Regulation, GDPR) was passed. However, large parts of the definitions and principles of the GDPR have been heavily influenced by the former German Data Protection Law (Bundesdatenschutzgesetz, BDSG) and overlap significantly.

Effective since May 2018, the GDPR now governs the collection and use of personal data in the European Union. It supersedes previous national law such as the Federal Data Protection Act, which is replaced by a new version complying with the GDPR's principles.

Consequently, little research and judicature exists to judge practical ramifications on data gathering and data protection with regard to Home Automation Systems. For this reason, the GDPR is only briefly touched in the context of this thesis. A detailed analysis of the exact changes introduced by the GDPR is outside the scope of this work.

Personal Data

The GDPR only applies to personal data. According to Article 4, Number 1, the term *personal data* "means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person".[5]

The term *personal data* refers to data about a single person as opposed to data about a group of persons. [Gol18, Art. 4, Recital 8] Data such as the temperature within a property or times at which persons enter it first and foremost relate to all persons living or residing in this flat. However, many properties are inhabited by single persons. The Federal Statistical Office of Germany estimates that about 40% of all households in Germany consist of a single person. [Sta13] Furthermore, data about a household consisting of multiple persons can still leak information about individuals: If, for example, data shows that nobody is at home at a given time, this tells something about each individual inhabitant. A HAS with cloud support can thus leak information about the registered owners to the service provider (or adversaries with access to this data). Occasional visits by other people do not fully break the link between the data and the habits of the HAS owner. They merely increase the error rates for the deduction of information. This question has been the topic of discussion in related fields as well: Researchers as well as practitioners regard IP addresses as personal data[VM16] even though the connection can be used by multiple persons.

Other than referring to an individual person, the data also has to refer to that person's personal matters. This is to be interpreted in a broad sense. [Gol18, Art. 4,

[5] In German law, a natural person is a human being (as opposed to a *legal person*, which is either a human being or a legal body such as a company or association).

Recital 6] Article 4, Number 1 explicitly lists location data and data pertaining to the cultural or social identity as examples. Data which can be used to deduce information about e.g. the presence of the inhabitants falls under this definition.

The last requirement of the legal definition is that the affected person has to be specific or determinable. Since customers generally register with the provider of cloud-based HAS services or the manufacturer for additional warranty and support, this requirement is often satisfied. A passive adversary can usually also link the data to a specific person as they know where the data was captured, i.e. they can link the data to a physical address which in turn can be linked to the person/s living there.

As a preliminary conclusion, data from a HAS must usually be considered to be personal data with respect to the GDPR. The only cases where this does not hold are if a provider or adversary gains access to data but does not have any information about the user.

Special Categories of Personal Data

EU data protection law includes dedicated rules for the handling of special categories of personal data. According to Article 9, Paragraph 1 of the GDPR, these special categories comprise a person's racial or ethnic origin, political opinions, religious or philosophical convictions, union membership, health or sex life. At first glance, HAS data is barely related to any of these categories. Only in rare cases can a third party deduce information about e.g. the religious convictions from data about presence and absence: If a user is regularly absent from 20 minutes before until 20 minutes after the weekly mass, they are likely religious. Prolonged presence can also indicate sickness. However, an interpretation broad enough to take all these side cases into account would lead to infeasible legal implications. Almost any data may in some way be loosely related to these categories, making the distinction between "regular" and special categories of personal data obsolete. Researchers thus suggest tying the distinction to the actual intention of the party using the data.[Gol18, Art. 9, Recital 13] Albeit not being an optimal solution (the classification of data can change at runtime, depending on the intentions of the involved parties), this is considered to be closer to the idea behind the regulation than other readings.

In conclusion, data from HAS usually does not qualify as belonging to these special categories.

2.5 (Desired) Security Goals

So far we have highlighted shortcomings and privacy leaks in existing Home Automation Systems. In order to systematically develop improvements, we need to establish security goals which are to be fulfilled by a secure and privacy-preserving HAS. For the definition and interpretation of these security goals, we refer to the terminology established by Pfitzmann and Hansen [PH10] in addition to the three basic security goals established by Voydock and Kent [VK83].

If an attacker learns information about the activities or habits of a HAS user, the privacy of the latter is violated. To keep this information private, the adversary must not be able to determine how and when the user interacts with the system. This means that a privacy-preserving HAS must offer

- confidentiality of the communication contents and
- unobservability of the interaction.

In this section we describe what these security goals mean and how they can be achieved in general. We then present existing security mechanisms to achieve parts of these goals. In the following chapters we describe an approach to solve the remaining issue of preventing leakage of information by traffic analysis.

2.5.1 Confidentiality

The "prevention of [the] release of message contents" [VK83], commonly dubbed as confidentiality, means that an outside observer is unable to learn the contents of the communication between two parties. Achieving confidentiality of message contents is an effective step in ensuring unlinkability (see below) of individual messages. If an attacker does not know the contents of a message, they can only try to link it to another message, a sender or a user by using metadata such as size or timing. Consequently, ensuring confidentiality and preventing information leakage from metadata can be used to offer unlinkability.

Confidentiality of message contents is usually achieved by applying encryption. It can also be indirectly achieved by using steganography: By offering unobservability of the contents [PK01], the attacker is unable to find the information and subsequently cannot learn the contents. However, steganography requires a carrier medium and requires a larger communication overhead in our scenario than encryption. We therefore focus on using encryption in the context of this work.

2.5.2 Unlinkability

In order to effectively hide user behaviour from passive adversaries, messages exchanged between devices in the HAS have to be unlinkable. An observer must not be able to link one message to another in order to prevent the identification of patterns, leading to the disclosure of activity.

Encryption is an obvious first measure to ensure unlinkability. However, there are side channels which may still leak information about the contents or origin of a message even if it is encrypted. The timing may exhibit patterns, e.g. a door lock may require a challenge-response protocol with three messages in quick succession while a window sensor will only send a single message on status changes. Addressing information may enable the adversary to link messages to the same destination device.

An ideal HAS provides perfect unlinkability for messages. While it is possible to achieve, it comes with serious downsides and is generally infeasible for regular use. As we show in Chapter 4, however, it is possible to achieve unlinkability to a certain degree: An attacker can only link messages with limited confidence and the conclusions drawn from this may be false.

2.5.3 Authenticity

While not directly affecting the unlinkability of messages within a HAS, authenticity is an important building block for privacy-preserving Home Automation Systems nevertheless. Sender authenticity ensures that an attacker is unable to impersonate genuine devices and thus cannot intrude the system to actively extract information. Establishing encryption keys by using an authenticated key exchange mechanism ensures confidentiality of the transmitted data. We therefore briefly describe some ideas on how to offer authenticity in a Home Automation network.

2.6 Existing Security Mechanisms

Some of the aforementioned security goals—or parts thereof—can already be achieved using available schemes and mechanisms. In order to refine the goals and limits of this thesis, we consider these existing technologies and describe how they can be applied to strengthen the security of Home Automation System against both active and passive attacks. However, this section is not merely a summary of previous and existing research. Many of the measures have not yet been implemented at all in consumer Home Automation Systems or have been implemented only in few.

It is entirely possible that more effective or more efficient solutions than the ones described hereafter exist. Nevertheless, this thesis does not aim towards improving the existing approaches. Its purpose is to analyse yet unsolved research problems and to provide scientific—rather than purely engineered—approaches and solutions. The goal of this section is to draw the line between scientific research problems and engineering problems.

2.6.1 Pseudonymous Device Addresses

In order to provide unlinkability of individual transmissions, it is necessary to strip them of any identifying information. This also includes addressing information such as persistent, unique device addresses or identifiers.

The most straightforward approach to remove identifying information is to encrypt the addressing information along with the message payload and broadcast the message to all devices in range. While this hides addresses from a passive observer, it also increases the computational effort for a given device to decide whether it is the intended receiver of a message. Depending on whether this increase is significant and depending on the ratio of received messages (the number of messages intended for a device divided by the total number of messages received by this device), it makes sense to slightly modify the approach.

Greenstein et al. have developed an approach to solve this problem and to completely remove all visible identifiers from IEEE 802.11 packets, leaving only the size and the timing. [Gre+08] The mechanism can be adapted for different link layer protocols. It does not introduce significant performance or traffic volume overhead, so the effect on power consumption is assumed to be negligible. Possible limitations of the approach are beyond the scope of this thesis. For this work, the confidentiality guarantees by SlyFi are sufficient to satisfy the assumptions of our model.

2.6.2 Key Distribution

Encrypted communication relies on the distribution of cryptographic keys. Network nodes which are supposed to communicate with each other via an encrypted channel usually agree on a symmetric key. Depending on the use case, the symmetric key can be derived from the nodes' asymmetric public keys via the Diffie-Hellman Protocol [DH76]. Depending on the network topology and the communication scheme, the problem is usually called *Key Distribution* or *Device Pairing*.

A multitude of different approaches exist for the distribution of cryptographic keys in networks. They range from simple pre-shared keys where the common key has to

be entered into each node over public key infrastructures where a central authority certifies public keys for authorised devices to completely decentralised protocols using reputation-based trust models.

For use in a HAS, not all approaches are similarly feasible. For example, entering a pre-shared key into each device can be cumbersome if new nodes are added or the key needs to be changed on a regular basis. Furthermore, including an interface for manual key entry into a Home Automation device might require more space and cost than a manufacturer or customer is willing to accept. A relatively complex public key infrastructure or a reputation-based system on the other hand are harder to implement and usually requires more interaction to include a new device.

Some approaches have found widespread acceptance and have been implemented in popular software.

Bluetooth Secure Simple Pairing features different *association models* (pairing mechanisms) which provide different levels of protection against an active attacker present during the connection setup. The so-called *Just Works* model requires the user to take a "leap of faith" and perform the pairing without any authentication of the devices. This method is also often found in use-cases of the software OpenSSH[6], if the user has no second channel to verify the target's public key during the first connection. If combined with a communication channel which only offers a limited transmission range, the assumption that no adversary is in range during the initial setup can be realistic for the majority of uses. An example of a suitable protocol for this is NFC [SSP].

IEEE 802.11 (commonly known as Wi-Fi) networks often use a pre-shared key to secure the communication. [WiFi] This key must be entered into each device in order for it to join the network.

Researchers have also proposed key distribution and exchange schemes specifically targeting low-cost embedded systems with limited sensors and input options.

Jun et al. have proposed a pairing mechanism which leverages the ability of a human to detect light and sound interference. [Han+14] Their goal is to remove the possibility of an active attacker inserting data into the key exchange. While originally developed for pairing smartphones with cars, the approach can be adapted to pair HAS devices. For the transfer via light signals, directed beams can be used within the user's home. For an active attack to be successful, the attacker would then have to place their transmitter either within the property or next to a window while directing it at the target device. This attempt can be easily spotted and requires tedious work for every single HAS installation. The transfer via sound signals can be applied as-is. A house or apartment is usually even safer from sound penetration

[6] https://www.openssh.com

than a car, so the attack is more difficult to launch and exhibits a higher risk of detection.

In conclusion, many approaches exist to exchange encryption keys between devices of a Home Automation System. The question which approach is the most suitable strongly depends on the communication protocol (e.g. Wi-Fi or ZigBee) and the physical characteristics of the devices (the amount and type of sensors or input options available). Choosing one or multiple mechanisms requires an analysis of feasibility, cost and usability. Since this is outside the scope of this thesis, we do not further compare the approaches.

2.7 Chapter Conclusion

In this chapter we have highlighted significant problems of privacy in Home Automation Systems. We have shown that systems without encryption allow adversaries to learn highly sensitive details about the inhabitants. We have proceeded to demonstrate that applying encryption alone does not protect the users completely from passive attacks. Traffic analysis using statistical tests can leak presence and possibly other information to outside observers.

In the second part of this chapter we have given an overview of the past and present legal situation regarding HAS data. We have highlighted problems in the legal frameworks and attempts of legislators to catch up with rapidly evolving technology. As we have exposed, laws alone cannot protect the privacy of HAS users. We therefore constitute that technical solutions must be found in order to protect against the attacks described here.

Chapter 3

A Privacy Model for Home Automation Systems

In order to systematically develop approaches for privacy in Home Automation Systems, we need to formalize the results from the previous chapter in terms of a model. This model can then be used to formally describe desired properties of privacy-preserving HASs and new approaches can be evaluated against it.

Research Question: *How can communication and traffic analysis attacks in HASs be modelled in order to develop and compare countermeasures?*

In this chapter we define such a formal model which can be used to develop and implement a secure and privacy-preserving Home Automation System. We first list assumptions that we make about the setting and identify problems which are either already solved or fall into other research areas. Based on the assumptions we model different aspects of Home Automation Systems so that manufacturers or system architects can formally verify the effectiveness and efficiency of their security measures.

3.1 Assumptions

The following list of assumptions has been published together with an initial version of the model. [Möl+18] In a later publication [Möl20], we have further detailed the assumptions and have slightly amended the definitions to correctly account for continuous timestamps.

3.1.1 Network Topology and Forwarding

As described in Section 1.1.1, there are different possible network topologies a HAS can be built upon. Most systems form mesh networks where devices can communicate directly with each other. Some systems also use a star topology, where communication either takes place between a peripheral device and the base station or is relayed via the latter.

We assume that the network graph is a clique with respect to intended communication. This means that no forwarding is necessary for nodes to communicate. The reason behind this assumption is that forwarding introduces a set of problems, but also opportunities which are already well understood. Results from research on Wireless Sensor Networks can be used to hide both the sender and the receiver of a data packet from an outside observer, leaking no information. The problem of information leakage from the path of messages through the network is best considered separately.

Furthermore, star topology networks or other forwarding architectures can easily be fit into our model: Disconnecting the transfer to and from the base station yields two transmissions which can be viewed as separate packets.

Temporal links between messages can appear even in networks without forwarding. For example, a user might have programmed their (mesh-network) HAS to switch on the light and turn on the heating whenever a door is opened. Thus, transmissions in close succession can happen in both mesh and star topology networks. Their commonness depends on the automation rules and user habits in addition to the network topology.

3.1.2 Encryption and Padding

We assume that the HAS uses strong encryption and padding for all links between two communicating nodes. Among other properties (which are not relevant for our model), these measures provide confidentiality of message contents and prevent classification of messages based on their contents. While side cases such as the initial key exchange between two devices are non-trivial—especially if the HAS is to be used by users with little to no technical experience—the encryption itself is not considered to be a research problem in this work. There are numerous approaches to provide confidentiality of message contents in systems with low computational power and limited power supply. [Ban+15, p. 1–2] We stress that messages must be re-encrypted in case forwarding does happen in order to prevent the adversary from being able to match the incoming and outgoing message. If decrypted message contents shall not be available at intermediate nodes, layered encryption can be used

without invalidating our model.

Padding messages to a uniform length is a compromise. While fixing some or all packets to the same length inevitably increases the communication volume and thus negatively affects battery life, there is no alternative to achieve indistinguishability. If the communication overhead is too high, graduated approaches can be investigated such as the division of traffic into a fixed number of classes and padding messages within one class to the same length. This sacrifices some amount of privacy, as an attacker is able to distinguish between the different classes. Another possibility is to pad messages to random lengths. This way the attacker can only identify message classes with a certain confidence. For these approaches to be effective, a detailed analysis taking the distribution of messages from the different classes into consideration has to be performed. This however only divides the problem into achieving indistinguishability within each class. The approach developed in the context of this thesis is more general and can be adapted to handle messages of different lengths. For the sake of understanding, we therefore do not consider variably padded packets within this work.

Last, we assume that message headers such as the intended receiver are either encrypted together with the message payload or are otherwise hidden from outside observers. This can be achieved by a mechanism such as SlyFy [Gre+08]. While this is a strong assumption and certainly not true for all available HAS products, we keep it for the sake of simplicity.

In the Section 3.2 we dive deeper into the assumptions on encryption and padding and examine the effects on our evaluation should they not hold.

Effectively, these assumptions guarantee that message payloads do not leak information to an observer.

3.1.3 Attacker's Mode of Operation

The attacker in our scenario is passive. Their goal is to learn information about the HAS's user by passively capturing traffic and analysing it. No active attacks such as traffic injection, node compromise or denial of service are launched.

This assumption is based on the feasibility of the attack. Traffic injection does not lead to a reaction by the system if encryption and authentication are correctly implemented. Thus, the attacker does not gain anything from using it. Node compromise requires either the presence of vulnerabilities in the nodes' software or physical intervention by the attacker. While the former cannot be accounted for in a theoretical approach, the latter significantly increases the effort the attacker has to expend. A large-scale attack on multiple properties at the same time becomes even more costly

and the attack would be less rewarding than personal observation. While denial of service attacks (e.g. by sending junk data on the communication channel using a strong transmitter) are possible, they offer no benefit for the attacker. Either the system will not react at all (providing no information for the attacker) or it will queue unsent messages (both dummy and genuine) to be transmitted after the attacks ends.

