Survey of Privacy Enabling Strategies in IoT Networks

Lukáš Hellebrandt
Brno University of Technology
Brno
Czech Republic
+420 54114 1179
ihelleb@fit.vutbr.cz

Ondřej Hujňák
Brno University of Technology
Brno
Czech Republic
+420 54114 1179
ihujnak@fit.vutbr.cz

Petr Hanáček
Brno University of Technology
Brno
Czech Republic
+420 54114 1216
hanacek@fit.vutbr.cz

Ivan Homoliak
Brno University of Technology
Brno
Czech Republic
+420 54114 1185
ihomoliak@fit.vutbr.cz

ABSTRACT

In this paper, we discuss privacy issues in modern networks for Internet of Things. We focus on anonymization of both devices and users in the context of both IP and non-IP networks.

We take a closer look on two current non-IP technologies – LoRaWan and ZigBee. Those represent two distinct groups of Internet of Things (IoT) networks – Low Power WANs covering large areas and providing connectivity as a service, and Wireless PANs following traditional scheme with a local network interconnecting IoT devices.

For both IP and non-IP networks we analyze possible approaches to preserve privacy of connected devices and identify open problems for future investigation. We propose strategies for ensuring privacy for IoT devices in IP, LPWAN and PAN networks based on their specific features and analyze possible problems of suggested strategies.

CCS Concepts

• **Security and privacy → Security services → Pseudonymity, anonymity and untraceability.**

Keywords

Privacy; Anonymity; IoT; Internet of Things; LoRa; ZigBee.

1. INTRODUCTION

With increasing usage of Internet of Things (IoT) devices, apart from other problems, there is the issue with their privacy. IoT devices deal with huge amounts of sensitive data [1] and users often do not realise the nature of the data they are sharing. Furthermore, the inference is very powerful, which in combination with the amount of information shared in IoT brings serious concerns. Through common attacks on privacy, as described in [1], adversary can collect information about the user leading to efficient surveillance including, but not limited to, real-time tracking of the user through connected IoT devices [2].

For achieving anonymity, anonymity networks are used. An anonymity network is a network of some nodes together with protocols for their communication which, through denying some information to the adversary, makes it hard for that adversary to decide which two parties are communicating.

There are multiple issues to address when trying to implement efficient anonymity of IoT devices such as number of connected devices, their diversity and energy constraints. The ever-increasing number of devices combined with multiple communication technologies used for last mile and huge differences between various IoT devices make implementing anonymization for those devices very challenging. Moreover, IoT devices need to have low energy consumption, so they offer relatively low computation power, which limits power demanding operations, such as cryptographical ones.

In this paper we try to address those issues and propose possible ways of achieving anonymity in the domain of IoT networks. We have divided those networks into two groups based on the area they are covering, as they differ significantly in technologies used. From each group we have chosen an example – LoRaWAN for Low-Power Wide Area Networks (LPWANs) that cover large areas, and ZigBee for Personal Area Networks (PANs) intended for much smaller networks. For both types of IoT networks we propose specific ways of ensuring privacy by anonymization and define which piece of information is concealed from whom.

First, we will cover some research that has been published in the area of security and privacy of IoT networks in Section 2. Then we will briefly introduce different networks used for interconnecting IoT devices in Section 3 and describe common anonymization techniques in Section 4. Our proposed anonymization approaches are stated in Section 5 and Section 6 that are further subdivided according to the network used. The last part, Section 7, concludes the paper and provides further research directions.

2. RELATED WORK

In this section we the cover current state of research in the area of privacy of IoT networks and devices. Over the past years privacy issues are receiving significant attention as the amount of exploitable information increases with the spread of sensors and use of clouds for storing vast amounts of information.

There are multiple ways of how to address privacy in IoT. Perera et al. [2] identifies general privacy concerns of IoT solutions and proposes set of legal methods including standardization and regulation to counter them. One of the concerns is identifiability based on fingerprinting and he suggests using anonymity technology such as Tor. On the other hand Yoshigoe et al. [3] examines a concrete solution, Samsung SmartThings, and identifies privacy issues with respect to it. He focuses on ensuring privacy between IoT gateway and cloud server and recommends using VPN and dummy traffic to achieve it.

