Understanding software obfuscation and diversification as defensive measures for the cybersecurity of Internet of Things

Sampsa Rauti
University of Turku
sjprau@utu.fi

Samuli Laato
Tampere University
samuli.laato@tuni.fi

Abstract

Internet of Things (IoT) has emerged as an umbrella term to describe connecting smart everyday objects (such as washing machines, toilets and sound systems), sensors and industrial machines to the internet. While IoT devices hold potential to greatly enhance quality of life through automating and optimizing mundane tasks, there are a great deal of security and privacy challenges. For this reason, practitioners and academics have explored various ways to enhance the multi-layered security of IoT devices. One of these methods is obfuscation, which has been successfully applied to make accessing devices more difficult for adversaries. In this study, we systematically processed the literature on applying obfuscation and diversification to improve IoT cybersecurity (81 articles) and clustered this research according the obfuscation target (code, data, interface, location, traffic). We then conducted a follow-up bibliometric review of the entire research profile of IoT cybersecurity (3,682 articles) to understand how these obfuscation and diversification approaches relate to the general cybersecurity landscape and solutions of IoT. We also derive a comprehensive list of benefits and shortcomings of enhancing IoT security through diversification, and present points of departure for future research.

Keywords: Obfuscation, IoT, cybersecurity, diversification, internet of things,

1. Introduction

The world is becoming increasingly inter-connected in both the physical and digital space. One manifestation of this trend is the growing use of commercial cyber-physical systems and smart devices in households as well as industrial production. When such systems are connected to one another over a network, the resulting network is called the Internet of Things (IoT), although there is some dispute over the exact definition of the term (S. Li et al., 2015). In this study, we define IoT as a broad umbrella term that generally refers to connecting smart house appliances (e.g., lamps, vacuum cleaners or refrigerators), industrial devices, urban infrastructure (e.g., lamp posts, security cameras, sprinkler systems) and transportation (cars, airplanes, drones) to the internet. IoT devices can be remote controlled, they can automatically utilize online data to optimize their performance, and they overall hold potential to automate mundane tasks, improve energy efficiency and even improve safety and security (Kumar et al., 2019; S. Li et al., 2015; Nord et al., 2019).

1.1. Cybersecurity solutions for IoT devices

Despite these many promises, the concept of IoT has been plagued by security concerns pertaining primarily to privacy and cybersecurity (S. Li et al., 2016; Lu and Da Xu, 2018). Regarding privacy, IoT devices often accumulate sensitive sensor data from users, and if leaked, this data may be used for nefarious purposes (Weber, 2015). To counter this, edge and fog computing approaches have been proposed where the users' data never leaves their house (X. Li et al., 2018). However, privacy issues may also arise through hacked IoT devices where an adversary gains access to the devices through e.g., weak passwords or the IoT devices running outdated systems with security vulnerabilities (Kolias et al., 2017). For example, a few of the largest botnet cases reported during the past decade (Mirai, Meris) have been running on IoT-devices (e.g., cameras and internet routers) [1]. In addition to being used as part of a botnet for distributed denial-of-service (DDoS) attacks, compromised IoT devices may be used for a wide variety of nefarious or otherwise unwanted purposes ranging from spying on users to mining cryptocurrencies for the benefit of the perpetrator (Vignau et al., 2019).

In summary, there are many characteristics in IoT devices that make it critical to improve their cybersecurity. These include the potential to collect

[1] Cybersecurity journalist Brian Krebs discusses IoT botnets and why they are popular in the following post: https://krebsonsecurity.com/2021/09/krebsonsecurity-hit-by-huge-new-iot-botnet-meris/, visited on the 13th of May, 2022

URI: https://hdl.handle.net/10125/103438
978-0-9981331-6-4

Table 1. Key cybersecurity characteristics of IoT devices

Concept	Description
Rarely updated	Very few IoT devices require updating beyond cybersecurity updates. (Remesh et al., 2020)
Minimal install	IoT devices typically run on low power components and require small tailored operation systems e.g., a minimal install Linux distribution. (Hahm et al., 2015; Zikria et al., 2018)
Focus on data	A lot of IoT devices have sensors that collect data – this allows them to optimize energy consumption etc. However, this data can also be highly sensitive. (Xu et al., 2019)
Communication with other devices	Extra care is needed to ensure that only trusted and desired parties are able to communicate with the IoT device. (Kolias et al., 2017)

private or sensitive information, the IoT devices being critical cyber-physical system that may cause harm in the real world through remote controlling (e.g. Rauti et al., 2020) and the increased power consumption, them being used in botnets and other negative aspects following from compromised IoT devices (Vignau et al., 2019). In order to improve the cybersecurity of IoT devices, it is paramount to understand their key characteristics and how they differ from other software systems and technologies. We list the key cybersecurity characteristics of IoT devices in Table 1.