Furthermore, active attacks can be observable by the victim and countermeasures can be applied (e.g. searching for the attacker's transmitter).

3.1.4 Attacker's Reception

We assume that the attacker has perfect reception as well as an accurate clock. This means that they are able to capture all traffic being transmitted by any node in the HAS. They do not suffer from reception errors and can save the time at which a message was captured. Using available low-cost hardware, these properties are easy to fulfil. Both experiments on real-world installations of HASs [Möl+14; MDG14] have proven this. While reception errors cannot be fully eliminated, their rate in the experiments was low enough to have no significant impact on the feasibility of attacks.

3.1.5 Wireless Device Fingerprinting

We assume that the attacker does not (or is not able to) launch triangulation or device fingerprinting attacks. While these are possible for static transceivers (and most HAS devices usually do not move) in wireless settings, they require a considerable amount of effort from the attacker. [Bag+14] Furthermore, countermeasures to device fingerprinting are a separate area of research. While they complement the results of our work, they are not investigated in this thesis.

3.1.6 Attacker's Awareness and Knowledge

We assume that the attacker is aware of the presence of any countermeasures. They also know the underlying ideas (e.g. algorithms) of these countermeasures, but do not know runtime information of the system (e.g. the internal state of PRNGs). Assuming the attacker does not know the algorithms could be easily invalidated by anyone acquiring the same HAS and reverse engineering the algorithms.

We make no assumptions about the a priori information available to the attacker. Instead, privacy goals are modelled with respect to a given set of tasks that the user

might perform. This allows the model to be applied to a wide variety of systems and scenarios. Reasonable a priori information can differ vastly depending on the setting, so any assumption only reduces the utility of the model. By stating privacy goals with regard to certain a priori information (in the form of tasks), we can develop viable approaches for a given use case and extend them whenever there are new findings without having to change the underlying model.

3.1.7 Transmission Errors

Especially in wireless networks the possibility of transmission or reception errors in non-negligible. However, we argue that the existence of transmission and reception errors as well as the subsequent retransmission of packets does not necessarily leak information about the identity, position or behaviour of HAS devices.

If the approach described in Section 2.6.1 is implemented, each message to the same destination device carries completely new addresses which are unlinkable to those before. Furthermore, devices use sliding windows when determining whether to accept a message to a given address. This means that if a packet is not acknowledged and has to be retransmitted, the retransmission can be sent to a different address and will still be received by the intended target. Furthermore, the payload of the packet will usually include a sequence number to thwart replay attacks (among other uses such as measuring link quality). Having this sequence number at the beginning of the packet payload and using a suitable encryption algorithm as well as mode of operation (such as AES in cipher block chaining mode) ensures that the encrypted payload is different for every message. This means that if message payloads appear random and uncorrelated to outside observers before the retransmission, the payload of the retransmission itself will also appear random and uncorrelated to the other messages. For an outside observer, the retransmission then looks like any other regular message and leaks no information.

3.2 Effects of Relaxations

It might be infeasible to pad all messages to the exact same length. While the heartbeat message of a smoke detector only requires a single bit of information to be transmitted, the measurements of a weather sensor can be several bytes long. Also, parts of the communication (such as addressing information or certain message types) might not be encrypted . Furthermore, the assumptions about the attacker might not always hold to their full extent in practice.

We therefore explicitly note the possibility of extending the attributes of a message

beyond our model. For our analysis, we model messages as relative timestamps. One can however annotate each timestamp with a message length, a receiver address, a wireless channel ID or any other information visible to a passive adversary. Instead of scalar message timestamps, messages can then be modeled as vectors whose elements carry the information visible to an adversary.

The principles of our model and analysis still apply and only the numbers will differ in practice, as it is still possible to calculate probabilities of encountering certain message sequences. For the sake of readability, this thesis assumes that only timestamps are visible and thus we model messages as scalar timestamps.

3.3 Dummy Traffic

Apart from encryption and pseudonyms, an effective and efficient way to generate dummy traffic is essential to a secure, privacy-preserving communication scheme for Home Automation networks. There are two main requirements for the dummy traffic generation algorithm: On the one hand, it has to effectively hide the regular (benign) network communication. On the other hand, it has to be energy-efficient as not to put excessive load on the relatively weak power supplies found in HA devices.

The idea of generating dummy traffic in order to hide genuine traffic is not new per se. As early as 1978 researchers working on security and end-to-end encryption determine that side-channel attacks on message timings can be hindered by applying dummy traffic. [PSK78] However, they also point out the limitations of this approach. Chaum [Cha81] also includes dummy messages in his original design of MIX networks to hide traffic from unauthorized observers.

Current anonymity overlay networks like Tor however do not use dummy traffic and explicitly exclude a global adversary from their threat model. [DMS04] A fundamental problem and one of the most important differences to Home Automation Systems is the source of participants: Anonymous computer networks are designed for everyone to be able to join and participate. This enables adversaries to mount active attacks by modifying other participants' traffic. Such *tagging attacks* are hard to counter and have been acknowledged by the designers. [DMS04]

In contrast to WANs or overlay networks for the internet, Home Automation Systems have the advantage of being closed systems with regard to the users and the control over participating devices. All participants are devices installed by the owner (unless other security measures fail and the attacker gains access to a node). Outsiders cannot join the network arbitrarily. However, the size of these systems also enable an attacker to monitor the whole network at once. The threat model must thus consider a global passive adversary.

3.3.1 Knowing When To Stop

The main difficulty of developing a low-latency dummy traffic generation scheme is evaluating the effectiveness of an approach. Given a generation scheme or a generated pattern, we want to know the chances of an attacker learning anything useful about the victim.

The intuitive and desirable method would be to quantify the amount of available information which is useful to an attacker, comparing a genuine traffic pattern to one augmented by dummy traffic. This, however, is problematic: While the characteristic properties of the traffic pattern (message timestamps and inter-message intervals) present an upper bound for the amount of information available, it is impossible to generalize how much of this information is useful to an attacker.

For example, one attacker might only be interested in learning whether the victim is present at a given time, which could be indicated by the overall system activity. Another attacker might want to know when the victim enters the living room, for which different information is necessary. The other way around, two traffic patterns with the exact same characteristics, i.e. messages of the same length at the same times, might appear for different reasons. Increased activity during the day can indicate that a person is at home or that there are more environmental events (lighting conditions, rain etc.) which the system reacts to.

A priori information plays another key role in modelling the general attacker. Since any findings about the victim are based on some kind of a priori knowledge, we would have to quantify this as well. The conclusion that a person is at home during the day is based on the fact that an increase of the system activity by a certain amount indicates presence. Deducing that a person leaves the house at 09:00 instead of entering it requires knowledge about the person's job or daily routine beforehand. Formalizing the infinite number of possible scenarios is an unachievable prerequisite for this approach to work.

A viable alternative, however, is to exclude fixed a priori information from the general model and instead include it as a parameter. This way, general analyses can be done on the model and concrete implementations can be evaluated with regard to fixed parameters—i.e. sets of a priori information. Furthermore, keeping the model general prevents it from failing to capture new attacks. If we were to only model the currently known extent of information leakage, the model would have to be revised as soon as new attacks are developed. Thus, we keep the model as general as possible and include all implementation- and use-case-specific properties as parameters which can be tuned for concrete evaluations.

3.3.2 System Model

The following definitions and parts of the overall model have been published. [Möl+18]

The HAS performs communication by sending packets which are observable by the adversary. Due to encryption, padding and the unavailability of device identification attacks, the attacker cannot learn the contents of packets. The only information available to the attacker is thus the timestamp of a message. Hence, we model the traffic output of the HAS as a *set of message timestamps* (or *fingerprints*) F.

F can further be divided into multiple (possibly empty) subsets:

$$F = R \cup E \cup D \tag{3.1}$$

where

- R is a set of "regular" messages. They are of no particular interest to the attacker other than ruling them out as candidates for other sets. In a real system, this set can comprise automation rules or reactions to environmental events (e.g. temperature changes).
- E is a set of "interesting" events. They can be results of direct user interaction (e.g. pressing a light switch) or anything else that is of particular interest to the attacker. If the attacker successfully identifies an event as belonging to I, they learn e.g. that the user was at home at a given time.
- D is a set of dummy traffic messages. Their only purpose is to make it more difficult for the attacker to identify events.

We argue that these subsets are enough to model passive attacks on HASs. Events from other use cases than the ones mentioned above can be considered as belonging to either group without leaking information. For example, events from remote user interaction (e.g. using an internet gateway) can be treated the same way as events in R as they can appear at random points in time and (unless the adversary can distinguish them from other events in R) leak no information about user presence.

Any of the subsets can be empty. If the HAS consists of a single actuator with a remote control (i.e. no automation rules), then $R = \emptyset$. If it consists of a single temperature sensor periodically sending data to a base station, then $E = \emptyset$.

We assume that the subsets of F are disjoint, i.e. $R \cap E = R \cap D = E \cap D = \emptyset$. If the system supports simultaneous message transmission on different channels, the

timestamps can be annotated with the channel the respective message was transmitted on. This also accounts for the case where the attacker is able to observe both messages. However, we assume that the channel on which a message was transmitted does not leak information to the attacker. If the system does not support multiple concurrent channels, there can be no two messages with the same timestamp. Avoiding message collisions or ensuring that they do not leak information to the attacker (e.g. by including collisions in dummy traffic) is considered to be feasible and therefore not covered in this thesis.

3.3.3 Attacker Model

As previously stated, the attacker is global and passive, and cannot break the encryption of packets. For a given capture interval $[x, y]$ (x and y are timestamps), they observe the output of the HAS $F^{x,y} \subseteq F$ containing all messages with timestamps in $[x, y]$. Formally if $t(m)$ denotes the timestamp of message m,

$$F^{x,y} = \{m \in F | t(m) \geq x \wedge t(m) \leq y\} \tag{3.2}$$

This rule reflects the assumptions that the attacker has perfect reception, but cannot tell different messages apart based on their contents. We define all possible subsets $F^{x,y} \subseteq F$ satisfying

$$x \geq \min_{m \in F}(t(m)) \wedge$$
$$y \leq \max_{m \in F}(t(m)) \wedge$$
$$x \leq y$$

(i.e. all contiguous subsets) as the *Subsequence Set* $\mathbb{S}(F)$. $\mathbb{S}(F)$ thus contains all possible traffic captures an attacker can observe over any time frame.

Note that when modelling a HAS, message timestamps may follow a random distribution. Consequently, $\mathbb{S}(F)$ is not necessarily a set of message sequences (or sets of concrete timestamps). In this case, $F^{x,y}$ could be modelled as a random distribution and $\mathbb{S}(F)$ would consequently be a set of random distributions—one for every possible capture interval.

3.3.4 Privacy Goals

The goal of the attacker is to learn information about the user by analysing the captured traffic. The goal of the system is to make the traffic patterns look the same regardless of how the user interacts with the system; thus keep the attacker from learning anything.

Toledo et al. have presented a model for information leakage in Private Information Retrieval settings. [TDG16] While the topic may not seem close to that of Home Automation System Privacy at first sight, we can apply their definitions to our model to express the attacker's and the system's goals.

In their scenario, users issue queries to database servers, some of which are corrupted and pass their observations to the adversary. The adversary then provides one user with two possible queries, of which the user randomly chooses and executes one using the database servers. The rest of the users—all of which are honest—executes different, attacker-provided queries. The adversary tries to identify which query the user chose based on the information provided by the corrupted database servers and communication metadata.

We do not consider the whole model for Private Information Retrieval, as the majority of it is not applicable to our setting. In particular, we do not have equivalents for [corrupted] databases/servers. We instead focus on the user (U_t in the PIR model) and the adversary (A), which directly correlate with our user and adversary. Furthermore, our Subsequence Set $\mathbb{S}(F)$ relates to the adversarial observation space (Ω): For a given capture time frame $[x, y]$, the observation space $\Omega_{x,y}$ is the corresponding distribution $F^{x,y} \in \mathbb{S}(F)$. An observation ($O$) is then a sample from this random distribution

The queries (Q_i, Q_j) from the PIR model can be transferred to the HAS scenario almost immediately. The attacker provides the user with two *tasks* T_i, T_j (e.g. "Open the front door at 9:00." or "Perform any interaction between 9:00 and 9:15.") of which the user randomly chooses one to execute. The attacker then captures the HAS's traffic (by obtaining a sample from the random distribution $F^{x,y}$) and tries to identify which of the two actions the user performed. Obviously, it only makes sense to look at a capture interval where traffic may be affected by the tasks.

We can thus formulate a notion of privacy in Home Automation Systems using the definition from Toledo et al. [TDG16] Since our concept does not revolve around information retrieval, but instead focuses on privacy against traffic analysis, we give the properties slightly different names. Note that the definition is also similar to that of ε-differential privacy by Dwork et al. [Dwo06] Aside from the constant leakage δ, the idea behind the definition is the same: The ratio of probabilities of two events being the source behind an observation is bound by an exponential term e^{ε}.

Definition 1. *A Home Automation System provides (ε-δ)-private communication if there are constants $\varepsilon \geq 0$ and $\delta < 1$, such that for any possible adversary-provided tasks T_i, T_j and for all possible adversarial observations O (being a particular random sample of the distribution $F^{x,y} \in \mathbb{S}(F)$) we have that*

$$Pr(O|T_i) \leq e^{\varepsilon} \cdot Pr(O|T_j) + \delta$$

This definition assumes that timestamps are discrete. If they are considered to be continuous and the elements $F^{x,y} \in \mathbb{S}(F)$ are thus density functions, we modify the definitions as follows: The conditions must hold not for all adversarial observations O but for all adversarial observations O and all sets of adversarial observations $\mathbb{O}$. In place of the probability $Pr(O|T)$ from the discrete case we use the probability $Pr(O \in \mathbb{O}|T)$. The definition of (ε-δ)-private communication then reads as follows:

Definition 2. *A HAS provides (ε-δ)-private communication if there are constants $\varepsilon \geq 0$ and $0 \leq \delta < 1$, such that for any possible adversary-provided tasks T_i, T_j, for all possible adversarial observations O and for all sets of adversarial observations $\mathbb{O}$ we have that*

$$Pr(O \in \mathbb{O}|T_i) \leq e^{\varepsilon} \cdot Pr(O \in \mathbb{O}|T_j) + \delta$$

As in the original work [TDG16], if $\delta = 0$ we call the stronger property ε-private communication. Note that in contrast to the definition of (ε-δ)-private PIR, we require $\delta < 1$. This only affects some cases where a particular observation is certain for one task and impossible for another. As systems where $\delta > 0$ already leak information to the adversary and are thus not desirable in practice, this additional constraint prevents the definition from being overly broad (i.e. (ε-δ)-private communication with $\delta = 1$ cannot be called (ε-δ)-private communication).

3.3.5 In Practice

In practice, the property of (ε-δ)-private communication is hard to achieve. If the tasks can be arbitrary, the attacker may choose them in a way so that they produce highly recognizable traffic patterns. For example, the tasks "Press the light switch for 10 times in 2 seconds." produces a very distinct traffic pattern which is unlikely to be observed if the task "Do not interact with the system for 10 minutes." is executed.

In order to account for this, we define a slightly weaker property for communication in Home Automation Systems.

Definition 3. *A Home Automation System provides (ε-δ)-indistinguishability for a set of tasks $\mathbb{T}$ if there are non-negative constants ε and δ, such that for all tasks $T_i, T_j \in \mathbb{T}$ and for all possible adversarial observations O (being a particular random sample of the distribution $F^{x,y} \in \mathbb{S}(F)$) we have that*

$$Pr(O|T_i) \leq e^{\varepsilon} \cdot Pr(O|T_j) + \delta$$

As with the definition of (ε-δ)-private communication, (ε-δ)-indistinguishability can be defined for continuous timestamps by using the probabilities $Pr(o \in O|T)$. If $\delta = 0$, we call the stronger property ε-indistinguishability.

The difference between this definition and that of (ε-δ)-private communication is that (ε-δ)-indistinguishability is only defined for a limited set of tasks $\mathbb{T}$. This way, we account for the fact that some tasks are theoretically possible, but unlikely to be encountered in practice. For example, a system might be able to provide unobservability of the user pressing a light switch twice within an interval of 10 minutes by generating dummy traffic and making sure that there are always at least two packets in every 10-minute interval. While this does not fulfill the goal of (ε-δ)-private communication, it covers much of the everyday activity and might be preferable to a system which offers full (ε-δ)-private communication at a higher energy consumption.

When considering real-world attack scenarios like the detection of user presence (cf. Section 2.2), the tasks provided by the attacker follow a particular pattern. Instead of choosing two unrelated tasks, the attacker wants to extract a certain piece of binary information from the captured data (such as "Did the user interact with the system?"). In this case, the tasks T_i and T_j from the definition are complementary: $T_j = \bar{T}_i$ (i.e. if T_i is "Interact with the system.", then $T_j = \bar{T}_i$ is "Do not interact with the system."). Due to this being an important special case of (ε-δ)-indistinguishability, we define the separate property of *(ε-δ)-unobservability*.

