Blockchain-based Consents Management for Personal Data Processing in the IoT Ecosystem

Konstantinos Rantos[1], George Drosatos[2], Konstantinos Demertzis[1], Christos Ilioudis[3] and Alexandros Papanikolaou[3]

[1]*Dept. of Computer and Informatics Engineering, Eastern Macedonia and Thrace Institute of Technology, Kavala, Greece*

[2]*Dept. of Electrical and Computer Engineering, Democritus University of Thrace, Xanthi, Greece*

[3]*Dept. of Information Technology, Alexander Technological Educational Institute of Thessaloniki, Thessaloniki, Greece*

Keywords: Privacy, Internet of Things, GDPR, Digital Consents Management, Blockchain, Ontology Matching.

Abstract: In the Internet of Things (IoT) ecosystem the volume of data generated by devices in the user's environment is constantly increasing and becoming of particular value. In such an environment the average user is bound to face considerable difficulties in understanding the size and scope of his/her collected data. However, the provisions of the European General Data Protection Regulation (GDPR) require data subjects to be able to control their personal data, be informed and consent to its processing in an intelligible manner. This paper proposes ADVOCATE, a framework that facilitates GDPR-compliant processing of personal data in IoT environments. The present work aims to assist stakeholders, i.e. Data Controllers and Processors, satisfy GDPR requirements, such as informing data subjects in a transparent and unambiguous manner about the data they will manage, the processing purposes and periods. Respectively, data subjects will be promptly and comprehensively informed about any processing requests addressed to them, create and edit processing policies, exercise their rights in access, correction, deletion, restriction and opposition to data processing. Simultaneously, a notary service using blockchain infrastructures will ensure consents' security and an intelligent service will inform data subjects about the quality of their consents.

1 INTRODUCTION

The rapid and sometimes uncontrolled growth of the Internet of Things (IoT) threatens to a great extent the users' privacy. Soon, users will be surrounded by a significant number of devices bearing sensors and actuators (according to some reports there will be more than 75 billion such devices by 2025 (Lucero, 2016)). These devices will collect data that can be used to monitor users and create user profiles, with or without their consent. Moreover, they can make automated decision with questionable, if no at all, technical and organisational measures to protect user rights and freedoms (Roman et al., 2011; Mendez et al., 2017). It is obvious that most of the users will not be able to cope with this vast amount of data, sufficiently understand the scope of the data collected and the different processing methods, and have control over their personal data in accordance with the requirements of the General Data Protection Regulation – GDPR (European Parliament and Council, 2016).

Although extensive research work has been carried out in the field of user-centric privacy solutions there seems to be a wide space of privacy issues that need to be investigated in the IoT ecosystem (Sicari et al., 2015), not to mention those that also consider the GDPR requirements. The latter aims to protect natural persons with regards to the processing and movement of their personal data. In this context, GDPR defines the conditions under which data processing is lawful which include processing (a) on the basis of a data subject's consent, (b) for the needs of a contract, (c) for compliance with a legal obligation, (d) to protect the vital interests of the data subject, (e) for the needs of the public interest and (f) for the legitimate interests of the controller.

In the IoT ecosystem, and especially in segments like smart health and smart homes, processing is mostly accomplished on the grounds of users' consents and for protecting the vital interests of the data subject. While for the latter condition there is no need to interact with the user, consents should be granted only as a result of a declaration provided by the data controller in an easily accessible manner for the data

Rantos, K., Drosatos, G., Demertzis, K., Ilioudis, C. and Papanikolaou, A.
Blockchain-based Consents Management for Personal Data Processing in the IoT Ecosystem.
In *Proceedings of the 15th International Joint Conference on e-Business and Telecommunications (ICETE 2018) - Volume 2: SECRYPT*, pages 572-577
ISBN: 978-989-758-319-3

subject. The data controller should use clear and plain language and allow separate consents to be given to different data processing operations. Even if consent is not required though, data subjects have the right to know about any personal data processing and the legal grounds of it.

ADVOCATE aims to provide an environment that protects the privacy of users in the IoT ecosystem, in line with the GDPR requirements. In particular, following a user-centric approach to the development of IoT solutions (Roman et al., 2013), this framework covers the main principles of GDPR according to which data controllers will, among others, be able to inform the users in a transparent and unambiguous manner about any information relevant to their personal data.