1.2. Obfuscation and diversification to enhance the multi-layered security of software systems

One of the important principles when developing software for IoT devices is to keep memory usage and computational requirements low so that IoT devices, with their limited memory and computation power, are able to operate smoothly (Hahm et al., 2015). This principle also holds for security solutions on IoT devices – performance, effectiveness and power consumption should not be sacrificed for security. This means many traditional solutions such as anti-virus programs are not a reasonable security solution for most IoT devices. Instead, computationally inexpensive and memory-efficient solutions are needed.

One such solution is interface diversification, which is an approach based on creating unique instances of software interfaces (Rauti et al., 2021). The program code is diversified so that different instances are syntactically different but functionality is not affected. Interface diversification can be achieved by employing various different source code obfuscation techniques (Collberg et al., 1997). Cohen presented one obfuscation approach in 1993 and proposed creating diversified versions of operating systems (Cohen, 1993). After this, there has been a large body of research concerning interface diversification (Hosseinzadeh et al., 2018), and in recent years, the idea has also been increasingly been applied to software running on IoT devices (Hosseinzadeh et al., 2016; Koivunen et al., 2016; Mäki et al., 2016).

Although there are billions of IoT devices connected to the internet, only relatively small set of different operating systems and programs are being used on these devices. This monoculture is not unique only to IoT devices, but is a key reason why obfuscation and diversification approaches hold so much potential in improving system security (Collberg, 2018). In other words, due to the identical design and well-known interfaces, large groups of IoT devices are susceptible to the same vulnerabilities and security attacks. Therefore, a malicious adversary can compromise a huge number of systems with a single attack, as evidenced by e.g., the Mirai botnet attacks (Kolias et al., 2017). Interface diversification is a way to add multiculturalism to the software design, which mitigates opportunities for non-targeted large-scale attacks (Rauti et al., 2021). Assuming a malicious attacker discovers how one unique IoT device is diversified, the other devices are still safe due to their unique and secret diversification. It would take more time and resources for the attacker to reverse engineer the diversification procedure, significantly slowing down the attacker. In the best scenario, the attacker is forced to build system-specific attack models, which renders various currently existing botnet approaches obsolete.

One of the main advantages of diversification is that the technique can improve system security without a significant increase in resource consumption (Rauti et al., 2021). For instance, using simple obfuscation techniques such as changing the names and parameter order of functions does not lead to increased computational power or memory usage. This makes the techniques particularly suitable for IoT devices that run preferably on low power and computational resources (Hosseinzadeh et al., 2016; Koivunen et al., 2016; Mäki et al., 2016). With the continuously increasing number of IoT devices, the incentives to attack the devices with bulk attacks also increase. For this reason, proactive protection techniques, in particularly those addressing the monoculture issue

of IoT (Collberg, 2018), should be given careful consideration. Here obfuscation and diversification appear as one of the most promising solutions.

1.3. Benefits and shortcomings of enhancing IoT security through diversification

Recent studies have emphasized that since IoT devices are relatively seldom updated and run a very limited and rather static set of software, internal interface diversification solutions may be particularly relevant and effective for boosting the security of these devices (Rauti et al., 2021). Interface diversification has the following favorable properties:

Proactiveness. Interface diversification can be considered a proactive security measure: unlike many traditional security solutions, diversification does not assume that the exploit works in a certain way or that the malicious binary follows a specific pattern. Previously unknown zero-day exploit will be rendered useless if they try to use the well-known interfaces (Cohen, 1993; Koivunen et al., 2016).

Passiveness. Interface diversification passively waits the malware to make its move. The solution does not waste resources in trying to prevent malware from infiltrating the system or executing. However, the harmful software is prevented from working in an intended manner.

Low performance requirements. When the diversification solution is kept relatively simple, for example by only diversifying system call numbers or names of library functions, the effects on the system performance are negligible or modest (Collberg et al., 1997). Obviously, this property is especially important in low-resource IoT devices.