First, we define complementary tasks. A task T is a set of possible outputs $T = O_1, O_2, \ldots, O_n$. Given the set of all possible tasks $\mathbb{T}_{all} = \bigcup T$ the complementary task $\overline{T}$ is defined as $\overline{T} := \mathbb{T}_{all} \backslash T$. This means that the complementary task is the set of all outputs that are never generated by performing T but are possibly generated by other tasks.

Definition 4. *A Home Automation System provides (ε-δ)-unobservability of a set of tasks* $\mathbb{T}$ *if it holds that*

$$\forall T \in \mathbb{T} : \bar{T} \in \mathbb{T}$$

and the system provides (ε-δ)-indistinguishability for $\mathbb{T}$*.*

Similar to the previous definitions, if $\delta = 0$ we call the stronger property ε-unobservability.

These definitions fully capture our system and attacker model as well as the attacks shown in real-world experiments [Möl+14; MS16; MDG14]. Building on an existing model of user behaviour, they can be used to prove privacy guarantees of a Home Automation System's dummy traffic generation scheme. The different levels of guarantees can be leveraged to prove a dummy traffic generator's suitability for any given scenario.

3.3.6 Examples

For trivial approaches it is easy to see whether they fulfil the privacy goals formulated in Secion 3.3.4.

No Dummy Traffic

In Section 2.2 have analysed a system which does not produce dummy traffic at all. We have shown that the system does not offer ε-unobservability for the tasks "Interact with the system during a one-hour period." and "Do not interact with the system for one hour." if the attacker knows certain thresholds beforehand or can determine them based on other a priori knowledge.

In our experiment, the attacker was able to determine conditions which, if met by the adversarial observation O, would reliably indicate user activity or inactivity. If the predicates $P(O)$ and $A(O)$ denote these conditions, then

$$\forall O : P(O) \Rightarrow$$

$$Pr(O|\text{“Interact with the system”}) > 0 \ \wedge Pr(O|\text{“Do not interact”}) = 0$$

$$\forall O : A(O) \Rightarrow$$

$$Pr(O|\text{“Interact with the system”}) = 0 \ \wedge Pr(O|\text{“Do not interact”}) > 0$$

As there is no constant ε satisfying $x \leq e^{\varepsilon} \cdot 0$ with $x > 0$, the system does not offer ε-unobservability. Consequently, the system also does not offer ε-indistinguishability for these tasks and in general, does not offer ε-private communication.

In our experiment, the probability of obtaining an adversarial observation meeting the condition if the user performed the given task was less than 1. Thus, it may be that the system offers (ε-δ)-unobservability.

Constant-Rate (Dummy) Traffic

As a contrast, we analyse the concept of *Constant Rate (Dummy) Traffic* (CRDT). We assume that the system is generating dummy traffic whenever there are no genuine messages to send and that it does not generate dummy traffic when there is a genuine message to be transmitted. Furthermore, time is divided into slots of fixed length. At the end of every timeslot, either a genuine or a dummy message is transmitted. Genuine messages are delayed until the end of the next free timeslot.

The technical details of implementing such a system are outside the scope of this thesis.

Formally, if S is the set containing the ending time of each timeslot and messages in R and E are delayed so that they only occur at the end of a timeslot ($R \subseteq S$, $E \subseteq S$), then dummy traffic is generated by the system so that $D = S \setminus (R \cup E)$.

By construction, the output of the system $F = R \cup E \cup D = S$ is exactly the same, no matter how the timestamps of genuine messages in R and E are distributed. Thus, for any interval $[x, y]$ the adversarial observation will be $O = S \cap [x, y]$, which is stochastically independent of the distributions of R and E. Consequently, for any task T to be executed by the user, it holds that $Pr(O|T) = Pr(O) = 1$. In conclusion, a system using constant-rate traffic provides (ε-δ)-private communication with $\varepsilon = \delta = 0$ (or (0-0)-private communication).

In practice, using CRDT may pose a problem. In order to keep the delay for user interaction reasonably low, the overall traffic rate must be very high (determined by the largest tolerable delay). However, this can lead to the system violating regulatory thresholds about wireless transmissions or draining the battery of connected devices. In wired systems, this is generally not an issue. In a wireless setting, other approaches which minimize the generated amount of traffic have to be evaluated. This—along with a comprehensive analysis of the practicability of CRDT—is done in Chapter 4. The chapter also features the proof that systems not using CRDT and not delaying genuine messages cannot offer (ε-δ)-private communication.

3.4 Chapter Conclusion

In this chapter, we have formalized notions for privacy and privacy guarantees in Home Automation Systems. The model we established is agnostic to user behaviour and can thus be employed to any HAS and any user. The privacy guarantees capture intuitive and desirable properties of privacy-preserving HASs and allow for formal proofs given concrete approaches or systems. Given that the occurence of patterns can be measured in real systems, the guarantees can also be verified using actual communication data.

The main goal of establishing the model is to enable researchers to develop approaches for privacy-preserving HASs and compare them quantitatively. The privacy parameters ε and δ can be compared to the energy consumption and thus form clear and unambiguous metrics. Parametrized approaches can be tuned to the desired values of ε and δ or to a given energy consumption and give the respective other value as a result.

Chapter 4

Privacy-Preserving Communication

In this chapter we build on the previously established model and develop an approach suitable for concealing activity in Home Automation Systems. The main contributions of this chapter have been published. [Möl20] In particular, they are as follows:

1. We provide a quantitative evaluation of dummy traffic generation in HASs with respect to the traffic volume as well as energy consumption overhead. While many related works evaluate the traffic overhead, little is known about the impact on energy efficiency. This is especially important in HASs where most devices are battery powered. We show that traffic overhead is not proportional to an increase in power consumption and highlight how this can be leveraged to achieve privacy at moderate cost in terms of energy consumption.

2. Given the heavy impact of Constant-Rate Dummy Traffic (CRDT) on power consumption, we propose a new dummy traffic generation scheme. By relaxing the targeted privacy goals, we propose a stochastic dummy traffic generation mechanism which allows for easy tuning of privacy versus traffic overhead. We call this scheme Naive Exponential Dummies (NED).

3. We formalize inherent shortcomings of zero-latency dummy traffic generation schemes as well as NED in particular and discuss the important breakpoints in energy efficiency and privacy.

4. We compare the performance of Constant-Rate Dummy Traffic and our stochastic NED scheme using data from real-world HAS installations which include authentic user interactions. We evaluate both approaches with respect to traffic overhead and energy efficiency.

4.1 Modelling Activity

The security goals formulated in Chapter 3 are based on probabilities of traffic patterns occurring in the HAS's output. Calculating these for a given setup can be approached either using a stochastic model of genuine traffic or empirically using previously sampled data.

Establishing a comprehensive stochastic model of genuine traffic in general requires a large, representative body of traffic data from HASs. However, a general stochastic model has the disadvantage of discriminating users whose systems are considered outliers of this model. Their privacy may be violated even though the same measures are applied as in other systems. Therefore, we employ an approach which works on a single system without the need for a general stochastic model. The probabilities and privacy guarantees can be calculated using sampled data from that system alone and dummy traffic can be generated accordingly to provide maximum privacy given the particular distribution of the HAS.

4.2 Data Sets

For the quantitative part of this chapter we use the same data set as in Chapter 2. However, we split the data of Candidate 2 into two subsets to account for the different setups in the two locations where the system is used. We designate these as Systems 2.1 (office) and 2.2 (home), respectively.

To provide a meaningful benchmark and to prove the usability of our method, we have chosen systems which cover typical use cases and setups. Systems 1, 2.2 and 3 are installed in homes and are used by the inhabitants during their daily routine. System 2.1 however provides insights into the performance in extreme situations, as the communication follows a condensed, regular pattern in contrast to the more distributed patterns in Systems 1 and 3. System 2.2 is installed in an office with little to no activity on weekends and an evenly distributed pattern during working hours. All systems are built using commercially available hardware. While we argue that the evaluation demonstrates the applicability of our approach, we stress our consideration from Section 4.1 that HAS usage patterns can differ widely and that the performance for a given HAS can only be reliably computed by using that HAS's own data. Edge cases will likely exhibit different results than those presented in the following sections, but the method for obtaining these results can be applied without modifications.

System 1

System 1's data sample consists of 45,679 messages which were recorded over a timespan of 3,066,575 s (35.49 days, the compensation for receiver outages is incorporated). The individual message timestamps were saved with a precision of 1 s.

The minimum interval between two subsequent messages is 0 seconds. Due to the precision of the timestamps, this merely indicates that multiple messages were transmitted within the same 1-second interval. The maximum interval between two subsequent messages is 3668 s, which means that during a time frame of one hour, no action was performed and thus no message was transmitted. The mean time interval between subsequent messages is 67.14 s; the median is 64 s and the standard deviation is 49.67 s.

System 2.1

The first part of the data sample of System 2 exhibits a data rate more than 3 times as high as that of System 1: 33,708 packets were recorded over a timespan of 720,112 s (8.33 days). Similar to System 1, the message timestamps were measured with a precision of 1 s.

Also similar to System 1, the minimum inter-arrival time between two subsequent messages is 0 seconds. The maximum inter-arrival time, however, is only 153 s. The mean is 21.36 s, the median is 14 s and the standard deviation is 27.49 s. The fact that these values are significantly lower than those for System 1 hint at a more intense usage.

System 2.2

The second part of the data sample of System 2 shows extreme values. Only 999 packets belong to this data set, being captured over a timespan of 1,161,565 s (13.44 days). There are significant gaps in the sequence of message timestamps, which can be explained if we assume that the owner lives alone and is in the office during the day. The minimum inter-arrival time between two subsequent messages is again 0 s, but the maximum is 61,645 s. While at first glance this suggests that outages occurred to the listening device during the capture phase, this is not the case: In fact, a total of 20 gaps are longer than 8 hours.

This is reflected by the statistics: The mean interval is 1163.89 s, although the median is 0 s. This means that at least 50 % of all messages are sent within a single second after their predecessor. The reason for this large spread—the standard

deviation is 5875.77 s—might be attributed to intense usage during user presence. The user is absent during the day, but interacts with the system frequently in the morning and evening, generating transmissions in rapid succession.

System 3

A total of 40,336 messages were captured over a duration of 3,259,153.87 s ($\sim$ 38 days). This places the overall transmission rate close to that of System 1. The minimum inter-arrival time is the same as for the other systems: 0 s. The largest gap between two subsequent messages is 25,050 s, but the mean interarrival time is 80.80 s. Consequently, the standard deviation is 469.68 s and thus the second highest. Interestingly, the median is 0 s, which indicates that similar to System 2.2, a large part of the communication happens in heavily concentrated bursts.

4.3 Constant-Rate Dummy Traffic (CRDT)

In order to establish a baseline against which we can evaluate new approaches, we first investigate the energy efficiency of Constant-Rate Dummy Traffic.

Research Question: *Is Constant-Rate Dummy Traffic feasible to implement in low-latency, wireless, battery-powered Home Automation Systems?*

CRDT offers (ε-δ)-private communication (cf. Chapter 3). To achieve Constant-Rate Traffic throughout the HAS, two steps are necessary: First, genuine traffic has to be shaped to achieve a fixed maximum traffic rate. Then, times of inactivity must be padded with dummy traffic to achieve the same traffic rate. Our evaluation is based on exactly this mechanism. We assume an ideal CRDT scheme where during each timeslot, a dummy message is sent if and only if there is no genuine message to be transmitted.

The only variable in CRDT is the rate at which traffic is generated or permitted. A lower bound for this rate can be estimated by taking an exemplary use case: The user presses a switch and expects the light to turn on (or some other action to happen). The maximum acceptable value for this response time sets the minimum rate at which traffic must be permitted and generated. This value cannot necessarily be used directly, though. Due to the fact that multiple transmission requests may arise within a single timeslot, the actual reaction to the user's input may be delayed for more than a single period. Therefore, the rate has to be adjusted depending on the user and system behaviour.

Research in usability engineering suggests that a response time of 0.1 s is acceptable in most cases and a response time of more than 1 s is not acceptable.[Mil68; Nie93; BEN12] To have a conservative estimate, we assume a maximum delay of 1 s for any message. While delays of more than 1 s might be acceptable for devices such as thermostats or temperature sensors, splitting these from the rest of the system has no effect on the resulting minimum data rates.

4.3.1 Applying CRDT to the sample data

Figure 4.1 illustrates the delays of genuine messages given the aforementioned boundary conditions (max. 1 s delay).[1]

System 1

The minimum transmission rate which delays genuine messages for at most 1 s is $4\,\mathrm{P\,s^{-1}}$. 2263 messages (4.95 %) have to be delayed in order to reach this rate. The mean delay is 0.01 s and the median is 0 s. This means that at least 50 % of all genuine messages are not delayed at all. This is not surprising: The precision of the recorded timestamps is 1 s, so messages in this example do not have to be delayed by default to match the ending times of 1-second timeslots. Furthermore, the median of the (unmodified) inter-arrival times is 64 s. This means that 50 % of the genuine messages are sent at least 64 s after the previous message and thus only have to be delayed if the previous message has been delayed for more than 64 s.

As a second step, we fill the empty timeslots with dummy packets to reach a fixed transmission rate of $4\,\mathrm{P\,s^{-1}}$. In order to achieve this, a total of 12,220,618 dummy messages have to be generated, increasing the overall traffic by a factor of 267.53.

System 2.1

The higher overall traffic rate of System 2.1 also affects the transmission rate for CRDT: In order to delay genuine messages for at most 1 s, $7\,\mathrm{P\,s^{-1}}$ have to be sent. Accordingly, 8794 genuine messages (26.09 %) have to be delayed. The median delay is again 0 s—the same as for System 1. The mean and standard deviation are 0.04 s and 0.07 s, respectively—slightly higher than for System 1.

Due to the significantly higher transmission rate of genuine packets, less dummy messages have to be generated. The total number is 5,007,070—a 148.54-fold in-

[1] Within this thesis, the number of packets is denoted by the letter P.

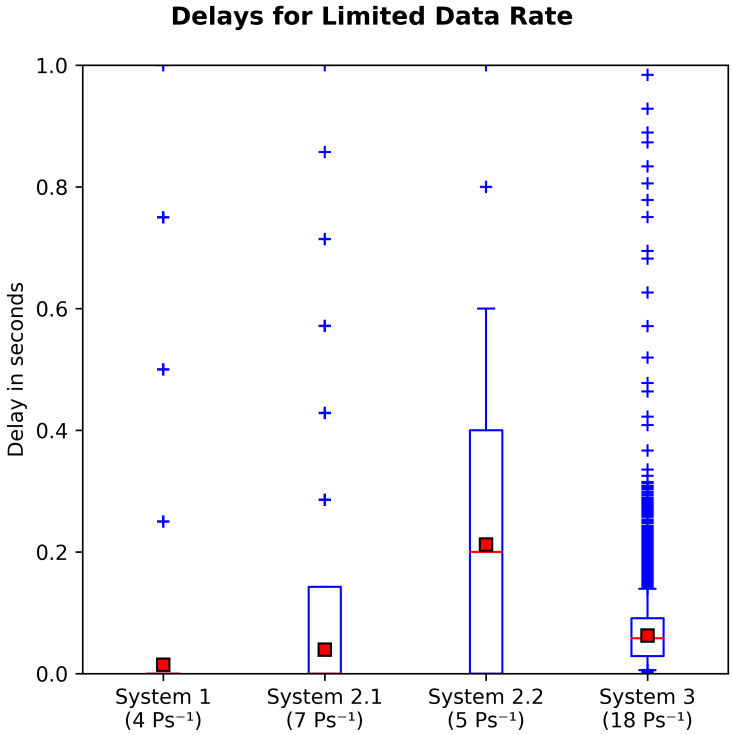

Figure 4.1: *Boxplots of transmission delays when limiting the data rate for a maximum delay of* 1 s*. The boxes extend from the first to the third quartile. The whiskers extend to the furthest sample point within* $1.5 * IQR$ *of the box where* IQR *is the interquartile range. Outliers are marked with blue plus signs. The red bars mark the median and the red squares mark the arithmetic means. Note that for System 1, the box seems to disappear because more than three quartiles are at* 0 s*. For System 2, the whisker disappears because there is no sample point within* $1.5 * IQR$ *of the box.*

crease. However, the relation is not proportional: Even though the genuine message rate is 3 times that of System 1, the number of dummy packets to be generated is only 50 % lower.