2 RELATED WORK

The need to provide users with the ability to control their personal data generated by smart devices in their environment is widely recognised (Russell et al., 2015). The European Research Cluster on the Internet of Things (IERC) also highlights this need with an additional emphasis on the GDPR (IERC, 2015). However, preserving user privacy in the IoT is not an easy task (Zhang et al., 2014). Several research efforts have been made in the direction of developing suitable protocols for security and privacy in the IoT, since this is an area that currently attracts a significant amount of research. One of the ways to deal with the aforementioned challenges is to devise appropriate ways for applying policy management access control, which has been identified as an important research opportunity by IERC as well as by other researchers in the field (Stankovic, 2014; Sicari et al., 2015).

Some indicative examples are presented below that aim at providing solutions within this scope. However, they focus more on satisfying certain requirements of the GDPR and particularly the need for getting the user consent prior to any data processing. The framework proposed in (Cha et al., 2018), allows users to set their privacy preferences for the IoT devices they interact with. Additionally, Blockchain technology is employed to both protect and manage the privacy preferences that each user of the system has set, thus ensuring that no sensitive data has been accessed without their consent. The use of Blockchain gateways is also proposed in (Cha et al., 2017), where the setup is tailored for use with IoT scenarios. Good practices to be considered for obtaining user consent for IoT applications in the healthcare domain are also proposed in (O'Connor et al., 2017).

ADVOCATE addresses the challenges related to privacy protection in the IoT, especially with regards to the management of consents as GDPR requires and tries to close a significant gap in this area. It allows users manage their consents and formulate their personal data disposal policies considering the corresponding systems recommendations. Similarly, it provides data controllers with a useful tool for being GDPR-compliant.

3 PROPOSED ARCHITECTURE

The ADVOCATE approach concerns a set of sensors in the user environment that collect data related to the data subject. Such environments are a smart home, a patient health monitoring system, or activity monitoring sensors. The use of a portable device, such as a mobile phone, provides a user-friendly environment for data subjects to interact and manage their personal data disposal policy and their consents. It also provides a way for data controllers to interact with data subjects and obtain the necessary consents. The proposed architecture focusing, for the sake of simplicity, on smart cities and health ecosystems, is depicted in Figure 1. ADVOCATE, visualised as a cloud services platform, consists of the functional components described in the following sections. The main ADVOCATE components are analysed in the following sections while their role in consents management is depicted in Figure 2.

3.1 Consent Management Component

At the core of ADVOCATE is the *consent management* component, responsible for managing users' personal data disposal policies and the corresponding consents, including generation, updates and withdrawals. Utilising this component, data subjects are able to generate generic, domain-specific, or context-based privacy policies comprising a set of rules that correspond to data subjects' consents. The latter are the result of requests placed by data controllers for access to specific IoT data, for definite processing purposes and periods, in line with the GDPR requirements.

Having an interoperable mechanism for placing requests, granting permissions and formulating policies requires relying on common rules for defining sets of data, as well as the meaning and the use thereof. The foundations of this mechanism are data privacy ontologies which ensure that data controller

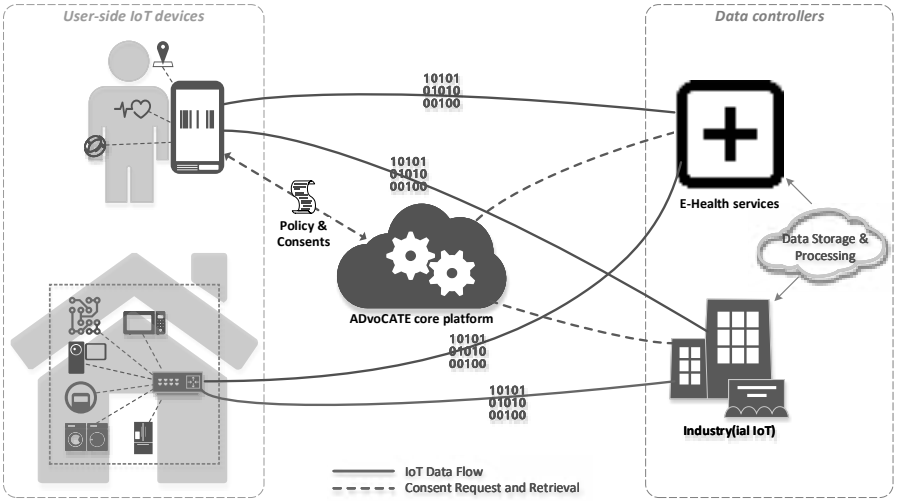

Figure 1: ADVOCATE infrastructure.

access requests are projected to data subjects in a unified and unambiguous manner.