CSAI 2017, December 5–7, 2017, Jakarta, Indonesia

ACM ISBN 978-1-4503-5392-2/17/12...$15.00
https://doi.org/10.1145/3168390.3168440

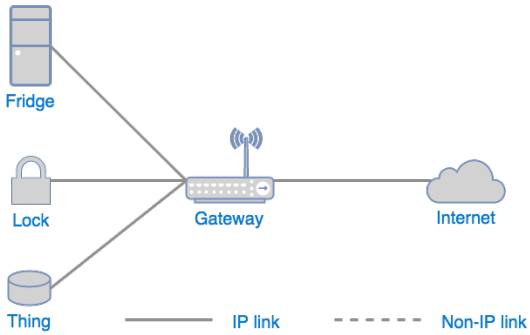

Figure 1. IP device network architecture

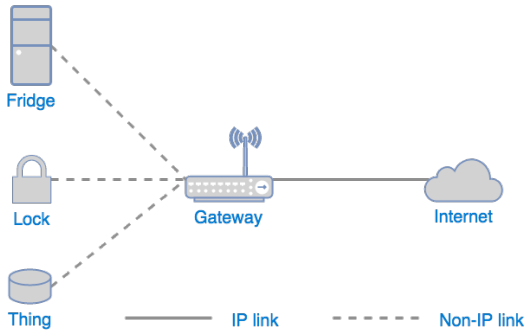

Figure 2. Non-IP device network architecture

Many applications of IoT focus on collecting data from various sensors owned by large number of participants and computes statistics based on those data. Yao et al. [4] notices that adversary with access to those data can gain sensitive information about participants and designs a new protocol for anonymous sensor data collection. This protocol ensures that it is impossible even for the Application Server to link data with participants.

Even though the privacy issue is recognised, it is not receiving enough attention when developing IoT protocols. And even in the case of including privacy features in the protocol, they are often not used or not strong enough, which was shown for example, by Das et al. [5] on the Bluetooth protocol. Using VPN for securing connection between gateway and cloud server increases security of the connection, but it might still not stand against a powerful adversary. For those reasons we have decided to focus on use of anonymity networks in IoT.

3. BACKGROUND ON IOT DEVICES

The definition of IoT is not yet settled and some experts consider as IoT only such devices that communicate via specially developed networks, while others include IP devices as well. The main difference is that, because the Internet network uses IP for communication, while IP devices can connect directly to the Internet, non-IP devices require a gateway to make translation between their protocol and IP. Moreover, some protocols do not allow direct connection or connection through a gateway, and require using cloud server for every communication with the device. To cover wide range of possible options ensuring privacy in IoT devices we decided to include both IP and non-IP devices.

3.1 IP IoT Devices

These are devices powerful enough to have their own IP network interface for communication with the outer world and are connected directly through IP networks such as Ethernet or Wi-Fi (see Figure 1). Communication with these devices does not differ from common Internet traffic and such devices can be seen as lightweight servers providing its services.

3.2 Non-IP IoT Devices

These are devices using one of the specific (often proprietary) network protocols for reasons of computational power, power consumption, or specific use. These protocols are not compatible with the IP protocol and to communicate with the Internet they need to utilize a special gateway that translates between the protocol used and IP as depicted in Figure 2.

There are two distinct types of IoT networks:

- **Low-Power Wide Area Networks (LPWANs)** that are intended for covering very large areas such as countries. This type of networks consists usually of some base stations spread in the locality and concentrating cloud server that provides access to network services. The infrastructure is often owned by some service provider who sells connectivity. Examples of these networks are Narrowband, LoRaWAN and SigFox.
- **Personal Area Networks (PANs)** are local networks intended for one building — such as a smart home. They usually consist of some nodes and one gateway connecting the network to the Internet. The whole infrastructure is usually owned by the user and devices may or may not connect to some cloud service on the Internet. This type includes networks such as Bluetooth, ZigBee or Z-Wave.