System 2.2

The bursty usage pattern of System 2.2 strongly affects the fixed transmission rate: In order to not delay genuine messages for more than 1 s, at least $5\,\mathrm{P\,s^{-1}}$ have to be sent. Despite this high rate, 59.46 % (594 packets) of all genuine traffic is delayed by an average of 0.21 s. The median is 0.20 s and the standard deviation is 0.22 s.

Due to the high transmission rate and the large gaps between genuine messages, the amount of dummy traffic to be generated is extremely high. A total of 5,806,822 dummy packets are required, leading to a 5812.63 times higher traffic volume.

System 3

System 3 is distinct from the other setups in that the message timestamps were captured with a higher precision. Among other things, this means that messages often have to be delayed for less than a second in order to match the ending time of a 1-second timeslots.

The minimum transmission rate for a delay of at most 1 s is $18\,\mathrm{P\,s^{-1}}$ and thus more than twice as high as that of System 2.1. The number of delayed messages shows that System 3 exhibits part of System 2.2's properties: Despite a high transmission rate, more than half of the messages (20,935 messages or 51.90 %) have to be delayed. However, the delays themselves are quite small. The mean and median are only 0.06 s and the standard deviation is 0.04 s.

The traffic increase factor is 1453.40 and thus between System 2.2 and the others. The total number of dummy packets is 58,624,417.

4.3.2 Evaluation of CRDT

After the application of CRDT, we measure its effect on the power consumption of HASs by applying different models to the resulting data. The constraint of having a maximum delay of 1 s for genuine messages results in different data rates: $4\,\mathrm{P\,s^{-1}}$ for System 1, $7\,\mathrm{P\,s^{-1}}$ for System 2.1, $5\,\mathrm{P\,s^{-1}}$ for System 2.2 and $18\,\mathrm{P\,s^{-1}}$ for System 3. Table 4.1 contains the detailed results. The results of applying the different models are explained in the following sections.

4.3.3 Traffic Overhead and System Responsiveness

All Systems exhibit bursts of traffic. This becomes very obvious in System 2.2, which has the lowest traffic rate before introducing CRDT. However, the usage scenario of the system explains the traffic pattern: The HAS is almost exclusively used in the morning and evening. However, many interactions are user-initiated and thus require quick feedback. The data rate thus has to be high enough to guarantee the upper bound of 1 s during the bursts. However, the high rate has to be kept during the whole day, effectively introducing a traffic overhead which is orders of magnitude larger than the amount of genuine messages.

System	1	2.1	2.2	3
Unmodified Data				
Timespan (days)	35.49	8.33	13.44	37.72
Messages	45,679	33,708	999	40,336
Inter-Arrival Times				
Minimum	0 s	0 s	0 s	0 s
Maximum	3668 s	153 s	61,645 s	25,050 s
Mean	67.14 s	21.36 s	1163.89 s	80.80 s
Median	64 s	14 s	0 s	0 s
Standard Deviation	49.67 s	27.49 s	5875.77 s	469.68 s
After applying CRDT				
Data Rate ($\mathrm{P\,s^{-1}}$)	4	7	5	18
Genuine Messages Delayed	4.95 %	26.09 %	59.46 %	51.90 %
Traffic Increase (Factor)	267.53	148.54	5812.63	1453.40
802.11 PC Card [FN01] (20 B, 11 $\mathrm{Mbit\,s^{-1}}$)				
Idle	52.26 MJ	8.54 MJ	6.03 MJ	4.69 MJ
Increase by CRDT (Factor)	4.55×10^{-4}	6.11×10^{-4}	1.29×10^{-3}	5.71×10^{-3}
EYES nodes [vDL03]				
Unmodified	29.68 kJ	4.81 kJ	3.37 kJ	4.05 kJ
Increase by CRDT (Factor)	0.31	0.83	1.36	1.14
Mica2dot [Wan+05] (20 B, 0.5 % Duty Cycle for Receivers, 0 % for Senders)				
Unmodified System	2600.36 J	634.89 J	383.55 J	5179.13 J
Increase by CRDT (Factor)	5.33	9.13	17.52	14.12
Telos [PSC05] (20 B, 0.5 % Duty Cycle for Receivers, 0 % for Senders)				
Unmodified System	1473.50 J	272.22 J	191.37 J	2395.58 J
Increase by CRDT (Factor)	0.44	0.98	1.02	2.28
HomeMatic				
Unmodified System	37.93 kJ	6.69 kJ	4.29 kJ	52.21 kJ
Increase by CRDT (Factor)	2.31	5.37	8.68	10.12

Table 4.1: *Effects of enforcing Constant-Rate Dummy Traffic in the sample installations under different energy consumption models. Numbers larger than* 10^{-2} *are rounded to two digits after the decimal point. The increase (both traffic and energy consumption) is given as a factor. A factor of 0 means no increase in traffic or energy consumption, whereas a 1 means that the original value doubled. Note that for System 3 using the 802.11 PC Card, a different idle power consumption was used.*

In System 2.1, several heating actuators are installed and could be subjected to a different maximum message delay. However, we recall from Section 2.1.4 that a remote control is also part of the installation and this direct user interaction requires a low response time. This remote control is responsible for the required data rate of $7\,\mathrm{P\,s^{-1}}$, removing any effect from a separate handling of heating actuators.

Decreasing the rate is no option: If we fix the rate to 1 packet/s, genuine messages are delayed for up to 59 s (in System 3), while the traffic still increases by a factor of almost 80 in the same system. The other systems perform similarly.

When taking the calculated data rates for CRDT, the mean delay for genuine messages is 0.2 s for System 2.2 and lower than 0.1 s for all other systems.

This demonstrates two intuitive properties of CRDT: On the one hand, the limitation of the transmission rate can lead to significant delays in regular traffic, negatively affecting responsiveness and thus usability. On the other hand, the introduction of dummy packets leads to an increase in overall traffic by a factor of at least 20. While the impact of each part can be lessened by adjusting the transmission rate, this immediately worsens the impact of the other. Increasing the transmission rate improves the responsiveness, but negatively affects the power consumption.

As a preliminary conclusion, we see that CRDT introduces a significant traffic overhead if the system responsiveness is not to visibly deteriorate. While this confirms the first intuition, its actual impact on energy consumption is not necessarily as strong.

4.3.4 Energy Consumption

It is widely perceived that in terms of required energy, communication is more expensive than e.g. computations. Wander et al. have stated that "the power required to transmit 1 bit is equivalent to roughly 2090 clock cycles of execution on the microcontroller" [Wan+05]. By this principle, many protocols for Wireless Sensor Networks (WSNs) have been developed, minimizing communication as far as possible. However, the cost of communication in Home Automation Systems has not been analysed thoroughly yet. It is unclear whether e.g. the idle power consumption of the devices outweighs the sporadic bursts of communication so that there is little need for further optimizations of the traffic volume. Naturally, manufacturers try to optimize their systems' battery lifetimes and thus put effort in minimize the power consumption. However, they usually do not publish specifications or information on the focus of their research.

In order to estimate the increase in power consumption by CRDT, we apply five different models to our data. Feeney et al. have conducted measurements on regular

802.11 PC cards, which can be installed in laptops. [FN01] Van Dam et al. have implemented a power-saving MAC protocol for WSNs and evaluated its performance on *EYES* nodes – battery-powered devices using an energy efficient microcontroller and a wireless transceiver. [vDL03] Wander et al. have evaluated the power consumption of public-key cryptographic protocols using the *Mica2dot* platform. [Wan+05] The Mica2dot platform uses the same transceiver as HomeMatic devices, which were used in two of our three analysed systems. Polastre et al. have developed a custom wireless node called *Telos* and compared its power consumption to others, including the Mica2dot. [PSC05] Due to the age of these models, we have performed our own measurements on HomeMatic hardware for the fifth model. The goal of this was to find out whether the energy consumption of wireless transceiver devices significantly changed in recent years or whether the existing models allow for a sensible approximation of (modern) HAS power consumption.

Table 4.1 shows the results of the evaluation. The following sections summarise these results and draw conclusions from them.

The five models lead to significantly different numbers for the total energy consumption, which can partly be explained by taking a closer look. Feeney et al.'s measurements were performed on laptop hardware. While laptops are built with energy efficiency in mind, the batteries are usually much larger than in HAS devices and the communication behaviour is significantly different. The other models use different kinds of hardware which were developed with particular use cases in mind.

802.11 PC Card, Feeney et al.

For the method presented by Feeney et al., we use the numbers of the $11\,\mathrm{Mbit\,s^{-1}}$ WaveLAN PC Card as it has a lower power consumption than the $2\,\mathrm{Mbit\,s^{-1}}$ model. For the calculations we use a packet size of 20 B – the same size used by van Dam et al. – for comparability. We also calculated the energy consumption using a packet size of 2048 B, but the conclusion is similar. With a packet size of 2048 B, the maximum increase when using CRDT is by a factor of 0.07 (or 7 %) and thus acceptable for most scenarios.

Feeney et al. did not take into account the time it takes to send a certain amount of data. They merely calculated the idle power consumption and then the additional energy consumption required to send a packet of a certain size. Thus, we can multiply the total sample time of each system by the idle power consumption. HomeMatic components as found in Systems 1 and 2 can communicate directly and without the need for a base station. We therefore use the "ad hoc" idle power consumption for these two systems. The exact setup of System 3 is unknown, but the publications suggest that an Wi-Fi (IEEE 802.11) network is used. We therefore use

the "BSS" mode idle power consumption for this setup.

In contrast to the other models, Feeney et al. have specifically considered the case that transmissions are received by all nodes within range of the transmitter, but only some of those are the intended receivers. The other nodes need to discard these packets after examining the addressing data—a process which requires a certain amount of energy. Feeney et al. have measured this energy and we have used the value for calculating the energy consumption of CRDT. It does however not play a significant role in the result, as the idle power consumption massively outweighs any other factor.

According to the study, the transmission of broadcast messages has a lower energy consumption than the transmission of point-to-point messages. We therefore use the broadcast energy consumption for our calculations. For the calculations it does not matter which device is sending which packet. However, when one device is sending data, at least one other device is receiving it and the remaining (non-receiving) devices need to discard the packet. This model is the only one with precise data available on the energy required to discard a packet.

The results of the application of this model are decisive: The energy consumed during idle phases is orders of magnitude higher than the energy consumption of message transmissions – both genuine and dummy. Thus, the negative effect of CRDT is likely unobservable. If a Home Automation System is therefore implemented using similar hardware, CRDT can be implemented with a negligible impairment of battery lifetime and system responsiveness.

EYES Nodes, Van Dam et al.

Van Dam et al. have developed and implemented a power-saving MAC protocol using so-called EYES nodes. The published data on power consumption is brief, but can be used to get a rough estimate of the energy consumption of a Home Automation System. Based on the graphs in the paper, we estimate that the maximum transmission rate of EYES nodes using the T-MAC protocol is about $58\,\mathrm{P\,s^{-1}}$. Since the relation between transmission rate and power consumption does not appear to be linear, we interpolated the missing data using a second-degree polynomial based on the power consumption of $1\,\mathrm{P\,s^{-1}}$, $10\,\mathrm{P\,s^{-1}}$ and $58.17\,\mathrm{P\,s^{-1}}$.

We then calculated the original transmission rates for every device in the evaluated systems during each 1 s-timeslot. Here we assume that all transmissions originating from a single sender and being sent during the same timeslot are targeted towards the same receiver. Furthermore, we assume that no two devices transmit data to the same receiver during the same timeslot. While this might not be entirely accurate, we

assume the error introduced by this to be negligible. This assumption is supported by the results of the other models.

Applying the consumption values from the model to this data results in an estimate of the energy consumed by the original system. We then introduce dummy traffic into the system at the given rates. While doing so, we distribute the dummy traffic evenly among all idle nodes, so that each node only sends at most one dummy packet during each 1 s-timeslot. This minimises the overall energy consumption and results in a more conservative estimate on the impact of CRDT.

The resulting total energy consumption of the systems is less than 0.06 % of that of Feeney's model. However, due to the smaller idle power consumption, the impact of transmissions on the total consumption is much higher. The energy consumption of the system is increased by at least 31 % (System 1) and up to 136 % (System 2.2). The large discrepancy between the different systems in this model can be explained by examining their usage patterns. While System 1 exhibits activity throughout the day from interaction and automation rules, System 2.2 almost solely handles bursts of interactions after long periods of inactivity.

The *relative* difference between the different systems is similar to the one in Feeney et al.'s model. System 3 being an outlier can be explained by the different model variation used previously in the first calculation.

If hardware similar to EYES nodes is is used to implement a Home Automation System, it becomes necessary to develop alternative approaches to CRDT in order to protect the users' privacy.

Mica2dot, Wander et al.

In order to measure the energy efficiency of different public-key cryptographic protocols, Wander et al. have performed measurements using the Mica2dot platform. The platform uses the same CC1100 transceiver as HomeMatic devices which make up 3 of our 4 analysed systems. When applying this data to our scenario, we make several assumptions: We assume the packet length is 20 bytes – similar to the other models. Furthermore, we have to use different power consumptions for active and inactive microcontrollers. Therefore, we differentiate between two categories of devices:

Senders (e.g. light switches) only initiate communication themselves. They react to events such as the pressing of a button and then begin communicating with other devices. The processors of Senders can therefore lie dormant for most of the time and only wake up when there is an event to be processed. This matches our observations during our own measurements which are described in Section 4.3.4.

Receivers (e.g. door locks or thermostats) react to messages from other devices. They therefore have to wake up periodically to check if there is a transmission to be processed and reacted upon. This mechanism is called *radio duty cycling*[2] Its exact parametrisation depends on the communication protocol and design decisions of the manufacturer. In our experiments we found out that HomeMatic Receivers wake up approximately once every 350 ms. Their processors are active about 0.5 % of the time. We apply this duty cycle to our calculations.

To calculate the energy consumption in the CRDT scenario, we assume that all dummy traffic is transmitted by Receiver nodes. This leads to a more conservative estimate than assuming Senders transmit all dummy traffic: Receivers exhibit a 0.5 % duty cycle anyway, so the transmissions of dummy packets pose a smaller impact on power consumption. In theory, the radio duty cycle could be combined with the carrier sense period before sending a packet. Sender devices, on the other hand, would have to wake up in order to be able to send packets.

Calculating intermediate results for the use of CRDT reveals that the transmissions force the Receivers to spend more than 0.5 % of the time in active mode. Transmitting 4 packets of 20 B each at a speed of 12.8 kbit s^{-1} takes about 0.05 s. This means that even if the dummy traffic generation is evenly distributed among all 9 Receivers and the genuine messages transmitted by Senders are subtracted, the Receivers have to spend more than 0.5 % of the total time transmitting data. This is to be expected, as packet transmissions have to take longer than one duty cycle by design, in order to give the receiving node a chance to detect the transmission. In practice, other tasks such as reading sensor values further increase the processing time of Receiver nodes.

The impact of CRDT on the energy consumption according to Wander et al.'s model is enormous: For System 1, the energy consumption is five times that of an unmodified system. For System 2.2, it the factor is as high as 17. Increasing the length of the packets further increases the impact of CRDT. A packet size of 200 B leads to an increase factor of at least 45. In conclusion, CRDT is infeasible for systems using similar hardware.

Telos, Polastre et al.

Polastre et al. have developed a new type of wireless node and compared it to previous hardware such as the Mica2dot platform. Their goal was to create a more energy-efficient device. In order to apply the model to our data, we made similar assumptions as for the other models. We assumed a packet size of 20 B and –

[2] In fact, the CC1100 transceiver uses a technique called *Wake on Radio*, where only the radio chip employs duty cycles while the CPU sleeps. [Tex09]

similarly to the application of Wander et al.'s Mica2dot model – split the devices into Senders and Receivers.

The results match the goals the Telos project: The energy consumption less than half of the consumption of the Mica2dot nodes. The effect of CRDT on the overall energy consumption is also lower. However, the overall energy consumption when using CRDT is still 44 % to 228 % *higher* than the original energy consumption of the systems. This supports our thesis that CRDT is not feasible for use in specialised HASs.

The results also suggest that Telos is a viable technology for use in HASs in general. Among our evaluation, it exhibits the lowest energy consumption both before and after applying CRDT across all four tested systems.

HomeMatic Hardware

To check the models against hardware of an existing HAS, we performed our own measurements on HomeMatic hardware. The test setup is described below Figures 4.2 and 4.3. We were able to confirm the most important aspects of power consumption by measuring and comparing the data with the other models. We could confirm our classification of devices as Senders and Receivers. Senders lie dormant most of the time with an idle power consumption of around 0.4 mW. Sending a (genuine) packet requires about 17.22 mJ of energy, including sensing the carrier before transmitting and listening for replies or collisions after the transmission. The energy consumption of a pressed switch where the different states are highlighted is supplied in Figure 4.2. Receivers, on the other hand, periodically wake up to listen for incoming transmissions. The duty cycle is 0.5 % and each spike requires about 80.79 µJ of energy in addition to the idle power consumption. The measurement for a door lock actuator without any activity on the carrier is illustrated in Figure 4.3

According to our measurements, the HAS systems consumed energy in the order of several kJ (over the full duration of 8 to 38 days) for communication. This is a reasonable amount for a battery-powered system: A single alkaline AA battery can supply around 10 kJ of energy.