Orthogonality. Interface diversification can be seen as a part of multi-layered security scheme. Diversification is orthogonal: it can be used together with many other security approaches. Traditional solutions such as intrusion detection systems and cryptography can be combined with interface diversification to enhance overall security (de Haro-Olmo et al., 2020). This is an important property, because interface diversification is not a silver bullet that works against all attack scenarios.

Counterbalancing poor security. Interface diversification counterbalances the poor security of IoT devices by providing an additional layer of security. Even if a malicious program finds a vulnerability invades the device, it cannot use essential interfaces of the target system. This is especially important because software on IoT devices is often not updated regularly.

Invisibility. When diversification is applied to internal interfaces of the system, a normal end user does not notice anything out of ordinary (Collberg et al., 1997). External interfaces that the user directly interacts (such as graphical user interfaces) with are left intact and not affected by diversification. Diversification also does not affect the software development process and programmers' work, because it can be applied automatically after the source code has been compiled.

The list of shortcomings of the approach is shorter, with perhaps the most important one being the monetary costs of implementing such solutions and challenges in deploying updates to the obfuscated devices (Koivunen et al., 2016). Even in cybersecurity some cost/gain balancing needs to be done, and some obfuscation approaches may be needlessly costly whilst offering security that could also be achieved through other means. Another shortcoming may be on usability. It is not always entirely clear who would be in charge of obfuscating the system and deploying the solution. There is also the additional work of ensuring that the system would operate as intended for the user even after such measures have been put in place.

1.4. Research question

In order to understand obfuscation and diversification as cybersecurity approaches for IoT devices, we wanted to observe the overall IoT cybersecurity landscape and locate how obfuscation and diversification fit in. Accordingly, we formulate the following research question (RQ):

RQ: How do obfuscation and diversification techniques compare and relate to the overall cybersecurity landscape of IoT devices?

In order to answer this question, we first systematically reviewed the academic literature on diversification and obfuscation techniques for IoT security (n=81), and extracted the approaches for enhancing the multi-layered security of IoT systems. In order to then understand these solutions as part of the overall IoT cybersecurity solutions landscape, we performed a bibliometric co-word analysis of the overall IoT cybersecurity research field (n=3,682) and evaluated obfuscation and diversification techniques in relation to this research profile. With this approach we contribute to the research field of IoT security by synthesising the academic knowledge on obfuscation and diversification techniques for improving the multi-layered security of IoT devices.

The rest of this study is structured as follows. First, we describe our methods for the two literature search processes and subsequent data analyses. Second,

we present our findings followed by discussion. We conclude the work by summarizing our key findings and by presenting an agenda for future research within this field.

2. Materials and methods

The research process in this study is depicted in Fig 1. We conducted two literature searches, one for obfuscation and diversification for IoT cybersecurity, and another to understand the overall IoT cybersecurity research landscape. We then combined our findings from these approaches to understand obfuscation and diversification as part of thee multi-layered security solutions for IoT devices. With this approach, we are able to conceptually root the research on obfuscation and diversification firmly within the broader IoT cybersecurity field. Furthermore, as we observe obfuscation and diversification from this broader vantage point, we are able to derive points of departure for future work related to aspects such as the practicality, applicability and feasibility of obfuscation and diversification for improving the multi-layered cybersecurity solutions of IoT devices.

2.1. Literature searches

2.1.1. The first search: obfuscation and diversification for improving IoT cybersecurity In March-April 2022, we gathered keywords related to IoT, cybersecurity and diversification. These keywords were gathered from reading existing literature reviews on IoT cybersecurity (Corallo et al., 2022; Kuzlu et al., 2021; Lee, 2020; Lu and Da Xu, 2018) and obfuscation and diversification (Hosseinzadeh et al., 2018). In addition, we read white papers and selected practitioner blog posts (e.g. [2]) on obfuscation and diversification for IoT. The final set of search terms resulting from this preliminary scoping are displayed in Table 2.

We chose to search for studies from the Elsevier Scopus research database due to its coverage of relevant research, and its high standards in indexing studies. Scopus contains research from several relevant information systems and computer science research databases such as IEEE Xplore, DBLP Computer Science Bibliography and ACM Digital Library. Furthermore, Scopus offers researchers a high level of control over the search terms and results curation was well as easy-to-use export tools. For these reasons, Scopus was estimated to be a good fit for this research.