The use of ontologies solves in principle the IoT device heterogeneity problems, contributing to the safety and privacy of users (Mozzaquatro et al., 2015). Ontologies can also provide a unified description for solving the problem of semantic heterogeneity in the field of Internet security with user-defined rules (Xu et al., 2017). Moreover, they contribute to the description of threats, attacks, impacts, controls and vulnerabilities and the definition of relationships among them, with many important advantages (Pereira and Santos, 2009). In ADVOCATE they will also facilitate intelligence policies analysis, as described in Section 3.3.

Several ontologies have been defined to support privacy and allow users to express access control rules for sharing data. The authors in (Passant et al., 2009) observe that protecting data does not merely mean granting access or not to the full data, but in most cases users require more fine-grained privacy preferences that define access privileges to specific data. The ontology that is close to the needs of ADVOCATE is proposed in (Bartolini et al., 2017). Although being a work in progress, the proposed model considers an early version of GDPR and defines an ontology to model data protection requirements. Adopting an appropriate ontology for data privacy will also facilitate specifying enforceable privacy policies. This implies that the description of policies is based on well-defined policy languages, such as the eXtensible Access Control Markup Language (XACML), to assist the decision making process. EnCoRe (EnCoRe Project, 2010) has also adopted XACML for enforcing policy based access control.

The adoption of a single ontology by ADVOCATE though, does not preclude the use of similar or competitive ontologies by participating data controllers. In that case ontology matching mechanisms have to be deployed to reduce semantic gaps between different overlapping representations of the same data privacy related information (Otero-Cerdeira et al., 2015). An ontology matching process typically uses several and different types of matchers such as labels, instances, and taxonomy structures to identify and calculate the similarity among ontologies. Ontology matching methods based on machine learning have been proved to provide more accurate and reliable matching results (Eckert et al., 2009).

In ADVOCATE, we can use ontology matching based on a semi-supervised learning approach. Given a small set of validated matching entity pairs, the method initially exploits the dominant relations in the similarity space to enrich positive training examples. After getting more training examples, a graph-based semi-supervised learning algorithm is employed to classify the remaining candidate entity pairs into matched and non-matched groups (Zhu et al., 2003).

3.2 Consent Notary Component

The *consent notary* component constitutes an important and structural component of the proposed architecture responsible for ensuring the validity, integrity and non-repudiation of data subjects' consents, as well as the privacy of contracting users. It assures that the generated consents (and the corresponding policies) are up-to-date and protected against malicious or unauthorized attempts to repudiate or alter them.

For the integrity and non-repudiation of consents, both state-of-the-art technologies of the Blockchain

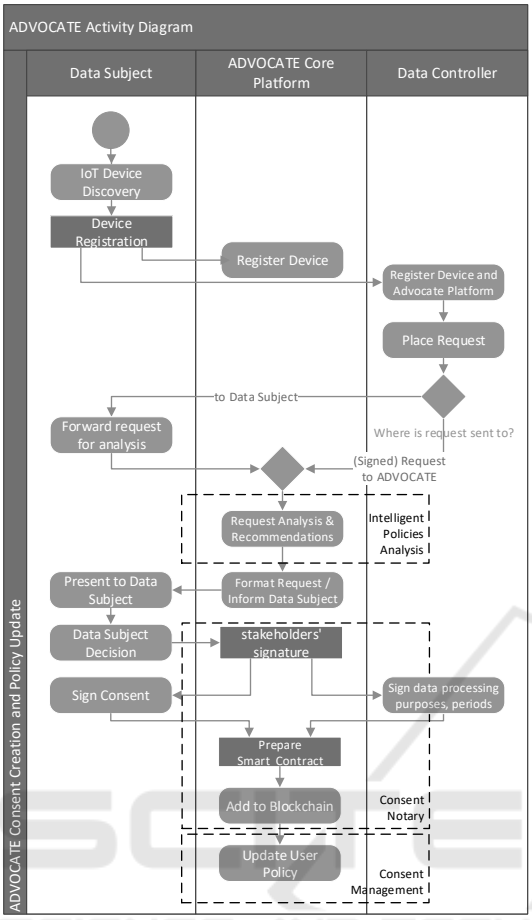

Figure 2: Interaction of ADVOCATE components during introduction of new consents.