We decided to use LoRaWAN as a representative of LPWAN and ZigBee as a representative of PAN because both are open and well documented protocols created specifically for IoT.

3.2.1 LoRaWan

LoRaWAN is a LPWAN specification intended for wireless battery operated Things in a regional, national or global network. High level architecture of LoRaWAN (Figure 3) consists of:

- **Devices** – the Things in the network
- **Gateways** – points through which the Devices are connected to the server. One of their interfaces is a wireless interface using Lora RF for connection to Devices, another is an Internet connection for connection to the Server.
- **Network Server** – a point with which all the Gateways communicate in order to transfer some data between Devices and Applications.
- **Applications** – services running on an Application Server implementing some functionality.

Whatever required functionality is implemented via the applications. Inputs of these applications are data sent by the devices and actions taken by the devices based on the applications' instructions. The applications run on the server that is interconnected with the devices through *multiple* gateways.

The Device that wants to communicate sends a *broadcast* message that can be received by *any* of the gateways that might receive it. All the receiving gateways then forward the message to the server. This allows scalability, redundancy and also localization based on known positions of the gateways that received the broadcast.

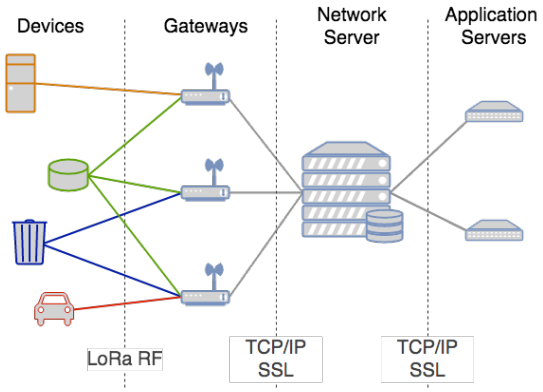

Figure 3. LoRaWAN architecture

The application acts based on data received from the devices, potentially sending them some instructions back through gateways.

3.2.2 ZigBee

ZigBee was developed to serve as a cheap and low-power-consumption way of interconnecting the IoT devices. The ZigBee network consists of:

- **End Device** – a device that has capability to send data to a Router. It does not receive nor transmit data continuously thus it has very low energy consumption.
- **Router** – a device that has all the capabilities of an End Device and additionally relays the data it receives to other devices. It can receive data continuously.
- **Coordinator** – a device that has all the capabilities of a Router and processes, stores and maintains information about the network. There is exactly one Coordinator in a ZigBee network.

As opposed to LoRaWAN, ZigBee has a classic network topology similar to Local Area Networks (LANs). Every device is connected to certain other devices in a relatively stable topology. There are three topologies supported in ZigBee: Star, Cluster Tree and Mesh (see Figure 4) and the network does support relocation of nodes within the topology. Devices can communicate both by broadcast messages readable by all devices in a network or by unicast messages that are encrypted so they can by decrypted only by the recipient.

For communication with the outer world over the Internet, there is a ZigBee Gateway. It is a ZigBee device with the capability of a Coordinator that has an IP connection to the Internet and thus can convert between ZigBee protocol and the Internet protocol transparently [6] and route properly.

4. BACKGROUND ON ANONYMIZATION IN THE CONTEXT OF IOT

In the current section we discuss basics of anonymization of communications using anonymity networks. This is a well-known problem for normal (IP) devices and there are existing, although limited, solutions. In the following, we discuss high level background on two possible options for ensuring anonymity in IoT devices.