[2]https://encyclopedia.thefreedictionary.com/internet+of+things, accessed April 5, 2022

Table 2. The search terms

Main word	Synonyms or closely related terms
IoT	internet of things, internet-of-things, smart home, smart devices, home automation
Cybersecurity	data security, cyber-security, information security, software security, security, privacy, trust
Diversification	diversification, obfuscation, randomization, randomisation

Using the keywords specified in Table 2, we conducted a search on Scopus in April, 2022. We limited the search to peer-reviewed studies only, which left us with 313 articles. We then proceeded to read the abstracts of the studies, excluding (1) articles which were not in English; (2) articles which were not peer-reviewed; and (3) articles which were not related to obfuscation and diversification techniques for IoT cybersecurity. During this process, we noticed that in particular the search term *diversification* was used to refer to various things other than the software/network cybersecurity techniques. Examples included (1) mentions where diversification was used to describe the proliferation, distribution or adoption of IoT devices in real world context; and (2) studies where ”diversification” was used to describe the growing variance in the types of available IoT devices. We followed the abstract screening with a full-text assessment of the remaining studies, and used the same criteria as in the previous step. These processes were carried out by the first author, and resulted in the final number of 81 articles to be included in the final synthesis.

2.1.2. The second search: the overall literature on IoT cybersecurity A preliminary search showed us that the amount of literature on IoT cybersecurity is enormous. For this reason, we chose the bibliometric co-word approach for understanding this research field, which is a particularly suitable method for bringing clarity to complex and large research fields (Laato et al., 2022; Malanski et al., 2021; Van Eck and Waltman, 2010). Similarly to the previous step we used Scopus. Since false positives are a critical concern in bibliometric reviews, we paid extra care in selecting the keywords. For example, we omitted general one word keywords such as *security*, *privacy* and *trust* which were part of the search string in the first search. Instead, we chose more descriptive terms such as *software security* and *information security*. Based on these keywords we formulated a search string combining the IoT and

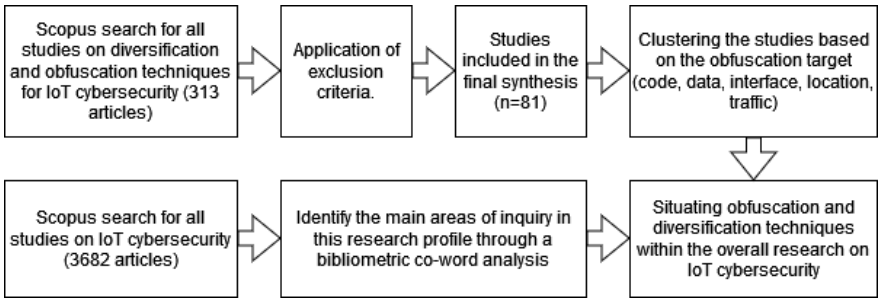

Figure 1. An overview of the research process in this study.

cybersecurity keywords. The final search was performed on the 11th of June, 2022. This search resulted into 4218 articles. The articles were limited to peer-reviewed studies only (journal articles, conference proceedings and book chapters), which resulted in the final number of 3,682 articles to be included in the bibliometric review.

2.2. Data analysis

We began our analysis by extracted from the initial set of papers (n=81) the target of the technical obfuscation. In all the papers of the final sample this was specified either explicitly or implicitly. We also extracted the publication years from the studies to see if obfuscation techniques were a growing, diminishing or stable trend within the broader IoT cybersecurity literature. We then moved to the larger sample of studies (n=3,682) and extracted basic information from the studies including (1) publication year; (2) document type; (3) subject area; (4) publication venue; (5) most popular keywords; and (6) country of the first author. From this information we are able to obtain an understanding of where the research has been conducted and published and when. This data could also reveal biases in the research field and offer opportunities for future research. Next, we conducted a co-word analysis to understand the research profile in more detail. Co-word analysis is a data mining technique that connects keywords that appear in the same paper together, forming a network of concepts that highlights their relationships (Van Eck and Waltman, 2010). In this study, we specified that only keywords that appeared in four or more studies are included in the final concept network. By setting this limit, the analysis result excludes weak relationships (that may be accidental) and thus increases the reliability of the result. We performed the analysis using the VOSviewer tool (Van Eck and Waltman, 2010), and looked at the author-given keywords. We iterated the analysis a couple of times and combined similar keywords together, and testing the outcome by tweaking how many times the keywords had to appear together in the sample of studies to be included in the Figure. The iterations and decisions in this process were influenced by the authors' evolving understanding of the research profile as they got more acquainted with the studies in the final sample.