The impact of CRDT on the overall energy consumption is significant. In a "busy" setting such as System 1, the energy consumption is more than tripled. This shows that the idle power consumption is relatively low in comparison to the power consumption of data transmissions. In a system handling bursts and long spans of inactivity such as System 2.2, the increase is nearly tenfold. This discrepancy again shows that CRDT performs worse in settings with sporadic transmissions. While the impact on System 3 is the highest, this value has to be interpreted with care:

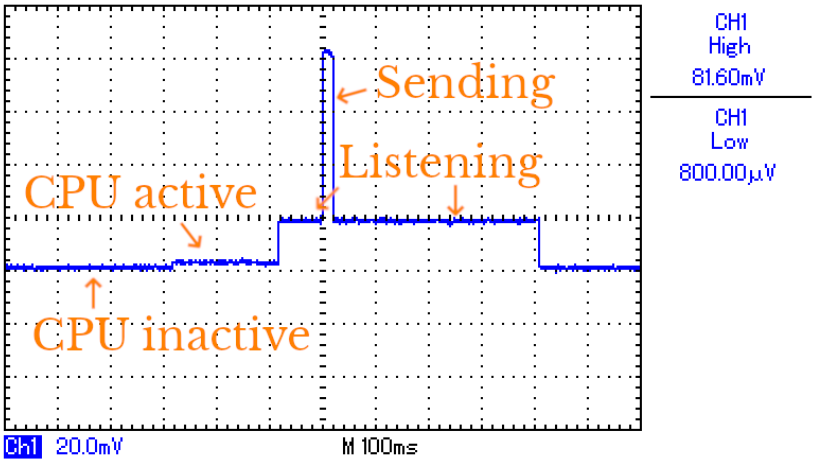

Figure 4.2: *Measurement of the voltage dropping across a* 2 Ω *shunt resistor connected in series to a HomeMatic light switch which was pressed. Y and X axis scales (major ticks) are given below the graph: X axis ticks are in intervals of* 20 ms *(minor ticks) and* 100 ms *(major ticks); Y axis ticks are in intervals of* 4 mV *(minor ticks) and* 20 mV *(major ticks). Measured properties of the curve are displayed on the right side: maximum voltage ("High") and minimum voltage ("Low"). The different states of the hardware are clearly identifiable. The switch is usually powered by two AAA batteries. For the measurement we used a laboratory power supply serving* 3 V.

The system is not built from HomeMatic components and therefore actual values might differ from the model. The discrepancy between busy systems and those

4.3.5 Conclusion of CRDT

From the analysis of CRDT power consumption we can draw two conclusions. The impact of CRDT on the power consumption strongly depends on the hardware used. On the one hand, there are systems where CRDT places a negligible strain on batteries and is therefore well suited to guarantee privacy. This is a strong contrast to the first intuition, which is that CRDT introduces too much traffic overhead to be feasible. On the other hand, many specialized systems with low idle power consumption are heavily impacted by CRDT. For those systems, different approaches need to be developed and implemented.

Additionally, regulatory thresholds need to be obeyed when increasing the trans-

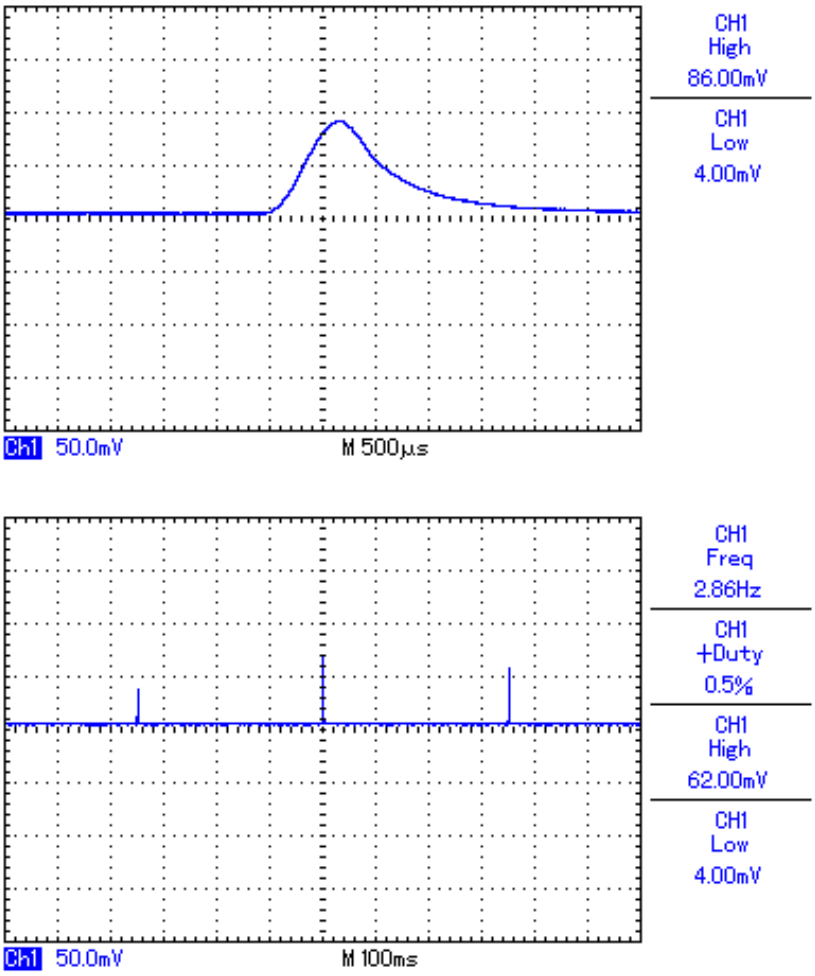

Figure 4.3: *Measurement of the voltage dropping across a* $2\,\Omega$ *shunt resistor connected in series to a HomeMatic door lock actuator. Y and X axis scales (major ticks) are given below the graphs, similar to Figure 4.2. Measured properties of the curves are displayed on the right side: maximum voltage ("High"), minimum voltage ("Low"), pulse frequency ("Freq") and duty cycle ("+Duty"). No traffic was sent or received during the measurement and the lock was not turned. The lock is usually powered by two AA batteries which were swapped for a laboratory power supply serving* $3\,\mathrm{V}$ *for the experiment. In the left picture a single pulse is measured in detail. In the right picture the duty cycle is illustrated. Note that due to limitations of the measuring equipment, the spikes in the right picture appear to be of different height. However, the pulses all follow the same shape as the one in the left picture.*

mission frequency. This imposes an upper limit on the packet rate which depends on the technologies used and can lead to observable delays and, consequently, a degradation in usability.

A possible alternative to achieving (ε-δ)-private communication with CRDT is to try and achieve (ε-δ)-indistinguishability or (ε-δ)-unobservability. For this, the HAS designer first has to choose a set of tasks $\mathbb{T}$ for which the definitions in Section 3.3.4 will hold. The following section deals with one such approach.

4.4 Naive Exponential Dummies (NED)

Research Question: *Is it possible to provide privacy guarantees at a lower power consumption overhead than CRDT?*

The power consumption of systems using CRDT has made it clear that low latency and constant-rate traffic are incompatible for most (optimised) HASs. We therefore relax our requirements on the privacy goals and present an approach using a probabilistic generation of dummy traffic. This approach introduces significantly less traffic overhead while introducing no latency to user interaction and offering (ε-δ)-unobservability for certain interactions. We call this approach *Naive Exponential Dummies (NED)*.

The approach works as follows: Genuine traffic is untouched by the system and is transmitted without delays. After every message (genuine or dummy), the system generates a random duration d from an exponential distribution with rate λ. If no genuine message is transmitted after this time d, a dummy message is generated and transmitted. If a genuine message does appear before, a new number is sampled from the same distribution.

The choice of the random distribution is arbitrary. For this work we chose the exponential distribution as it is generally used to model the time between subsequent events of a Poisson point process and because it offers provable and reasonable privacy as shown later in this section. However, we stress that no stochastic model for user interaction exists and that this choice does not constitute an effort to substitute one. Furthermore, we explicitly acknowledge the possibility of other distributions providing better results.

In reality, a system cannot choose the transmission time of a message with arbitrary precision (it is limited by the maximum transmission rate and other physical properties) and the precision with which the adversary can determine the timestamp of a captured message is also limited by the equipment. Therefore, a discrete model is more suitable than a continuous one. To convert the continuous exponential dis-

tribution into a discrete one, the system rounds down the drawn number to the nearest possible transmission time. Alternatively, it can draw a number directly from a geometric distribution with success (sending) probability $p = 1 - e^{-\lambda}$.

Note that the algorithm does not necessarily exclude 0 as a possible outcome. In practice, systems may limit the number of messages transmitted in the same timeslot or may disallow simultaneous messages altogether. In these cases, the value 0 can be interpreted as the *next possible* transmission time rather than the current time.

If the timestamps are to be modelled as continuous or if the timestamp precision is too high to provide meaningful results, they can be transformed into a discrete model nevertheless. By using Apthorpe et al.'s approach [Apt+19], genuine traffic can be shaped into following a fixed maximum transmission rate without visibly affecting system responsiveness, similar to the functioning of CRDT.

In this section we focus on the discrete model as it captures the properties of real systems better than the continuous one. However, since the goals are also defined for continuous models, similar analyses can be performed for those cases or for approaches which require an underlying continuous model. We assume that for the sending probability p of the geometric distribution, it holds that $0 < p < 1$. Furthermore we assume a geometric distribution with the possible outcomes $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$[3]. As mentioned before, a result of 0 can be interpreted as the next possible transmission slot rather than the same.

We can immediately deduce an important property of NED:

Theorem 1. *Running the algorithm on n different nodes simultaneously with sending probability $1 - \sqrt[n]{1-p}$ leads to the same distribution of inter-arrival times (time between subsequent messages) for dummy traffic as running it on one node with sending probability p.*

Proof. Let $X_1, \ldots, X_n$ be stochastically independent random variables following geometrical distributions with success probabilities $p_1, \ldots, p_n$ and describing the inter-arrival times on n different nodes. Let $X = \min\{X_1, \ldots, X_n\}$ be the combined random variable describing the overall inter-arrival times.

Then, since the geometric distribution is memoryless,

[3]For the theoretical analysis in this thesis we define the natural numbers to include 0.

$$
\begin{aligned}
&Pr(X = k) \\
&= Pr(X \geq k) \cdot Pr(X = k | X \geq k) \\
&= Pr(X \geq k) \cdot (1 - Pr(X > k | X \geq k)) \\
&= Pr(X \geq k) \cdot (1 - Pr(X_1 > k \wedge \ldots \wedge X_n > k | X \geq k)) \\
&= Pr(X \geq k) \cdot (1 - \prod_{i=1}^{n} Pr(X_i > k | X_i \geq k)) \\
&= Pr(X \geq k) \cdot (1 - \prod_{i=1}^{n} (1 - p_i)) \\
&= Pr(X_1 \geq k \wedge \ldots \wedge X_n \geq k) \cdot (1 - \prod_{i=1}^{n} (1 - p_i)) \\
&= \prod_{i=1}^{n} (Pr(X_i \geq k)) \cdot (1 - \prod_{i=1}^{n} (1 - p_i)) \\
&= \prod_{i=1}^{n} (1 - p_i)^k \cdot (1 - \prod_{i=1}^{n} (1 - p_i)) \\
&= \left(\prod_{i=1}^{n} (1 - p_i) \right)^k \cdot (1 - \prod_{i=1}^{n} (1 - p_i))
\end{aligned}
$$

Therefore, X follows a geometric distribution with sending probability $1 - \prod_{i=1}^{n} (1 - p_i)$. If $\forall p_i : p_i = 1 - \sqrt[n]{1-p}$, then X follows a geometric distribution with sending probability $1 - \prod_{i=1}^{n} (1 - (1 - \sqrt[n]{1-p})) = p$ □

While this property is not groundbreaking, it serves an important purpose for practical system design: For NED, the sending probability p is the only parameter that needs to be synchronized between devices. Aside from this, the nodes can take decisions about the generation of dummy traffic locally and do not need to coordinate every transmission.

The sending probability p (or mean time between dummy messages λ) can be adjusted to reach a balance between privacy and energy efficiency. For $p = 0$, no dummy messages are generated so there is no impact on power consumption and none on privacy. For $p = 1$ NED generates CRDT at the maximum possible rate. Since the two corner cases have already been analysed, we excluded them in our assumptions

and focus on the impact of $0 < p < 1$ on privacy guarantees and energy efficiency in the following sections.

4.4.1 Privacy Guarantees of NED

In this section we analyse NED with respect to its privacy guarantees. Note that while NED uses a geometric distribution for the generation of dummy traffic, no assumption is made about the distribution of genuine user interaction. The proofs in this section hold for any distribution of genuine events.

ε-δ-private Communication

As a first step, we prove that NED and, more generally, any approach which neither uses CRDT nor delays genuine messages cannot offer ε-private communication.

Theorem 2. *NED and any approach which neither uses CRDT nor delays genuine messages does not provide ε-private communication.*

Proof. Let S be the set of all possible genuine message timestamps. For each element $s \in S$, we define a task T_s = "interact with the system in any way that invokes a message with timestamp s". Such a task must exist because genuine messages are not delayed. We also define a complementary task $T_{\bar{s}}$ = "interact with the system in any way so that no genuine message is generated at timestamp s".

Let $x \in S$ be a time at which a dummy packet is not necessarily generated (i.e. the probability of generating a dummy packet at x is less than 1). Such a time must exist since the dummy traffic generation scheme is not CRDT. Let $O_x = \emptyset$ now be an empty observation covering only the instant at time x. Then $Pr(O_x|T_x) = 0$ because executing T_x by definition invokes a message with timestamp x. However, $Pr(O_x|T_{\bar{x}}) > 0$ because executing $T_{\bar{x}}$ does not invoke a genuine message at time x. The latter probability is not necessarily 1, because the dummy traffic generation scheme *may* generate a dummy packet at time x.

There is no constant $\varepsilon > 0$ which satisfies $0 < Pr(O_x|T_{\bar{x}}) \leq e^{\varepsilon} \cdot Pr(O_x|T_x) = 0$. Consequently, the system does not offer ε-private communication.

For a continuous model, the same holds. Let $\mathbb{O}_x = O|x \notin \mathbb{O}$ be the set of all possible adversarial observations where no message appears at time x. Then by the same reason as above, $Pr(O \in \mathbb{O}_x|T_x) = 0$ and $Pr(O \in \mathbb{O}_x|T_{\bar{x}}) > 0$. Therefore, the system does not offer ε-private communication in the continuous model. □

The proof formally describes an intuitive but important property of NED and other latency-free approaches: If the attacker captures no messages within a certain time frame, they know that no task was executed that would have invoked a packet within this time frame. It also shows that in order to provide ε-private communication, any approach has to introduce artificial delays to genuine traffic, affecting the responsiveness of the HAS.

This is especially important if messages are not padded to a uniform length. If the goal is to provide ε-privacy for all classes of messages that an adversary is able to distinguish, then CRDT has to be applied to each. For our test data, this would introduce an overhead larger than padding all messages to the maximum length.

For the discrete version of NED specifically we can also prove that it does not offer (ε-δ)-private communication. The proof follows the idea that it is possible to generate arbitrarily improbable adversarial observations (e.g. long chains of messages) so that any constant δ will eventually be surpassed. Intuitively, this corresponds to a scenario where a user continuously presses a light switch over and over again.

Theorem 3. *NED does not provide (ε-δ)-private communication under a discrete time model.*

Proof. For all natural numbers $n \in \mathbb{N}$ we define a task T_n ="interact with the system in a way so that n consecutive messages are generated" and a corresponding observation $O_n = [0, n] \cap \mathbb{N}$ comprising n consecutive messages and covering a duration of exactly n. By construction, it holds that $\forall n \in \mathbb{N} : \Pr(O_n|T_n) = 1$ We also define a task T_0 ="interact with the system in a way so that no genuine messages are generated". For this task it holds that $\forall n \in \mathbb{N} : \Pr(O_n|T_0) = p^n \cdot (1-p)^n$ where p is the sending probability of NED (note that the factor $(1-p)^n$ describes the event that not more than one message is generated in each timeslot).

Assuming that NED does offer (ε-δ)-private communications, there must be two constants $\varepsilon > 0, \delta < 1$ so that for all tasks T_i, T_j and for all observations O it holds that $Pr(O|T_i) \leq e^{\varepsilon} \cdot Pr(O|T_j) + \delta$.