infrastructures (Nakamoto, 2008) as well as basic techniques of the Public Key Infrastructure (PKI) (Housley et al., 1999) can be utilised. The use of a blockchain infrastructure, that was firstly introduced to secure transactions in Bitcoin cryptocurrency (Nakamoto, 2008), constitutes a cutting-edge technology for data integrity. Blockchain infrastructure is a highly innovative technology with a wide range of applications that extend from cryptocurrency to IoT, smart contracts, smart property, digital content distribution and even to health data management (Yli-Huumo et al., 2016; Conoscenti et al., 2016; Mytis-Gkometh et al., 2018).

In ADVOCATE we focus on the concept of smart contracts that have been introduced by Ethereum (Buterin, nd), which defines the rules and penalties around an agreement in a similar fashion that a conventional contract does, but also automatically enforces these obligations. In public blockchains, such as Bitcoin and Ethereum, all the transactions are public and there is no direct link to the actual user identities. However, in applications that require non-repudiation, identities should be irrevocably maintained; this can be ensured by the appropriate use of public key infrastructure solutions (Housley et al., 1999). In the proposed framework, consents would be digitally signed by contracting parties to ensure the non-repudiation and the hashed version of the consents would be submitted to a blockchain infrastructure to ensure their integrity and anonymity.

The workflow of this component is presented as follows. First, it takes as input the agreed consent from the *consent management* component. This consent might be a new one, an update following changes occurred in the corresponding policies among the parties, or a withdrawal notice. Subsequently, a request is sent to both the data controller and the data subject to sign the consent using a public key infrastructure that will later allow the verification of their identities. This signed consent is then hashed (e.g. using SHA-256, SHA-512 or even Keccak-256 hash algorithm that is used by the Ethereum (Buterin, nd)) and the hash is deployed in a *smart contract* of the blockchain infrastructure.

The smart contract between a data subject and a data controller represents a specific consent (initial, updated or withdrawal) for a specific IoT device, and is only deployed in the blockchain that bears the initial consent. Each update on the initial consent or even its final withdrawal is managed by the smart contract and each *data contract* represents a different version of the consent. In this way, it is possible via the smart contract to validate if a particular consent is the last version of it. Thus, the usage of a blockchain infrastructure, apart from the consents' integrity, ensures the versioning and the withdrawal notice of consents. At the end, the *consent notary* component returns the current version of the signed consent to the *consent management* component, accompanied by the smart contract's address on the blockchain.

At any time, the data controller and the data subject (or any other representative party) can verify the validity of the consent: (1) by validating the digital signatures on the consent, and (2) by retrieving the respective data contract (i.e., the last version) from the blockchain infrastructure via the smart contract and comparing the retrieved hash with the new hash of the claimed digitally signed consent. The use of a blockchain infrastructure is crucial for our platform to secure the provided consents in a distributed and public verifiable way without one single trusted party. The digital consents are kept in the side of the contracting parties, the blockchain infrastructure only manages the hashes of them and the smart contracts are deployed using the proposed plat-

form's private key for signing the transactions in the blockchain. This distribution of information among the parties preserves data subjects' privacy their identity is only known among the contracting parties without any leakages from the blockchain infrastructure. Overall, the proposed platform provides a transparent way of storing and processing personal data in the IoT ecosystem that ensures the privacy of users.

3.3 Intelligence Component

The *intelligence* component is responsible for gathering and analysing policy data. It is a novel and flexible hybrid machine learning system that combines an Intelligent Policies Analysis Mechanism (IPAM) to detect conflicting rules or policies related to the disposal of personal data and an Intelligent Recommendation Mechanism (IReMe) to recommend more personalized intelligent rules in real-time for the users privacy policies.

Intelligent Policies Analysis Mechanism. The Intelligent Policies Analysis Mechanism (IPAM) automates export, analysis and correlation of data subjects' policies. It utilizes intelligent technologies to detect conflicting rules or policies related to the disposal of personal data and ensure that they cannot be used for profiling and identification of data subjects.