Finally, we compared the target areas of obfuscation from the initial set of papers (n=81) as well as the publication years of these studies to those within the broad IoT cybersecurity research (n=3,682). This approach allowed us to obtain an understanding of the trends and trajectories of obfuscation and diversification research within the overall IoT cybersecurity research domain.

3. Findings

Here we present our findings by first focusing on the overall bibliometric profile of the IoT cybersecurity field, and then connecting the findings from the obfuscation and diversification literature to this research field.

3.1. The bibliometric profile of the research field of IoT cybersecurity

As can be seen from Figure 2, the research field of IoT cybersecurity is growing strongly with more publications out each year than the year before. The field began growing rapidly in 2013, and the number of publications more than doubled each year until 2017. Afterwards the growth of the research field has continued steadily, but as of 2019 has also shown some signs of stabilizing. Perhaps unsurprisingly, the number of IoT cybersecurity publications is positively correlated with the number of publications within the entire field of IoT. According to Scopus, the number studies mentioning IoT in the title, abstract or keywords has been in the tens of thousands each year after the year 2017, peaking at 62,549 studies published in 2021.

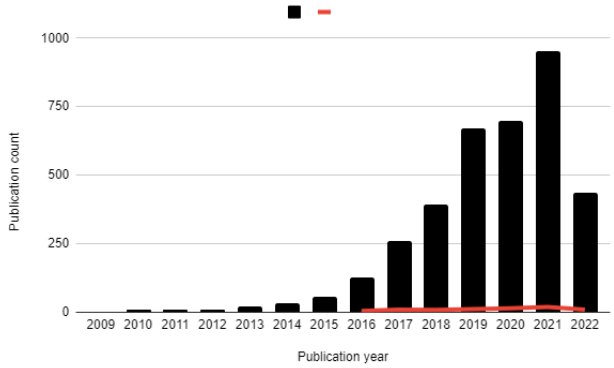

Figure 2. The publication years of the studies on IoT cybersecurity (black columns) compared to the publication years of diversification and obfuscation for IoT cybersecurity (red line). Both trends are similar in trajectory with no notable observable differences.

From here we can make the crude estimation that IoT cybersecurity research is roughly one seventh of the total number of IoT publications. Looking at obfuscation and diversification within the IoT cybersecurity field, we see that it is represents roughly 1/45 of the overall IoT cybersecurity research. As a trend, obfuscation for IoT research has been growing roughly at the same rate as the overall IoT cybersecurity research.

The majority of the studies within the field of IoT cybersecurity are published in conference proceedings (n=2,037) followed by journals (n=1,439). The remainder (n=206) are book chapters and other peer-reviewed publications. According to Scopus, these studies are overwhelmingly carried out in the field of computer science (n=2,972) or engineering (n=1,928), but significant number of studies are also conducted within the field of mathematics (n=575) and decision sciences (=564). There is also overlap in the field classifications, meaning some studies are interdisciplinary, and related to both mathematics and computer science. The majority of the research is produced by scholars from the USA (n=641) followed up by China (n=567), India (n=452), United Kingdom (n=303) and Australia (n=164). Altogether, the research has been carried out in 101 different countries. While there certainly is emphasis on USA, China and India, these numbers roughly correlate to the overall research output of these countries. Hence, we estimate that no significant country-related publication bias exists in this domain.

The results of the co-word analysis are displayed in Figure 3. The concept map in Figure 3 illustrates that while academics have studied many security technologies closely related to obfuscation and diversification, these techniques, and proactive cybersecurity measures in general, seem to be missing from the big picture. Next, we discuss these two in further detail with references to the studies.