There are multiple existing types of anonymity networks. Most anonymity networks are based on mixing - the way of achieving

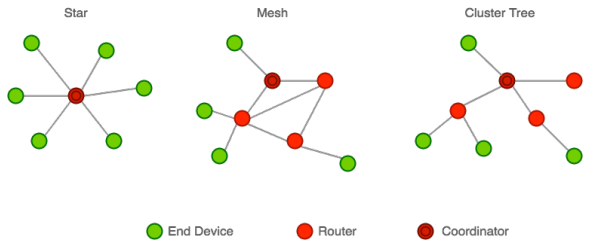

Figure 4. ZigBee topologies

anonymity through changing order and delaying of messages received by the device used for mixing – a mix. Mixes can be used in a network called mixnet [7]. There are also other methods such as reading a shared memory anonymously by accessing more than just the desired part. Using these strategies, the attacker cannot decide who has been communicating with whom.

For our purposes, we consider one of the options that is using some existing anonymity networks with necessary tweaks and directing the traffic of the IoT devices through this anonymity network. Important information when evaluating anonymity network is its threat model, anonymity set size, whether we need to trust parts of the anonymity network itself, or the network's latency. The most widespread anonymity network is Tor [8], which is a circuit-based mixnet. Through numerous layers of encryption of data including routing information, Tor allows us to create a circuit consisting of multiple relays that decrypt one layer at a time and using the decrypted routing information, those relays forward the message to the next relay. In this way, every relay only knows the previous and the next relay and while there exist some (mainly probabilistic) attacks, this is considered reasonably safe against non-global adversaries, who can, for example, carry on a successful traffic confirmation attack that is possible due to the fact that Tor is a low-latency network). Also, users of Tor do not need to trust all relays – it is sufficient that some of the relays in a circuit is honest. While there are other existing anonymity networks (with potentially better anonymity level), a crucial metric influencing anonymity level is anonymity set site – that means, how many users the adversary can pinpoint as possible origin of the message. Thus, large user base is important and for this reason, we consider Tor as an real option for an anonymization in the context of proposed techniques of this work.

Another option for ensuring anonymity is developing an entirely new anonymity network tailored especially for the needs of IoT devices. That means, taking into account IoT networks' limitations and possibly using their properties for a better or cheaper anonymity. Examples of such limitations are energy and computation resources consumption or perhaps more sensitive nature of transmitted data. On the other hand, large number of IoT devices might be useful for achieving better anonymity due to larger anonymity set.

It is worth noting that anonymity networks protect us from disclosing our identity and location on network level. Nevertheless, it is still possible to get deanonymized based on the data we send – if we, for example, send our exact location (based on GPS) as part of the data, we have effectively disclosed our location despite using an anonymity network.

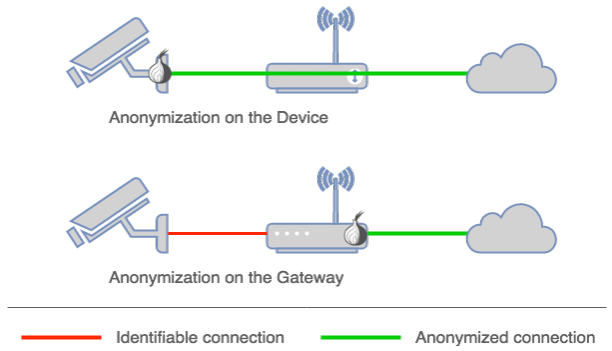

Figure 5. Different anonymization approaches for IP networks

5. ANONYMIZATION OF IP DEVICES

IoT devices with implemented IP stack interact with the network in the same manner as every other IP device. As anonymity networks for IP devices are already developed and in use, we can reuse these solutions in the domain of IoT as well. If the device is powerful enough, a client executing anonymization can be run directly on the device, otherwise it can run on the gateway between the device and the Internet. Those two approaches are described in the following sections and depicted in Figure 5.

5.1 Anonymization on the Device

If the device can run its own anonymity network client, it is possible to implement such client as a part of its software with possible hardware support, e.g. for encryption operations. This approach has a great advantage of all the data sent from the device being already anonymized. However, this does not seem to be necessary in most cases as our primary goal is not achieving anonymization against adversary who already has access/control of our LAN, but anonymize the traffic inside our network against adversaries who has not yet gained the access into it.