To achieve this we use Fuzzy Cognitive Maps (FCM), an artificial intelligence technique that incorporates ideas from recurrent artificial neural networks and fuzzy logic, to create a dynamic model of a decision support system. The proposed system controls the overall user privacy by creating automated control capabilities. In order to achieve this control, it is needed to map user's consent policies and to consider whether any changes to their privacy statement affects the overall privacy.

More specifically, in the proposed FCM, the nodes are linked together by edges and each edge that connects two nodes describes the change in the activation value. The direction of the edge implies which node affects the others. The causality relationship is positive if there is a direct influence relation, negative if there is an inverse influence relation and zero if the two nodes are uncorrelated. These relationships are described by the usage of fuzzy linguistics and they are fuzzified by using membership functions that taking values in the closed interval [-1,1]. Combining the theoretical background of fuzzy logic, the FCM cover the comparison and characterization purposes of the reference sets, towards modeling and solving complex problems for which there is no structured mathematical model (Demertzis et al., 2018).

Intelligent Recommendation Mechanism. The Intelligent Recommendation Mechanism (IReMe) is a computational intelligence and machine learning mechanism that is used to recommend rules for taking decisions. It offers personalized real-time information for the users privacy policies by utilizing Cognitive Filtering (CF) (Yang et al., 2016). The CF recommends items based on a comparison between the content of the items and user's profile. The content of each item is represented as a set of descriptors or terms. The user's profile is represented with the same terms and built up by analyzing the content of items that have been already checked by the user.

In the IReMe, we use a hybrid method consisting of neighborhood-based CF and content-based filtering which is a robust model-based method that improves the quality of recommendations (Ya-Yueh Shih and Duen-Ren Liu, 2005). The aim of this approach is to achieve more personalized intelligent rules and real-time recommendations for the users privacy policies in order to avoid any data leakages. Also, this hybrid method is more versatile, in the sense that it works best when the user space is large, it is easy to implement, it scales well with no-correlated items and does not require complex tuning of properties.

4 CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a framework that addresses a major emerging need regarding users' privacy protection in the IoT. Being part of a work in progress, this framework lays the foundations for the establishment of trust relationships between data subjects and controllers towards a GDPR-compliant IoT ecosystem. The aim is to develop a user-centric solution that will allow data subjects to formulate and manage their consents policies responding to unambiguous access requests placed by data controllers. Furthermore, we utilize blockchain technology to support the integrity, the non-repudiation and the versioning of consents in a public verifiable way without any trusted party.

As future work, we intend to investigate further the issues that surround each of the components that comprise the proposed framework with an emphasis on the development of GDPR-compliant data privacy ontologies, the *consent notary* and the *intelligence component*. In addition, we aim to study the use of policies, created by our system, in policy-based access control systems, thereby developing an integrated personal data management system in the IoT.

REFERENCES

Bartolini, C., Muthuri, R., and Santos, C. (2017). Using ontologies to model data protection requirements in workflows. In *New Frontiers in Artificial Intelligence*, volume 10091, pages 233–248. Springer.

Buterin, V. (n.d.). A next-generation smart contract and decentralized application platform. Available online at: https://github.com/ethereum/wiki/wiki/White-Paper.

Cha, S. C., Chen, J. F., Su, C., and Yeh, K. H. (2018). A blockchain connected gateway for ble-based devices in the internet of things. *IEEE Access*, PP(99):1–1.

Cha, S. C., Tsai, T. Y., Peng, W. C., Huang, T. C., and Hsu, T. Y. (2017). Privacy-aware and blockchain connected gateways for users to access legacy iot devices. In *6th Global Conference on Consumer Electronics (GCCE)*, pages 1–3. IEEE.

Conoscenti, M., Vetrò, A., and Martin, J. C. D. (2016). Blockchain for the Internet of Things: A systematic literature review. In *13th International Conference of Computer Systems and Applications (AICCSA)*, pages 1–6. IEEE.

Demertzis, K., Iliadis, L. S., and Anezakis, V.-D. (2018). An innovative soft computing system for smart energy grids cybersecurity. *Advances in Building Energy Research*, 12(1):3–24.