Table 3. The obfuscation targets

Obfuscation target	Number of publications
Data obfuscation	22
Code obfuscation by malware	17
Location obfuscation	11
Code obfuscation	10
Traffic obfuscation	8
Route obfuscation	7
Interface obfuscation	4
IP address obfuscation	2

3.2. Categories of diversification and obfuscation approaches within the landscape of IoT cybersecurity research

Table 3 shows different categories of obfuscation related to IoT cybersecurity. Most of the obfuscation schemes introduced in analyzed papers concentrated on obfuscating data (n=22). The data processed by IoT devices can also be obfuscated to protect users' privacy or intellectual property. While encryption is usually the primary method for protecting data from adversaries, using obfuscating techniques instead of encryption is often necessary when it comes to IoT devices with limited resources and low computational power (Khan et al., 2017). A special category of data obfuscation in mobile IoT devices is location obfuscation (n=11) which aims to preserve the user's location privacy while preserving service utility (Butun et al., 2019).

Many obfuscation approaches concentrate on traditional code obfuscation (n=10), in other words, obfuscating the internal structure of programs in order to make it more difficult for the adversaries to understand reverse engineer, and modify programs. For example, Nausheen and Begum propose protecting mobile eHealth applications using code obfuscation techniques (Nausheen and Begum, 2018), while Pastrana et al. present an obfuscation mechanism against code reuse attacks for embedded devices (Pastrana et al., 2016).

Several internal interfaces of IoT devices can also be obfuscated (n=4) to prevent malware authors from abusing the device's resources. Koivunen et al., for example, propose obfuscating several internal interfaces of IoT devices, such as system call interfaces and operating system libraries (Koivunen et al., 2016). Interface diversification is a lightweight protection