Another argument against embedding the anonymization into a device is the lack of generality. We would have to implement support for anonymization network into each device, either as a part of its firmware or as an additional module. We would also need to make sure the device indeed uses only the anonymized connection and does not disclose its identity and/or location by other means. Example of such revelation could be downloading firmware update directly (not through anonymity network) or disclosing the information in the readable data as mentioned in Section 4.

5.2 Anonymization on the Gateway

There are also devices with IP network interface but not powerful enough to run an anonymity network client, perhaps because of energy consumption and computational power requirements of cryptographical operations. Such devices devices are, however, usually connected to the Internet through an IP gateway that is connected to the Internet (Figure 1). This gateway can also serve as an "anonymization box" — a gateway that sends all the traffic directed to the Internet through an anonymity network client, while anonymizing the communication. Every device or even every connection (in the case of TCP) could have its own anonymous identity ("circuit", in case of Tor).

There already exist such anonymization gateways for the Tor network [9]. These devices can be reused in the case of IoT devices with IP network interface – this would be transparent as the gateway would not even know whether the device is IoT-based and would just anonymize its communication as if it would be a regular IP device (e.g. a computer). This solution is advantageous because there are implementations of such gateways already present, and mainly because this would not require any change to the IoT devices themselves, thus being much more universal and scalable.

Concerns, however, are that firstly the data travel unanonymized through LAN and secondly this approach, without any further improvements, only works for gateways we trust. We can usually trust the gateway that is fully under our control, however, this is a big issue for public gateways. And indeed, we also need to have public anonymization gateways as there are IoT devices that are expected to move, for example wearables. Creating a protocol that would ensure anonymization by an untrusted gateway (or at least detection of its malicious behavior) is a problem not resolved yet.

Another problem of this approach is that the gateway must be used to connect to the Internet by the IoT device. Not all devices will be connected through an anonymization gateway. If the device is not connected through it, it is not anonymous and it only has limited means of detecting whether it is being anonymized or not. Creating a protocol that would ensure the device is being anonymized — either through communication with the gateway or through some heuristics, like hop count — is an open problem.

6. ANONYMIZATION OF NON-IP DEVICES

More challenging task is to anonymize devices with network interfaces other than IP. The most spread type of non-IP devices that may benefit of anonymization are IoT devices and their networks, so we will look deeper into the two networks we have chosen: LoRaWAN and ZigBee. Each of these need a different approach to anonymization and it is also crucial to define *who* we are anonymizing *against whom*.

6.1 Proposed Anonymization for LoRaWAN Devices

For the purpose of LoRaWAN, we consider "anonymization" as making the server unaware of whose device has sent data from where or whose device has received data. That means the Server only knows it has received data, the data were processed by some spplication and possibly that application has sent some instruction to another unknown device.

The part where we deny the information about *the owner* of the device seems straight-forward. If the user registers his devices and the applications anonymously, it seems impossible to identify him or even get his IP address – because he does not have any. It is, however, still possible to assign multiple instances of communication to the same anonymous identity. This seems to be unavoidable because devices need to be authenticated in the LoRaWAN network.

A much harder challenge (and a problem to be resolved) is hiding the device's *location*. Each gateway informs the server of its own location. Therefore, when the server receives some data through a certain gateway, it can easily conclude that the device is somewhere within its transmission range from the gateway. It seems like an easy task to use own gateway that would not send its location or fake it, therefore hiding the device's location as well. There are, however, at least two ways the device's location may be discovered (depicted in Figure 6):

- Another gateway receives the broadcast from the device (green path in Figure 6). This Gateway would also relay the data, however, with correct location information. We are not aware of any way of secure unicast to the