For $m := \left\lceil \log_p \left(\frac{1-\delta}{e^{\varepsilon}} \right) \right\rceil$ it then holds that $Pr(O_m|T_m) = 1$ by the construction above. Due to the privacy guarantee, it holds that

$$1 = Pr(O_m|T_m) \leq e^{\varepsilon} \cdot Pr(O_m|T_0) + \delta = e^{\varepsilon} \cdot p^m \cdot (1-p)^m + \delta$$

Due to the construction of m and since $0 < p < 1$, it further holds that

$$e^{\varepsilon} \cdot p^m \cdot (1-p)^m + \delta \leq e^{\varepsilon} \cdot \frac{1-\delta}{e^{\varepsilon}} + \delta = 1$$

However, since $0 < p < 1$, we can come up with another sample O_{m+1} for which it holds that

$$\begin{aligned} & e^{\varepsilon} \cdot Pr(O_{m+1}|T_0) + \delta \\ =& e^{\varepsilon} \cdot p^{m+1} \cdot (1-p)^{m+1} + \delta < e^{\varepsilon} \cdot p^m \cdot (1-p)^m + \delta \end{aligned}$$

And since $e^{\varepsilon} \cdot p^m + \cdot(1-p)^m + \delta = Pr(O_m|T_0)$ as shown above, it holds that

$$Pr(O_m|T_m) > e^{\varepsilon} \cdot Pr(O_{m+1}|T_0) + \delta$$

This violates the condition of (ε-δ)-private communication. □

This proof illustrates another intuitive property of NED: If the user invokes a sequence of messages that is unlikely to be generated from dummy traffic, then the attacker has a high confidence in identifying this interaction.

ε-indistinguishability

If a system does not equalise message sequences with regard to the number of messages and their inter-arrival times, approaches like NED that are not bounded cannot offer ε-indistinguishability for tasks invoking different message sequences. The theorem and its proof can be visualised using the following example: Suppose that pressing a light switch makes the system transmit 3 consecutive messages and that opening a door makes it transmit 4 messages. If the system now uses NED for the generation of dummy traffic, it is possible that after pressing a light switch, still only 3 messages are transmitted within the observed time frame. If an attacker captures this output, they can be certain that the light switch was pressed rather than the door being opened. Since ε-indistinguishability requires one of the tasks to be performed, this information is leaked to the attacker.

For the proof, we assume that performing a task T invokes a message sequence $S = \{s_1, s_2, \ldots, s_n\}$ where s_i is a random variable following a (usually, but not necessarily

bounded) probability distribution and describing the time difference between the i-th message and the point at which the task was performed. Note that using slightly different models (e.g. s_i describing the inter-arrival time between messages i and $i-1$) requires minor adaptions, but does not invalidate the proof.

Theorem 4. *Let T_A, T_B be two tasks and let $A = \{a_1, a_2, \ldots, a_m\}, B = \{b_1, b_2, \ldots, b_n\}$ be the message sequences invoked by performing T_A or T_B, respectively.*

If the tasks do not fully overlap, or formally if

$$\exists i \in [1,m] \forall x \in \mathbb{N} :$$
$$Pr(a_i = x) > 0 \Rightarrow \forall j \in [1,n] : Pr(b_j = x) = 0$$

then NED (or any other unbounded probabilistic approach) does not provide ε-indistinguishability for the set $\{T_A, T_B\}$.

Proof. Let i be such that $\forall x \in \mathbb{N} : (Pr(a_i = x) > 0 \Rightarrow \forall j \in [1,n] : Pr(b_j = x) = 0)$ (cf. the precondition). Let $Q \subseteq \mathbb{N}$ be the set for which $\forall q \in Q : Pr(a_i = q) > 0$, therefore $\forall q \in Q, j \in [1,n] : Pr(b_j = q) = 0$.

Then for all observations O where $Pr(O|T_B) > 0$ and $O \cap Q = \emptyset$ it holds that $Pr(O|T_A) = 0$. Such observations must exist because elements from Q are not generated by performing T_B and if the dummy traffic generation algorithm is NED or another unbounded probabilistic scheme, then there are observations O with $O \cap Q = \emptyset$ which may occur when performing T_B.

Consequently, there is no constant $\varepsilon \geq 0$ which satisfies $0 < Pr(O|T_B) \leq e^{\varepsilon} \cdot Pr(O|T_A) = 0$. The system therefore does not offer ε-indistinguishability for the set $\{T_A, T_B\}$. □

We can conclude that for strict ε-indistinguishability, the message sequences invoked by the tasks have to be modified so that no possible sequence of one tasks is impossible for another. In practice, this may be achieved by equalising the length of message sequences, e.g. by ensuring that after every user interaction, a fixed number of messages is transmitted in the same fixed or equally distributed intervals. Apthorpe et al. follow this idea in their approach named STP. [Apt+19]

However, it is not always necessary for a HAS to offer (ε-δ)-private communication or (ε-δ)-indistinguishability. If the user wants to e.g. only guarantee that an attacker is unable to find out whether they are at home, it might be sufficient to provide (ε-δ)-unobservability for a set (or sets) comprising "normal" tasks which involve the user directly, such as opening doors and pressing switches (a reasonable number of times).

ε-δ-unobservability

As a last step we analyse NED with regard to (ε-δ)-unobservability. We show that NED achieves this goal and calculate the values of ε and δ for a given scenario. For the proof we assume the following case: As in the previous section, performing a given task T invokes a number of messages $S = \{s_1, s_2, \ldots, s_n\}$ where each s_i is a random variable following any probability distribution and describing the timing of message i relative to the task's execution time. S matches the set E of interesting events or messages from Chapter 3; we use a different letter here to avoid confusion with the constant e.

Theorem 5. *Let T be a task invoking a sequence of messages $S = \{s_1, s_2, \ldots, s_n\}$ and let $\overline{T}$ be the complementary task invoking no genuine message. Then NED offers (ε-δ)-unobservability of $\{T, \overline{T}\}$.*

Proof. Let O be any adversarial observation of duration l. We distinguish between the following (possibly overlapping) cases:

1. $Pr(O|T) > 0$ Then O will unconditionally include a number m subject to $0 \leq m \leq n$ of genuine messages generated by executing T and a number $d = |O \cap D| \geq 0$ of dummy messages. An upper bound for this probability $Pr(O|T)$ can be computed by calculating the probability of observing exactly the same d dummy messages in an observation of length l which already includes the m genuine messages:

   $$Pr(O|T) \leq p^d \cdot (1-p)^l$$

   Then, the probability of observing the same pattern with no genuine messages is

   $$Pr(O|\overline{T}) = p^m \cdot p^d \cdot (1-p)^l$$

   If we insert these two terms into the equation for (ε-δ)-unobservability, we get

   $$\begin{aligned} & Pr(O|T) \leq p^d \cdot (1-p)^l \\ \leq & e^{\varepsilon} \cdot Pr(O|\overline{T}) + \delta \\ = & e^{\varepsilon} \cdot p^m \cdot p^d \cdot (1-p)^l + \delta \end{aligned}$$

   We see that for $\varepsilon = -\ln(p^m) > 0$ (since $0 < p < 1$ and therefore $p^m < e$) and $\delta = 0$ the two sides become the same, satisfying the condition. Since $m \leq n$, the value $\varepsilon = -\ln(p^n)$ is a lower bound.

For the opposite direction ($Pr(O|\overline{T}) \leq e^{\varepsilon} \cdot Pr(O|T) + \delta$), we need a lower bound for the probability $Pr(O|T)$. Since the probability of any timestamp being yielded by NED is larger than 0, the probability of observing any message cannot be lower than the probability of this message being a dummy generated by NED. This value is the sending probability p. Therefore, a lower bound is $Pr(O|T) \geq p^{m+d} \cdot (1-p)^{l}$. Inserting this into the equation we get

$$\begin{aligned} Pr(O|\overline{T}) &= p^{m} \cdot p^{d} \cdot (1-p)^{l} \\ &= p^{m+d} \cdot (1-p)^{l} \leq Pr(O|T) \\ &\leq e^{\varepsilon} \cdot Pr(O|T) + \delta \end{aligned}$$

We immediately see that the condition is satisfied for $\varepsilon = \delta = 0$.

2. $Pr(O|T) = 0$ Then the equation $Pr(O|T) \leq e^{\varepsilon} \cdot Pr(O|\overline{T}) + \delta$ is satisfied for any value of $\varepsilon > 0$ and $\delta < 1$.

   As for the opposite direction, it is obvious that $Pr(O|\overline{T}) > 0$ since NED can yield any set of timestamps. The maximum possible value of this probability is therefore an upper bound for the constant δ. This maximum is trivial to compute: Since $Pr(O|T) = 0$, O must not have a packet over a period where T would generate one. An upper bound for observing this "silence" is $1 - p$. Hence, we get that $\delta \leq 1 - p$.

□

4.5 Evaluating NED

In order to evaluate NED against real HAS data, we implemented the algorithm and ran it against our evaluation data. During the implementation we made the following design decisions:

- The algorithm generates samples from an exponential distribution and rounds the result down to the nearest possible transmission time.
- Since our sample data contains messages with the same timestamp due to limitations of the capturing hardware, we allowed 0 as a possible result for the next dummy message.

The sample data of each system was used as a single trace of genuine traffic. The generated dummy traffic was added on top of the genuine messages similar to how

NED could be implemented in practice. As stated in Section 4.4, we placed no assumption on the distribution of genuine traffic. Instead, we used the realistic sample data as-is.

For each system, we generated dummy traffic using six different values for the mean inter-arrival time λ of the exponential distribution. We also performed an analysis using no dummy traffic. We generated up to 1000 observations from the resulting system output, making sure that at least 40 of them included user interaction. The duration of these observations was set to 10 s. We performed the evaluation with longer samples as well, but reached the same conclusion. For each sample, we counted the number of times it occurred in the system output including dummy messages. The occurrences were split between those where the same kind of user interaction happened as in the sample and those where different or no user interaction was recorded. We ran the complete simulation multiple times to see if the calculated values are stable. Since NED can generate extreme traffic patterns (no or very high dummy traffic), it is possible to see high variations in the results, although with a low probability. The results below are the average of 55 runs; the values did not differ significantly.

Using this data we estimate values for ε and δ by counting occurences of individual observations and applying the formula from Definition 4. We also estimate the effect on the energy consumption according to the HomeMatic model from Section 4.3.4. The results are summarised in Table 4.2.

NED offers significant privacy improvements for moderate increases of overall traffic. For higher values of λ, the traffic overhead increases quickly and so do the privacy guarantees. When NED is configured to send approximately one packet every second, the parameter ε approaches 0 in all four systems. The constant privacy leakage δ also becomes diminishingly small. However, the energy consumption is merely doubled for the first two systems. For System 3, the increase is only 50 %.

Even low values of λ alread provide significant improvements over non-anonymised systems: When transmitting one dummy every 10 seconds, the constant leakage δ is below 0.5 in all systems, while the increase in energy consumption is below 20 % for all four systems and below 10 % for three of them.

A suitable compromise between privacy and energy consumption seems to be around $\lambda = 0.5$. ε drops to at most half the value it has in the unmodified systems. Due to it being in the exponent when comparing probabilities, the factor e^{ε} drops by more than 97 %. Furthermore, the constant leakage δ drops below 0.01 in all systems, which means that the chance of an attacker being able to learn a definitive piece of information is close to zero. The increase in energy consumption is between 6 and 50 % for all systems except System 2.2.

System	**1**	**2.1**	**2.2**	**3**
No dummy traffic				
TI	0.00	0.00	0.00	0.00
ε	8.77	7.62	7.83	7.63
δ	0.87	0.70	0.99	1.00
$\boldsymbol{\lambda = 0.001}$ (∼1 P every 20 min)				
TI	0.07	0.02	1.16	0.08
ε	8.57	7.73	7.87	7.63
δ	0.86	0.70	0.99	0.99
ECI	6.05×10^{-4}	7.23×10^{-4}	1.73×10^{-3}	5.57×10^{-4}
$\boldsymbol{\lambda = 0.01}$ (∼1 P every 100 s)				
TI	0.67	0.22	11.73	0.80
ε	7.20	7.76	7.97	7.67
δ	0.79	0.64	0.90	0.91
ECI	5.79×10^{-3}	7.95×10^{-3}	0.02	5.57×10^{-3}
$\boldsymbol{\lambda = 1/60}$ (∼1 P every minute)				
TI	1.13	0.36	19.39	1.35
ε	7.08	7.25	8.19	7.65
δ	0.74	0.60	0.85	0.85
ECI	0.01	0.01	0.03	0.01
$\boldsymbol{\lambda = 0.1}$ (∼1 P every 10 s)				
TI	7.07	2.25	122.33	8.08
ε	5.27	6.74	6.92	7.50
δ	0.32	0.26	0.37	0.37
ECI	0.06	0.08	0.18	0.06
$\boldsymbol{\lambda = 0.5}$ (∼1 P every 2 s)				
TI	43.55	13.84	754.83	40.40
ε	3.91	$<10^{-10}$	$<10^{-10}$	3.97
δ	5.93×10^{-3}	4.78×10^{-3}	7.11×10^{-3}	6.71×10^{-3}
ECI	0.38	0.50	1.13	0.28
$\boldsymbol{\lambda = 1}$ (∼1 P every second)				
TI	115.39	36.69	1995.63	80.85
ε	$<10^{-10}$	$<10^{-10}$	$<10^{-10}$	$<10^{-10}$
δ	5.56×10^{-5}	1.46×10^{-3}	8.55×10^{-4}	7.50×10^{-9}
ECI	1.00	1.33	2.98	0.56
CRDT (for reference, rates from Sec. 4.3.2)				
TI	267.53	148.54	5812.63	1453.40
ε	0	0	0	0
δ	0	0	0	0
ECI	2.31	5.37	8.68	10.12

Table 4.2: *Results for using NED in Home Automation Systems. TI stands for traffic increase and ECI means Energy Consumption Increase according to the HomeMatic model. Both are given as a factor, where a value* 0 *means no increase and a value of* 1 *means that the original value doubled. TI and ECI for CRDT is supplied for comparison. Since CRDT offers (0-0)-private communication, the values for* ϵ *and* δ *are 0.*

4.5.1 Behaviour of ε and δ over Time

We have investigated how ε and δ develop over different stretches of time. First we consider a simple theoretic case: A user performs an interaction at the same time every day. The interaction invokes a chain s of n consecutive messages. The probability of observing s at this time each day is 1. The probability of observing n consecutive dummies in absence of the interaction is p^n. The probability of observing a sample of length n with at least one gap in the absence of interaction is $1 - p^n$. Thus, for one day the privacy parameters are $\delta = 1 - p^n$ and $\varepsilon = -n \cdot \ln p$ ((ε-δ)-unobservability).

When monitoring the system for multiple days (or task execution periods), however, the adversary observes the same pattern at the same time. Thus, after k days of monitoring, the probability of observing the same pattern of n consecutive dummy messages at the same time each day is p^{nk} and the probability of observing anything else than this pattern (i.e. at least one gap) is $1 - p^{nk}$. This means that the privacy parameter $\varepsilon = -nk \cdot \ln p$ linearly rises with k while $\delta = 1 - p^{nk}$ converges to 1. As an example, for $\lambda = 0.5 \Rightarrow p \approx 0.39$, $n = 4$, ε starts at $\varepsilon \approx 3.73$ for $k = 1$ and rises to $\varepsilon \approx 26.12$ for $k = 7$. δ starts at $\delta \approx 0.61$ and decreases to $\delta \approx 0.03$ for $k = 7$.

We have also tried to extract the behavior of ε and δ over time from the sample data. However, we observe that tasks are not as regular as in our theoretical example: With $\lambda = 0.1$, the parameter ε for System 3 starts at 2.97 for a timeframe of 6 hours and rises to 9.45 when taking a week's traffic data. For the full dataset it then decreases back to 7.50. This can be explained when examining the parameter calculation: For a given interaction in a short timeframe, a particular sample is unique. When observing longer stretches of time however, the same interaction can invoke different observations (as message delays might slightly differ), resulting in different samples. For example, pressing a light switch twice does not necessarily generate packet sequences with the exact same inter-arrival times.

Furthermore, short timeframes result in a high variance of the privacy parameters due to the limited data. In order to properly analyze real behavior of ε and δ over time, larger data sets are needed. The theoretical and practical values have been plotted in Figure 4.4, although the explanatory power of this graph is limited due to the aforementioned constraints.