Eckert, K., Meilicke, C., and Stuckenschmidt, H. (2009). Improving ontology matching using meta-level learning. In *The Semantic Web: Research and Applications*, volume 5554, pages 158–172. Springer.

EnCoRe Project (2010). Ensuring Consent and Revocation. www.hpl.hp.com/breweb/encoreproject/.

European Parliament and Council (2016). Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Official Journal of the European Union. http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX\%3A32016R0679.

Housley, R., Ford, W., Polk, W., and Solo, D. (1999). Internet X.509 public key infrastructure certificate and CRL profile. RFC 2459, RFC Editor.

IERC (2015). European Research Cluster on the Internet of Things, Internet of Things: IoT governance, privacy and security issues.

Lucero, S. (2016). Iot platforms: Enabling the Internet of Things. White paper, IHS Technology.

Mendez, D. M., Papapanagiotou, I., and Yang, B. (2017). Internet of Things: Survey on security and privacy. arXiv:1707.01879v2 [cs.CR].

Mozzaquatro, B. A., Jardim-Goncalves, R., and Agostinho, C. (2015). Towards a reference ontology for security in the Internet of Things. In *International Workshop on Measurements Networking (M&N)*, pages 1–6. IEEE.

Mytis-Gkometh, P., Drosatos, G., Efraimidis, P. S., and Kaldoudi, E. (2018). Notarization of knowledge retrieval from biomedical repositories using blockchain technology. In *Precision Medicine Powered by pHealth and Connected Health*, volume 66 of *IFMBE*, pages 69–73. Springer.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. https://bitcoin.org/bitcoin.pdf.

O'Connor, Y., Rowan, W., Lynch, L., and Heavin, C. (2017). Privacy by design: Informed consent and internet of things for smart health. *Procedia Computer Science*, 113:653–658.

Otero-Cerdeira, L., Rodrguez-Martnez, F. J., and Gmez-Rodrguez, A. (2015). Ontology matching: A literature review. *Expert Systems with Applications*, 42(2):949–971.

Passant, A., Laublet, P., Breslin, J. G., and Decker, S. (2009). A URI is worth a thousand tags: From tagging to linked data with MOAT. *International Journal on Semantic Web and Information Systems*, 5(3):71–94.

Pereira, T. and Santos, H. (2009). An Ontology Based Approach to Information Security. In *Metadata and Semantic Research*, volume 46, pages 183–192. Springer.

Roman, R., Najera, P., and Lopez, J. (2011). Securing the Internet of Things. *Computer*, 44(9):51–58.

Roman, R., Zhou, J., and Lopez, J. (2013). On the features and challenges of security and privacy in distributed internet of things. *Computer Networks*, 57(10):2266–2279.

Russell, B., Garlat, C., and Lingenfelter, D. (2015). Security guidance for early adopters of the Internet of Things (IoT). White paper, Cloud Security Alliance.

Sicari, S., Rizzardi, A., Grieco, L. A., and Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks*, 76:146–164.

Stankovic, J. A. (2014). Research directions for the Internet of Things. *IEEE Internet of Things Journal*, 1(1):3–9.

Xu, G., Cao, Y., Ren, Y., Li, X., and Feng, Z. (2017). Network security situation awareness based on semantic ontology and user-defined rules for Internet of Things. *IEEE Access*, 5:21046–21056.

Ya-Yueh Shih and Duen-Ren Liu (2005). Hybrid recommendation approaches: Collaborative filtering via valuable content information. pages 217b–217b. IEEE.

Yang, Z., Wu, B., Zheng, K., Wang, X., and Lei, L. (2016). A survey of collaborative filtering-based recommender systems for mobile internet applications. *IEEE Access*, 4:3273–3287.

Yli-Huumo, J., Ko, D., Choi, S., Park, S., and Smolander, K. (2016). Where is current research on blockchain technology? – A systematic review. *PLOS ONE*, 11(10):e0163477.

Zhang, Z.-K., Cho, M. C. Y., Wang, C.-W., Hsu, C.-W., Chen, C.-K., and Shieh, S. (2014). IoT security: Ongoing challenges and research opportunities. In *7th International Conference on Service-Oriented Computing and Applications*, pages 230–234. IEEE.

Zhu, X., Ghahramani, Z., and Lafferty, J. (2003). Semi-supervised learning using gaussian fields and harmonic functions. In *IN ICML*, pages 912–919.