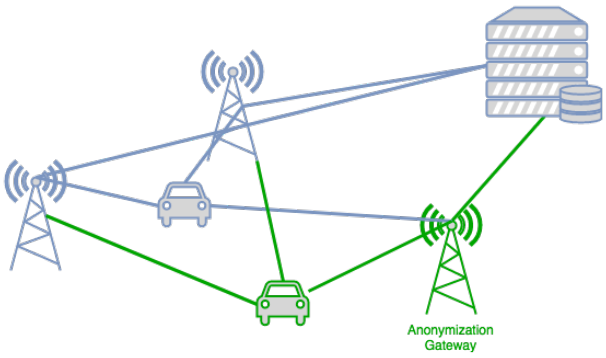

Figure 6. Anonymization gateway in LoRaWAN ecosystem

gateway. We would need to change the protocol for communication between the gateway and the device – so, we would need to make changes to both of them and violate the standard) to create a secure encrypted channel with our trusted gateway only.
The encryption is essential to ensure that gateways not under our control do not understand and relay the messages not intended for them, because such gateways cannot be trusted to follow our modifications. For this reason not only payload, but even headers have to be encrypted.

- Our gateway receives broadcast from another device not under our control (blue line in Figure 6). This device's broadcast is also received by at least one other gateway. If both gateways relay this data to the server as expected, the server can infer the gateway's position by knowing it is somewhere near the other gateway position of which the server knows. This would disclose location of all the devices using our gateway including our devices. A fix to this seems to be straightforward – our gateway should not relay traffic from any other device than ours.

The first point seems problematic mainly due to required changes to the device, the gateway and protocol used by these, and therefore needs further investigation. Also note that one of the most advertised LoRaWAN's advantages is the possibility of device localization.

6.2 Concealing the Location of LoRaWAN Gateways

Another approach to anonymity in LoRaWAN can be achieved by hiding the location of the gateway itself. A typical use case is a user who wants to use the devices that use LoRaWAN but his area is not covered by LoRaWAN yet. He therefore decides to create his own gateway. He, however, does not want the server to know the location of the area in question.

As the communication between the gateway and the server is through the Internet, which is the IP network, we could use one of the solutions shown in Section 5. However, while this approach would hide the gateway's IP address, the location may be revealed the same way as mentioned in Section 6.1 (through other gateways receiving the same message). After all, part of anonymizing the device's location seems to be anonymizing locations of all the gateways it uses.

6.3 Proposed Anonymization of ZigBee Devices

For the purpose of ZigBee, we understand "anonymization" as making whatever party on the Internet, with whom the device that is part of a user's network is communicating with, unaware of who is its owner and where the device is located.

As every device has to route all its data sent to the Internet through the gateway, we can use the gateway for running an anonymity network client. This way the gateway would actually anonymize the end device by anonymizing itself, as the IP packet originates from the gateway. If the device itself does not leak any location or identity related information as a part of the data sent, this seems to be secure.

One concern is whether this architecture cannot be compromised in the same way as described for LoRaWAN in section 6.1. We conjecture that it cannot happen, as the ZigBee is protocol intended for a very small area and no other gateway should even be within the transmit range of the device. Even if such a gateway would be in a transmit range, there are unicast links in ZigBee topology – the data is not broadcast to be understood by all (which is ensured by encryption). Also, ciphertext comparison (as described in section 6.1) will not work – a gateway changes ciphertext appearance by decrypting the data received from the device. Even though the routers and coordinator relay the data to other devices (so there may be other routers or a coordinator on a way between the communicating end device and gateway) it does not seem to be a problem as long as the adversary's device is not part of the user's topology.

As we have effectively reduced the task of anonymization in ZigBee to the same task as in section 5 – because the gateway is an IP device as any other – we can also opt for not changing the ZigBee gateway itself but rather use a cascade consisting of ZigBee gateway and anonymization gateway similar to an architecture described in Section 5.2.

6.4 Anonymization in Public Networks

With high proliferation of IoT devices used on the go such as wearables and sensors on cars, and the expected expansion of smart cities, there is urgent need for anonymization. In this environment, public gateways are expected for IoT devices (such as public WiFi access points) through which those devices would send their data.