4.6 Chapter Conclusion

In this chapter we have achieved two major goals of this thesis: We have analysed the impact of Constant-Rate Dummy Traffic on both the responsiveness and en-

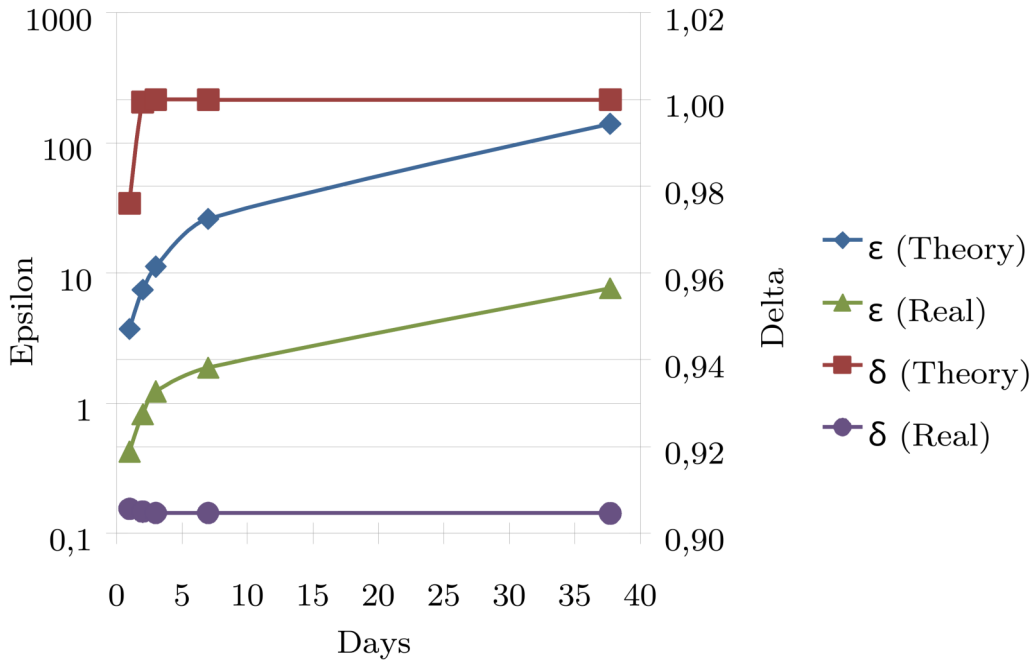

Figure 4.4: *Simulated and measured development of ϵ and δ over time ((ε-δ)-unobservability). The scale for ϵ is logarithmic for increased readability. Values are calculated for $\lambda = 0.5$ and real values are estimated from the data of System 3. Note that the real values are subject to limited data and to variations in message sequences for individual interactions.*

ergy efficiency of Home Automation Systems. As we have shown, CRDT can be implemented with acceptable impact on responsiveness and thus user experience. Depending on the hardware used, the approach can offer perfect (ε-δ)-private communication at moderate cost in terms of power consumption. We have also identified cases where CRDT is infeasible to implement and have introduced Naive Exponential Dummies as a robust alternative. We have formally analysed NED with regard to the security guarantees and have shown its limits as well as its strengths. In the last part of this chapter we have simulated NED on realistic HAS data and were able to show that it can offer reasonable guarantees at moderate cost in terms of energy efficiency. NED therefore poses a strong first step towards practical energy-efficient dummy traffic generation for Home Automation Systems.

Chapter 5

Conclusion

In this thesis we have investigated the problem of privacy in Home Automation Systems. As a first step, we have tried to answer the question:

What information can a passive attacker learn about the user of a Home Automation System if communication is unencrypted?

We have shown that current, commercially available systems are vulnerable to passive attacks. For this we have performed the first extensive analysis of 3 HomeMatic installations, capturing traffic over the course of several weeks and *identifying device classes, automation rules and user habits*. We have demonstrated that without encryption, these systems leak detailed information about everyday activities to any outside observer and that these eavesdropping attacks are possible using cheap, readily available hardware. [Möl+14]

We have then demonstrated that encryption alone does not fully protect the users' privacy. By applying statistical tests to timing metadata of two different Home Automation Systems we could prove that presence information can still be leaked even if the adversaries do not have access to the contents of data packets. The question we tried to answer was:

If an attacker has captured 1 hour of traffic from a user's HAS and knows whether the user was present at that time, can the attacker deduce the user's state by capturing another hour of traffic?

We were able to identify *cases in which user presence could be deduced with absolute certainty*, though it is unclear whether these results are generalizable to all systems and users. [MS16] Machine learning algorithms can slightly improve the prediction

accuracy, but so far do not significantly enhance an attacker's capabilities beyond individual statistical tests.

We have also inspected the legal situation with respect to both criminal and data protection law. We have asked ourselves:

Is Home Automation System communication protected by data protection law? If yes, which implications does this have for users and providers?

and

Does criminal law successfully capture the attacks described here? Or are amendments to the law necessary in order to provide legal protection for users?

Regarding data protection law, we have shown that *HASs generally process personally identifiable information* and that *providers must therefore take precautions to protect this data from unlawful access.* [MS15] We have highlighted shortcomings of past and present legislations and have shown that while *options for criminal prosecution exist*, legislation cannot replace technical countermeasures. This is especially true as the *law can only follow known developments in computer science and only marginally regulate future developments.* [MV16] However, we have also found that *there are international attempts to unify data protection and computer-related criminal law in order to facilitate privacy protection and prosecution of attacks.* [Möl+18]

As a consequence of these findings, we have investigated methods for modelling passive attacks on (encrypted) Home Automation System communication in order to answer the question

How can communication and traffic analysis attacks in HASs be modelled in order to develop and compare countermeasures?

Based on existing and well-established approaches from the field of Differential Privacy and Private Information Retrieval, we have constructed a *general model for passive attacks on (encrypted) Home Automation System communication.* We have formalized the notion of privacy in HASs and have *defined several privacy goals* which can be guaranteed by suitable traffic shaping schemes. Our model is *agnostic to communication protocols, other countermeasures such as header encryption, and user behaviour.* It allows the development and comparison of communication algorithms with respect to information leakage. To the best of our knowledge this is the *first model established for the particular case of Home Automation Systems*, where the mere existence of communication can already disclose sensitive data. [Möl+18]

Finally, we have analyzed the applicability of the model and the performance of two countermeasures. In the first step we have asked ourselves:

Is Constant-Rate Dummy Traffic feasible to implement in low-latency, wireless, battery-powered Home Automation Systems?

We have demonstrated the applicability of the model by performing a quantitative analysis of Constant-Rate Dummy Traffic. We have shown that contrary to the first intuition, *constant-rate communication can offer perfect privacy at moderate cost in terms of power consumption*, depending on the hardware used. We have also identified settings which make CRDT infeasible due to a significant power consumption overhead. [Möl20] As a second step, we have therefore tried to answer the question:

Is it possible to provide privacy guarantees at a lower power consumption overhead than CRDT?

For cases where CRDT is infeasible, we have proposed the *randomized dummy traffic generation algorithm NED*. NED offers *parametrization to strike an arbitrary balance between power consumption and unobservability* of genuine communication. We have proven the privacy guarantees which NED can offer and have demonstrated that it can provide *reasonable privacy at a lower power consumption overhead than CRDT*. [Möl20]

5.1 Outlook

The contributions of this thesis constitute a strong base for further research into privacy in Home Automation Systems. While the proposed NED algorithm offers a way to tune privacy against power consumption, it is unlikely to offer optimal efficiency for a given set of privacy parameters. Further research in this area could provide helpful insights into common user behaviour in order to further reduce the power consumption overhead.

5.1.1 Learning the Underlying Distribution

It is clear that an optimal approach cannot be based on fixed parameters concerning the distribution of messages over time. In any HAS there will be longer periods of increased or decreased activity. For example, the owner of a newly installed HAS might be excited to try out many features, but might gradually interact less with the system. Different setups may also exhibit different distributions: As the number of installed devices increases, so does the expected number of messages in a given time frame. An algorithm which does not adapt to the environment is likely to either lose efficiency over time or degrade in performance with respect to the privacy goals.

NED can be readjusted and its performance can be continuously monitored by the system itself, as it can distinguish between dummy and genuine messages and therefore sample the values of the privacy parameters. However, this requires interaction of the user. Future approaches can adjust the dummy traffic generation based on the sampled parameters and reliably generate sequences which closely resemble genuine user interaction.

Large-scale sampling of user interaction and behaviour can be used to build accurate and general models of traffic distribution in order to tailor more efficient dummy traffic generation schemes.

5.1.2 Introducing Delays

NED is a latency-free algorithm. It does not delay genuine messages. However, as mentioned in Section 4.3, small delays are acceptable for most functionality and larger delays might be feasible for certain device classes such as thermostats. Shaping the traffic by introducing delays might simplify the construction of algorithms which provide stronger privacy guarantees than NED without significantly increasing (or possibly while decreasing) the power consumption overhead.

5.1.3 Application

We have demonstrated that privacy guarantees can be met by Home Automation Systems and that they can be tuned to reach a balance between privacy and energy efficiency. At the time of writing, no system is known to implement this mechanism. However, we hope that the results of this thesis will be transferred into practice to help improve and protect the privacy of Home Automation System users.

Bibliography

[Apt+19] Noah Apthorpe et al. "Keeping the Smart Home Private with Smart(er) IoT Traffic Shaping". In: *Proceedings on Privacy Enhancing Technologies* 2019.3 (2019 July 12), pp. 128–148. ISSN: 2299-0984. DOI: 10.2478/popets-2019-0040.

[BAC] *BACnet – A Data Communication Protocol for Building Automation and Control Networks*. ANSI/ASHRAE Standard 135-2016. 2016 Jan.

[Bag+14] Ibrahim Ethem Bagci et al. "Gathering Tamper Evidence in Wi-Fi Networks Based on Channel State Information". In: *Proceedings of the 2014 ACM Conference on Security and Privacy in Wireless & Mobile Networks – WiSec '14*. Oxford, United Kingdom: ACM Press, 2014 July 23, pp. 183–188. ISBN: 978-1-4503-2972-9. DOI: 10.1145/2627393.2627405.

[Ban+15] Subhadeep Banik et al. "Midori: A Block Cipher for Low Energy". In: *Advances in Cryptology – ASIACRYPT 2015. ASIACRYPT 2015*. International Conference on the Theory and Application of Cryptology and Information Security. Ed. by Tetsu Iwata and Jung Hee Cheon. Auckland, New Zealand: Springer, Berlin, Heidelberg, 2015 Nov. 29, pp. 411–436. DOI: 10.1007/978-3-662-48800-3_17.

[Bas+15] Boulat A. Bash et al. "Hiding Information in Noise: Fundamental Limits of Covert Wireless Communication". In: *IEEE Communications Magazine* 53.12 (2015 May 29), pp. 26–31. DOI: 10.1109/MCOM.2015.7355562. arXiv: 1506.00066.

[BDK01] Peter Bergstrom, Kevin Driscoll and John Kimball. "Making home automation communications secure". In: *Computer* 34.10 (2001), pp. 50–56. ISSN: 00189162. DOI: 10.1109/2.955099.

[BEN12] Calum Benson, Adam Elman and Seth Nickell. *GNOME Human Interface Guidelines 2.2.3*. 2012. URL: https://developer.gnome.org/hig-book/3.12/index.html.en (accessed 2020-09-02).

[Blu20] Christoph Blumtritt. *Smart Home Report 2020*. 2020 Aug. URL: https://www.statista.com/study/42112/smart-home-report/ (accessed 2020-09-03).

[Cai+12] Xiang Cai et al. "Touching from a Distance: Website Fingerprinting Attacks and Defenses". In: *Proceedings of the 2012 ACM conference on Computer and communications security – CCS '12*. Raleigh, North Carolina, USA: ACM Press, 2012 Oct., pp. 605–616. ISBN: 978-1-4503-1651-4. DOI: 10.1145/2382196.2382260.

[Cha81] David L. Chaum. "Untraceable electronic mail, return addresses, and digital pseudonyms". In: *Communications of the ACM* 24.2 (1981 Feb. 1), pp. 84–90. ISSN: 00010782. DOI: 10.1145/358549.358563.

[Chi+19] Haotian Chi et al. *PFirewall: Semantics-Aware Customizable Data Flow Control for Home Automation Systems*. 2019 Oct. 21. arXiv: 1910.07987. URL: http://arxiv.org/abs/1910.07987 (accessed 2020-11-08).

[Cop+16] Bogdan Copos et al. "Is Anybody Home? Inferring Activity From Smart Home Network Traffic". In: *2016 IEEE Security and Privacy Workshops (SPW)*. IEEE Security and Privacy. San Jose, California, USA: IEEE, 2016 May, pp. 245–251. ISBN: 978-1-5090-3690-5. DOI: 10.1109/SPW.2016.48.

[Cop17] Bogdan Copos. "Modeling Systems Using Side Channel Information". Dissertation. University of California Davis, 2017.

[CP03] Haowen Chan and Adrian Perrig. "Security and privacy in sensor networks". In: *Computer* 36.10 (2003 Oct.), pp. 103–105. ISSN: 00189162. DOI: 10.1109/MC.2003.1236475. pmid: 20960968.

[CWC13] Mauro Conti, Jeroen Willemsen and Bruno Crispo. "Providing Source Location Privacy in Wireless Sensor Networks: A Survey". In: *IEEE Communications Surveys & Tutorials* 15.3 (2013), pp. 1238–1280. ISSN: 1553-877X. DOI: 10.1109/SURV.2013.011413.00118.

[Das+18] Debajyoti Das et al. "Anonymity Trilemma: Strong Anonymity, Low Bandwidth Overhead, Low Latency – Choose Two". In: *2018 IEEE Symposium on Security and Privacy*. IEEE, 2018 July 26, pp. 108–126. ISBN: 978-1-5386-4353-2. DOI: 10.1109/SP.2018.00011.

[DH76] Whitfield Diffie and Martin Hellman. "New directions in cryptography". In: *IEEE transactions on Information Theory* 22.6 (1976), pp. 644–654. ISSN: 1557-9654. DOI: 10.1109/TIT.1976.1055638.

[Die+18] Constanze Dietrich et al. "Investigating System Operators' Perspective on Security Misconfigurations". In: *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. CCS '18: 2018 ACM SIGSAC Conference on Computer and Communications Security. Toronto Canada: ACM, 2018 Jan. 15, pp. 1272–1289. ISBN: 978-1-4503-5693-0. DOI: 10.1145/3243734.3243794.

[DMS04] Roger Dingledine, Nick Mathewson and Paul Syverson. "Tor: The second-generation onion router". In: *Proceedings of the 13th USENIX Security Symposium*. 13th USENIX Security Symposium. San Diego, CA: USENIX Association, 2004, pp. 303–320.

[Dwo+10] Cynthia Dwork et al. "Differential Privacy Under Continual Observation". In: *STOC '10 Proceedings of the forty-second ACM symposium on Theory of computing*. New York, New York, USA: ACM, 2010, pp. 715–724. ISBN: 978-1-4503-0050-6. DOI: 10.1145/1806689.1806787.

[Dwo06] Cynthia Dwork. "Differential Privacy". In: *Automata, Languages and Programming – 33rd International Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, Proceedings, Part II*. Ed. by Michele Bugliesi et al. Vol. 4052. Theoretical Computer Science and General Issues. Springer-Verlag Berlin Heidelberg, 2006, pp. 1–12. ISBN: 978-3-540-35907-4. DOI: 10.1007/11787006_1.

[Fis16] Thomas Fischer. *Strafgesetzbuch: StGB*. 63rd ed. C.H.BECK, 2016. ISBN: 978-3-406-68260-5.

[FN01] Laura Marie Feeney and Martin Nilsson. "Investigating the energy consumption of a wireless network interface in an ad hoc networking environment". In: *Proceedings IEEE INFOCOM 2001 – Conference on Computer Communications – Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies*. Vol. 3. Anchorage, Alaska, USA: IEEE, 2001, pp. 1548–1557. ISBN: 0-7803-7016-3. DOI: 10.1109/INFCOM.2001.916651.

[FY63] Ronald Aylmer Fisher and Frank Yates. *Statistical tables for biological, agricultural and medical research*. 6th ed. Edinburgh: Oliver and Boyd, 1963.

[GL14] Simson Garfinkel and Heather Richter Lipford. "Usable Security: History, Themes, and Challenges". In: *Synthesis Lectures on Information Security, Privacy, and Trust* 5.2 (2014 Sept. 20), pp. 1–124. ISSN: 1945-9742, 1945-9750. DOI: 10.2200/S00594ED1V01Y201408SPT011.

[Gol18] Peter Gola, ed. *Datenschutz-Grundverordnung*. 2nd ed. München: C.H.BECK, 2018.

[Gre+08] Ben Greenstein et al. "Improving wireless privacy with an identifier-free link layer protocol". In: *MobiSys'08 – Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services*. 2008. ISBN: 978-1-60558-139-2. DOI: 10.1145/1378600.1378607.

[Han+14] Jun Han et al. "Short paper: MVSec: secure and easy-to-use pairing of mobile devices with vehicles". In: *Proceedings of the 2014 ACM conference on Security and privacy in wireless & mobile networks – WiSec '14*. New York, New York, USA: ACM Press, 2014, pp. 51–56. ISBN: 978-1-4503-2972-9. DOI: 10.1145/2627393.2627400.