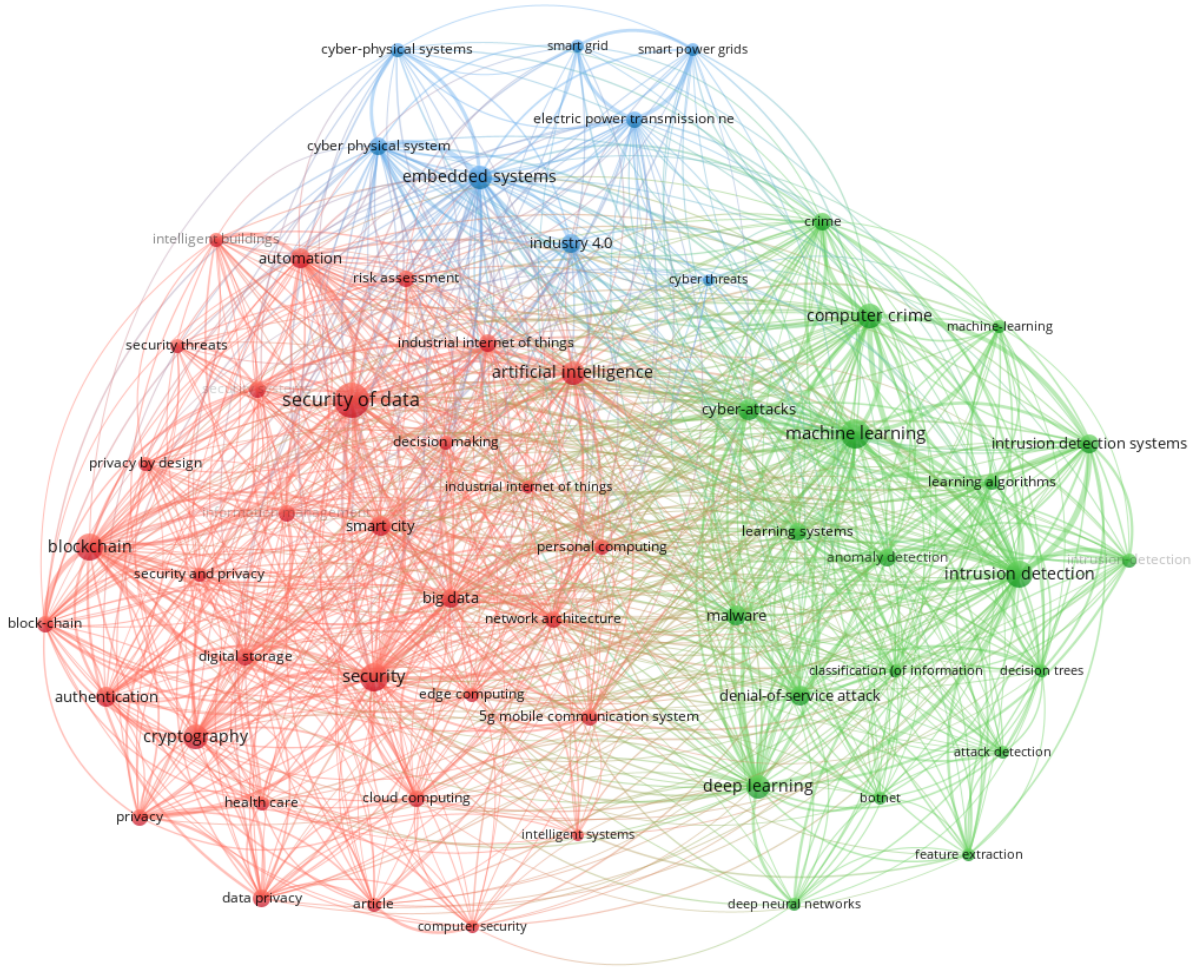

Figure 3. Visualizing the bibliometric co-word analysis.

mechanism that does not require lots of computational resources unlike many traditional software security mechanisms.

In the studied papers, software obfuscation is also regularly used by malware authors to hide the malicious nature of their code and executables (n=17). As malware authors produce several diversified functionally equivalent versions of their harmful programs, approaches for measuring the similarity of these diversified pieces of malware have to be developed (e.g., Venkatraman and Alazab, 2017).

When it comes to obfuscation on network level, obfuscating the contents and patterns of network packets is a popular approach (n=8). Datta et al. introduce a library that replaces standard networking functions and obfuscates traffic patterns of an IoT device by using payload padding, fragmentation mechanisms, and randomly generated fake traffic (Datta et al., 2018). The way packets are routed in a network (n=7) can also be obfuscated (Bin-Yahya and Shen, 2022). For example, Bin–Yahya and Shen (Bin–Yahya and Shen, 2021) present a proactive route mutation scheme that alters the routes in wireless sensor networks to prevent reconnaissance and sniffer attacks. Research has also looked at IP address obfuscation (n=2) and reassigning IP addressess as a moving target defence approach in order to prevent attackers targeting IoT devices (He et al., 2021).

Turning to the larger IoT security picture of Fig 3, we can see that obfuscation and proactive security methods in general are absent in the picture. However, software and network level obfuscation contributes to many general principles in the red area, such as security, data security and cryptography. Software security is enhanced through or protecting internal structure of programs and diversifying interfaces, making it difficult for malware to attach itself to programs or interfaces and abuse them to achieve

its goals (Cohen, 1993). On the other hand, data privacy in IoT systems can also be protected by using obfuscation like lightweight encryption when computation-intensive encryption methods cannot be used to ensure confidentiality of data (Yavari et al., 2017).

In the green area, central themes are malware, machine learning, intrusion detection. Obfuscation is connected to machine learning and artificial intelligence mainly through efforts by researchers to use these approaches to classify and understand obfuscated malicious code (Dib et al., 2021). Obfuscated malicious programs and network traffic can also be detected by intrusion and anomaly detection tools.

Finally, the blue area highlights industrial internet, and applications of IoT such as smart power grids and power transmission networks. Such parts of critical infrastructure that may never receive security updates can greatly benefit from supplementary security measures such as diversification (Koivunen et al., 2016).

4. Discussion

Our findings contribute to the research on IoT cybersecurity (Kuzlu et al., 2021; Lee, 2020; S. Li et al., 2015) as well as obfuscation and diversification (Hosseinzadeh et al., 2018; Rauti et al., 2021). We reiterate the claims from previous studies that obfuscation and diversification are particularly suitable approaches for improving the multi-layered security of IoT devices (Hosseinzadeh et al., 2016). Accordingly, this study raises questions as to why these techniques have seen so little attention in the academic research despite the excellent alignment to the context of IoT cybersecurity. To investigate this issue further, we propose three key avenues for future research. First, research is needed with commercial obfuscation and diversification tools and products. This would provide insights into the size of this industry, who are the main customers and what kind of systems are worth protecting. This leads us to our second point of departure for future research, which is to conduct experiments with diversifying various interfaces across various devices, and to measure the effectiveness of these techniques against cyber attacks. This research would allow us to gain a better understanding of what types of systems, as well as components of systems, are worth protecting via diversification. Third and finally, we propose that in the future researchers would carry out comparison analyses between obfuscation and other cybersecurity measures. Such an approach would allow the academic community as well as practitioners to understand how feasible and effective these solutions are in real world environments. These future research directions along with descriptions of them are displayed in Table 4.