If the privacy will not be a part of the network protocols used, privacy could be ensured by the public gateways. Because such Gateways are managed by some third party, they cannot be trusted and it would be a very advantageous property if the device would be able to discover whether it is actually anonymized. Protocol to achieve such feature is a future work.

7. CONCLUSION AND FUTURE WORK

Because in the era of IoT, the amount of data collected from various sensors shall increase dramatically and the nature of the data is often sensitive, we have decided to investigate ways of ensuring privacy in IoT networks. We have briefly summarized approaches used in IP networks with focus on anonymity networks. Then, we have divided IoT networks into two groups according to their intended size and use, and explained features of those networks on two examples – ZigBee and LoRaWAN.

We have identified the need to anonymize the Things in the IoT and divided the problem into multiple cases based on the network protocol used: IP and non-IP. We have proposed two approaches how to anonymize IP devices based on their capabilities. For IoT WANs we have suggested ways to anonymize both the devices connected to such networks and the gateways forming the network itself, and described issues that can occur. In PANs we reduced the privacy problem to the same as the one within the IP devices and we emphasized the issues arising from using public networks for Internet connection.

Finally, we pinpointed some problems that we plan to deal with in our future work. Especially finding a way to hide the location of a device in IoT WANs and reliably identify on the go whether the device is anonymized or not are interesting challenges that deserve closer examination.

8. ACKNOWLEDGEMENT

This article was created within the project Reliability and Security in IT (FIT-S-17-4014).

9. REFERENCES

[1] Abomhara, M., Køien, G. M. 2014. *Security and privacy in the Internet of Things: Current status and open issues.* International Conference on Privacy and Security in Mobile Systems (PRISMS). pp. 1–8.

[2] Perera, C., Ranjan, R., Wang, L., Khan, S. U., Zomaya, A. Y. 2015. *Big data privacy in the internet of things era*. IT Professional, vol. 17. pp. 32-39.

[3] Yoshigoe, K., Dai, W., Abramson, M., Jacobs, A. 2015. *Overcoming invasion of privacy in smart home environment with synthetic packet injection*. TRON Symposium (TRONSHOW), IEEE, 2015, pp. 1–7.

[4] Yao, Y., Yang, L. T., Xiong, N. N. 2015. *Anonymity-based privacy- preserving data reporting for participatory sensing*. IEEE Internet of Things Journal, vol. 2, no. 5, pp. 381–390

[5] Das, A. K., Pathak, P. H., Chuah., C. N., Mohapatra, P. 2016. *Uncovering Privacy Leakage in BLE Network Traffic of Wearable Fitness Trackers*. Proceedings of the 17th International Workshop on Mobile Computing Systems and Applications, ser. HotMobile '16, pp 99-104.

[6] Song, Q. D. X. S. Y., Chen, D. 2007. *The Design and Realization of Embedded Zigbee Gateway based on ARM9*. Microcomputer Information. vol. 35. page 066.

[7] Jakobsson, M., Juels, A., Rivest, R. L. 2002. *Making mix nets robust for electronic voting by randomized partial checking*. USENIX security symposium. San Francisco, USA. pp. 339–353.

[8] Dingledine, R., Mathewson, N., Syverson, P. 2004. *Tor: The second-generation onion router.* Technical Report. Naval Research Lab Washington DC.

[9] Herrmann, D., Lindemann, J., Zimmer, E., et al. 2015. *Anonymity Online for Everyone: What is missing for zero-e ort privacy on the Internet?* International Workshop on Open Problems in Network Security. Springer. pp. 82–94.

Authors' background

Name	Email	Position (Prof , Assoc. Prof. etc.)	Research Field
Lukáš Hellebrandt	**ihelleb@fit.vutbr.cz**	**PhD candidate**	**Anonymity networks**
Ondřej Hujňák	**ihujnak@fit.vutbr.cz**	**PhD candidate**	**Security of IoT networks**
Petr Hanáček	**hanacek@fit.vutbr.cz**	**Prof.**	**Information security**
Ivan Homoliak	**ihomoliak@fit.vutbr.cz**	**Researcher**	**Network security**