[Hel13] Andreas Hellmann. "Exploration und Evaluation von Sicherheitsaspekten drahtloser Hausautomationssysteme". Master's Thesis. University of Paderborn, 2013.

[JBC16] Andreas Jacobsson, Martin Boldt and Bengt Carlsson. "A risk analysis of a smart home automation system". In: *Future Generation Computer Systems* 56 (2016 Mar.), pp. 719–733. ISSN: 0167739X. DOI: 10.1016/j.future.2015.09.003.

[Jun+12] Markus Jung et al. "Privacy enabled Web service access control using SAML and XACML for home automation gateways". In: *2011 International Conference for Internet Technology and Secured Transactions*. 2011 International Conference for Internet Technology and Secured Transactions. Abu Dhabi, United Arab Emirates: IEEE, 2012 Feb. 9, pp. 584–591. ISBN: 978-1-908320-00-1.

[Kaa+17] Kim J. Kaaz et al. "Understanding user perceptions of privacy, and configuration challenges in home automation". In: *2017 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. 2017 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC). Raleigh, NC: IEEE, 2017 Oct., pp. 297–301. ISBN: 978-1-5386-0443-4. DOI: 10.1109/VLHCC.2017.8103482.

[Kes04] Gary C. Kessler. "An Overview of Steganography for the Computer Forensics Examiner". In: *Forensic Science Communications* 6.3 (2004).

[Kir+08] Csaba Kiraly et al. "Traffic Flow Confidentiality in IPsec: Protocol and Implementation". In: *The Future of Identity in the Information Society*. Third IFIP WG 9.2, 9.6/11.6, 11.7/FIDIS International Summer School on The Future of Identity in the Information Society. Boston, MA: Springer US, 2008, pp. 311–324. ISBN: 978-0-387-79026-8. DOI: 10.1007/978-0-387-79026-8_22.

[KM16] Jochen Krüger and Frederik Möllers. "Metadaten – eine neue juristische Problemkategorie im Rahmen der elektronischen Aktenführung?" In: *MultiMedia und Recht* 19.11 (2016), pp. 728–731. ISSN: 1434-596X.

[KNX] *Information technology – Home electronic systems (HES) architecture – Part 3-1: Communication layers – Application layer for network based control of HES Class 1.* ISO/IEC standard 14543-3-1:2006. 2006 Sept.

[Kol33] Andrei Nikolajewitsch Kolmogorow. "Sulla Determinazione Empirica di una Legge di Distributione". In: *Giornale dell'Istituto Italiano degli Attuari* 4 (1933), pp. 1–11.

[Lei+16] Dominik Leibenger et al. "Privacy Challenges in the Quantified Self Movement – An EU Perspective". In: *Proceedings on Privacy Enhancing Technologies* 2016.4 (2016), pp. 315–334. ISSN: 2299-0984. DOI: 10.1515/popets-2016-0042.

[Leu+18] Patrick Leu et al. "I Send, Therefore I Leak: Information Leakage in Low-Power Wide Area Networks". In: *Proceedings of the 11th ACM Conference on Security \& Privacy in Wireless and Mobile Networks*. 2018, pp. 23–33. ISBN: 978-1-4503-5731-9. DOI: 10.1145/3212480.3212508.

[Lev+04] Brian N. Levine et al. "Timing Attacks in Low-Latency Mix Systems". In: *Financial Cryptography – 8th International Conference, FC 2004, Key West, FL, USA, February 9-12, 2004. Revised Papers*. Ed. by Ari Juels. Vol. 1. Lecture Notes in Computer Science. Germany: Springer-Verlag Berlin Heidelberg, 2004, pp. 251–265. ISBN: 978-3-540-22420-4. DOI: 10.1007/978-3-540-27809-2_25.

[Mat+08] Alfredo Matos et al. "Toward dependable networking: secure location and privacy at the link layer". In: *IEEE Wireless Communications* 15.5 (2008 Oct.), pp. 30–36. ISSN: 1536-1284. DOI: 10.1109/MWC.2008.4653129.

[MDG14] Thomas Mundt, Andreas Dähn and Hans-Walter Glock. "Forensic analysis of home automation systems". In: *7th Workshop on Hot Topics in Privacy Enhancing Technologies (HotPETs 2014)*. 2014.

[Mey+16] Dominik Meyer et al. "A threat-model for building and home automation". In: *2016 IEEE 14th International Conference on Industrial Informatics (INDIN)*. 2016 IEEE 14th International Conference on Industrial Informatics (INDIN). Poitiers, France: IEEE, 2016 July, pp. 860–866. ISBN: 978-1-5090-2870-2. DOI: 10.1109/INDIN.2016.7819280.

[Mil68] Robert B. Miller. "Response time in man-computer conversational transactions". In: *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*. San Francisco, USA: ACM New York, 1968, pp. 267–277. DOI: 10.1145/1476589.1476628.

[Möl+14] Frederik Möllers et al. “Extrapolation and Prediction of User Behaviour from Wireless Home Automation Communication”. In: *7th ACM Conference on Security and Privacy in Wireless and Mobile Networks (ACM WiSec 2014)*. ACM Conference on Security and Privacy in Wireless and Mobile Networks. Oxford, United Kingdom: ACM, 2014 July 23, pp. 195–200. ISBN: 978-1-4503-2972-9. DOI: 10.1145/2627393.2627407.

[Möl+17] Frederik Möllers et al. “Mit Schirm, Charme und Kamera – Verbotene Sendeanlagen i.S.d. § 90 TKG”. In: *Trends und Communities der Rechtsinformatik: Tagungsband des 20. Internationalen Rechtsinformatik Symposions IRIS 2017 / Trends and Communities of Legal Informatics: Proceedings of the 20th International Legal Informatics Symposium IRIS 2017*. Internationales Rechtsinformatik Symposion. Salzburg, Austria: Österreichische Computer Gesellschaft, 2017, pp. 683–690.

[Möl+18] Frederik Möllers et al. “Modelling Traffic Analysis in Home Automation Systems”. In: *Cryptology and Network Security: 16th International Conference, CANS 2017, Hong Kong, China, November 30–December 2, 2017, Revised Selected Papers*. Cryptology and Network Security. Ed. by Srđan Čapkun and Sherman S. M. Chow. Vol. 11261. Lecture Notes in Computer Science. Hong Kong, China: Springer International Publishing, 2018 Nov. 10, pp. 526–536. ISBN: 978-3-030-02641-7. DOI: 10.1007/978-3-030-02641-7_27.

[Möl20] Frederik Möllers. “Energy-Efficient Dummy Traffic Generation for Home Automation Systems”. In: *Proceedings on Privacy Enhancing Technologies* 2020.4 (2020), pp. 376–393. ISSN: 2299-0984. DOI: 10.2478/popets-2020-0078.

[MS15] Frederik Möllers and Christoph Sorge. “Hausautomationssysteme im Datenschutzrecht”. In: *Kooperation: Tagungsband des 18. Internationalen Rechtsinformatik Symposions IRIS 2015 / Co-operation: Proceedings of the 18th Legal Informatics Symposium IRIS 2015*. Internationales Rechtsinformatik Symposion IRIS. Salzburg, Austria: Österreichische Computer Gesellschaft, 2015, pp. 553–558. ISBN: 978-3-85403-309-7.

[MS16] Frederik Möllers and Christoph Sorge. “Deducing User Presence from Inter-Message Intervals in Home Automation Systems”. In: *ICT Systems Security and Privacy Protection: 31st IFIP TC 11 International Conference, SEC 2016, Ghent, Belgium, May 30 - June 1, 2016, Proceedings*. ICT Systems Security and Privacy Protection. Ed. by Jaap-Henk Hoepman and Stefan Katzenbeisser. Vol. 471. IFIP Advances in Information and Communication Technology. Ghent, Belgium: Springer

International Publishing, 2016, pp. 369–383. ISBN: 978-3-319-33630-5. DOI: 10.1007/978-3-319-33630-5_25.

[MV16] Frederik Möllers and Stephanie Vogelgesang. "Smart-Home-Systeme in Zeiten digitaler Kriminalität". In: *Datenschutz und Datensicherheit – DuD* 40.8 (2016), pp. 497–502. ISSN: 1614-0702. DOI: 10.1007/s11623-016-0645-3.

[Nie93] Jakob Nielsen. *Usability engineering*. 1st ed. Morgan Kaufmann, 1993. ISBN: 978-0-12-518406-9.

[OTP14] Simon Oya, Carmela Troncoso and Fernando Pérez-González. "Do Dummies Pay Off? Limits of Dummy Traffic Protection in Anonymous Communications". In: *Privacy Enhancing Technologies SE – 11*. Ed. by Emiliano De Cristofaro and StevenJ. Murdoch. Vol. 8555. Lecture Notes in Computer Science. Springer International Publishing, 2014, pp. 204–223. ISBN: 978-3-319-08505-0. DOI: 10.1007/978-3-319-08506-7_11. URL: http://dx.doi.org/10.1007/978-3-319-08506-7_11.

[Pan+11] Andriy Panchenko et al. "Website Fingerprinting in Onion Routing BasedAnonymization Networks". In: *Proceedings of the 10th annual ACM workshop on Privacy in the electronic society – WPES '11*. New York, New York, USA: ACM Press, 2011, p. 103. ISBN: 978-1-4503-1002-4. DOI: 10.1145/2046556.2046570.

[Pea92] Karl Pearson. "On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is Such that it Can be Reasonably Supposed to have Arisen from Random Sampling". In: *Breakthroughs in Statistics*. Ed. by Samuel Kotz and Norman L. Johnson. Springer Series in Statistics. Springer New York, 1992, pp. 11–28. ISBN: 978-0-387-94039-7. DOI: 10.1007/978-1-4612-4380-9_2. URL: http://dx.doi.org/10.1007/978-1-4612-4380-9_2.

[Ped+11] Fabian Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[PH10] Andreas Pfitzmann and Marit Hansen. *A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management*. Version 0.34. 2010 Aug. URL: http://dud.inf.tu-dresden.de/literatur/Anon%5C%5C_Terminology%5C%5C_v0.34.pdf (accessed 2020-09-02).

[Pio+17] Ania M. Piotrowska et al. "The Loopix Anonymity System". In: *Proceedings of the 26th USENIX Security Symposium*. Vancouver, BC, Canada: USENIX Association, 2017, pp. 1199–1216.

[PK01] Andreas Pfitzmann and Marit Köhntopp. "Anonymity, Unobservability, and Pseudonymity – A Proposal for Terminology". In: *Designing Privacy Enhancing Technologies SE – 1*. Ed. by Hannes Federrath. Vol. 2009. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2001, pp. 1–9. ISBN: 978-3-540-41724-8. DOI: 10.1007/3-540-44702-4_1. URL: http://dx.doi.org/10.1007/3-540-44702-4_1.

[PPW91] Andreas Pfitzmann, Birgit Pfitzmann and Michael Waidner. "ISDN-Mixes: Untraceable Communication with Very Small Bandwidth Overhead". In: *Kommunikation in verteilten Systemen – Grundlagen, Anwendungen, Betrieb GI/ITG-Fachtagung, Mannheim, 20.–22. Februar 1991, Proceedings*. Ed. by Wolfgang Effelsberg, Hans W. Meuer and Günter Müller. Vol. 267. Informatik-Fachberichte. Germany: Springer-Verlag Berlin Heidelberg, 1991, pp. 451–463. ISBN: 978-3-540-53721-2. DOI: 10.1007/978-3-642-76462-2_32.

[PSC05] Joseph Polastre, Robert Szewczyk and David Culler. "Telos: enabling ultra-low power wireless research". In: *IPSN 2005. Fourth International Symposium on Information Processing in Sensor Networks, 2005.* Boise, ID, USA: IEEE, 2005, pp. 364–369. ISBN: 0-7803-9201-9. DOI: 10.1109/IPSN.2005.1440950.

[PSK78] Michael A. Padlipsky, David W. Snow and Paul A. Karger. *Limitations of end-to-end encryption in secure computer networks*. The MITRE Corporation, 1978 Aug. 1. URL: http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA059221 (accessed 2015-12-02).

[Sha+08] Min Shao et al. "Towards Statistically Strong Source Anonymity for Sensor Networks". In: *ACM Transactions on Sensor Networks (TOSN)* 9.3 (2008 Apr.), 34:1–34:23. ISSN: 1550-4859. DOI: 10.1145/2480730.2480737.

[SSP] *Bluetooth® Secure Simple Pairing Using NFC*. Application Document. 2014 Jan. 9.

[Sta13] Statistisches Bundesamt. *Wirtschaftsrechnungen, Fachserie 15, Sonderheft 1*. Wiesbaden, 2013, p. 19.

[SW06] Vitaly Shmatikov and Ming-Hsiu Wang. "Timing Analysis in Low-Latency Mix Networks: Attacks and Defenses". In: *Computer Security – ESORICS 2006 – 11th European Symposium on Research in Computer Security, Hamburg, Germany, September 18-20, 2006, Proceedings*. Ed. by Dieter Gollmann, Jan Meier and Andrei Sabelfeld. Vol. 4189. Security and Cryptology. Hamburg, Germany: Springer-Verlag Berlin Heidelberg, 2006, pp. 18–33. ISBN: 978-3-540-44601-9. DOI: 10.1007/11863908_2.

[TDG16] Raphael R. Toledo, George Danezis and Ian Goldberg. "Lower-Cost ε-Private Information Retrieval". In: *Proceedings on Privacy Enhancing Technologies* 2016.4 (2016), pp. 184–201. ISSN: 2299-0984. DOI: 10.1515/popets-2016-0035. arXiv: 1604.00223.

[Tex09] Texas Instruments. *Application Note AN047: CC1100/CC2500 – Wake-On-Radio*. Dallas, 2009.

[vDL03] Tijs van Dam and Koen Langendoen. "An Adaptive Energy-efficient MAC Protocol for Wireless Sensor Networks". In: *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems*. Los Angeles, California, USA: ACM, 2003, pp. 171–180. ISBN: 1-58113-707-9. DOI: 10.1145/958491.958512.

[VHM16] Stephanie Vogelgesang, Stefan Hessel and Frederik Möllers. "Hardware-Keylogger: Die Tastatur in der Hand des Feindes". In: *Datenschutz und Datensicherheit – DuD* 40.11 (2016), pp. 729–734. ISSN: 1614-0702. DOI: 10.1007/s11623-016-0692-9.

[VK83] Victor L. Voydock and Stephen T. Kent. "Security Mechanisms in High-Level Network Protocols". In: *ACM Computing Surveys* 15.2 (1983 June 1), pp. 135–171. DOI: 10.1145/356909.356913.

[VM16] Stephanie Vogelgesang and Frederik Möllers. "Personenbezug bei dynamischen IP-Adressen – Anmerkung zur Entscheidung des EuGH vom 19.10.2016". In: *Datenschutz-Berater: Informationsdienst der Verlagsgruppe Handelsblatt* 2016.12 (2016), pp. 233–235. ISSN: 0170-7256.

[Vog+17] Stephanie Vogelgesang et al. "Auf der Jagd nach Schwachstellen – Eine strafrechtliche Bewertung von Portscans". In: *Datenschutz und Datensicherheit – DuD* 41.8 (2017), pp. 501–506. ISSN: 1614-0702. DOI: 10.1007/s11623-017-0820-1.

[Vog16] Stephanie Vogelgesang. "Datenspeicherung in modernen Fahrzeugen – wem „gehören" die im Fahrzeug gespeicherten Daten?" In: *juris – Die Monatszeitschrift* 3.1 (2016), pp. 2–8.

[Wan+05] Arvinderpal S. Wander et al. "Energy Analysis of Public-Key Cryptography for Wireless Sensor Networks". In: *Third IEEE International Conference on Pervasive Computing and Communications*. IEEE, 2005, pp. 324–328. ISBN: 0-7695-2299-8. DOI: 10.1109/PERCOM.2005.18.

[Wei+15] Bruce D. Weinberg et al. "Internet of Things: Convenience vs. privacy and secrecy". In: *Business Horizons* 58.6 (2015 Nov.), pp. 615–624. ISSN: 00076813. DOI: 10.1016/j.bushor.2015.06.005.

[WiFi] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. IEEE, 2016. ISBN: 978-1-5044-3645-8.

[Yan+08] Yi Yang et al. "Towards Event Source Unobservability with Minimum Network Traffic in Sensor Networks". In: *Proceedings of the first ACM conference on Wireless network security – WiSec '08*. New York, New York, USA: ACM Press, 2008, pp. 77–88. ISBN: 978-1-59593-814-5. DOI: 10.1145/1352533.1352547.

[Zig] *Zigbee Specification*. ZigBee Document 05-3474-21. Zigbee Alliance, 2008, pp. 1–604.