Table 4. Future research agenda on obfuscation and diversification techniques for IoT devices.

Focus area	Description of the future research topic
Case studies with commercial products	Studies on applying obfuscation/diversification in commercial products are largely missing. There is little academic knowledge on the applications of these approaches in commercial IoT products.
System comparison studies	There are various IoT devices, some with more processing power than others. Feasibility analyses are needed on what kinds of systems, and what parts of those systems, are worth protecting with diversification and obfuscation.
Approach comparison studies	It remains unclear why obfuscation and diversification have seen relatively little attention in academic IoT cybersecurity research. An important avenue for assessing the feasibility of this approach is to examine implementation costs, costs on usability, expected value and security enhancement against various attacks.

As all studies our research has limitations, two of which in particular require further elaboration. First, we only focused on the scientific literature, but previous work have advocated for the importance of also including grey literature (Mahood et al., 2014). This limitation can be seen as intrinsic to bibliometric studies and those applying co-word analysis, since grey literature sources often lack keywords and other bibliometric information that would be required to objectively compare the grey literature sources to academic studies. Second, the broader bibliometric search may have contained some false positives, which we sought to mitigate by only including domain-specific and precise keywords. Overall we estimate that the literature reviews in this study provides valuable insights on the literature profile despite these minor limitations. However, we encourage future research to look at the grey literature on the topic to compare and contrast our observations.

5. Conclusion

The trend of connecting various devices and sensors to the internet continues to this day, and as a consequence, we are seeing more and more cyber-physical systems and IoT devices in our daily lives. While these developments offer enormous benefits with regards to automation and optimization, there are cybersecurity concerns. The constantly shifting and changing nature of the cybercrime landscape requires multi-layered proactive measures. In this study, we reviewed the literature on obfuscation and diversification techniques for IoT security. We extracted the various targets of obfuscation within the research field of IoT cybersecurity and examined how obfuscation and diversification relate to the entire multi-layered cybersecurity environment of IoT devices. In summary, by building multi-layered defence mechanisms to cyber-physical systems, we ensure that even if some defences fall, the entire system is not compromised. As a proactive invisible solution that consumes no to little energy, we encourage practitioners to look further into obfuscation and diversification approaches for improving IoT cybersecurity.

References

Bin-Yahya, M., & Shen, X. (2022). Secure and energy-efficient network topology obfuscation for software-defined wsns. *IEEE Internet of Things Journal*.

Bin–Yahya, M., & Shen, X. S. (2021). Srrm: Ranking-based route mutation scheme for software-defined wsns. *2021 IEEE Global Communications Conference (GLOBECOM)*, 01–06.

Butun, I., Österberg, P., & Gidlund, M. (2019). Preserving location privacy in cyber-physical systems. *2019 IEEE Conference on Communications and Network Security (CNS)*, 1–6.

Cohen, F. B. (1993). Operating system protection through program evolution. *Comput. Secur.*, *12*(6), 565–584.

Collberg, C. (2018). Code obfuscation: Why is this still a thing? *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy*, 173–174.

Collberg, C., Thomborson, C., & Low, D. (1997). A taxonomy of obfuscating transformations.

Corallo, A., Lazoi, M., Lezzi, M., & Luperto, A. (2022). Cybersecurity awareness in the context of the industrial internet of things: A systematic literature review. *Computers in Industry*, *137*, 103614.

Datta, T., Apthorpe, N., & Feamster, N. (2018). A developer-friendly library for smart home iot privacy-preserving traffic obfuscation. *Proceedings of the 2018 Workshop on IoT Security and Privacy*, 43–48.

de Haro-Olmo, F. J., Varela-Vaca, Á. J., & Álvarez-Bermejo, J. A. (2020). Blockchain from the perspective of privacy and anonymisation: A systematic literature review. *Sensors*, *20*(24), 7171.

Dib, M., Torabi, S., Bou-Harb, E., & Assi, C. (2021). A multi-dimensional deep learning framework for iot malware classification and family attribution. *IEEE Transactions on Network and Service Management*, *18*(2), 1165–1177.

Hahm, O., Baccelli, E., Petersen, H., & Tsiftes, N. (2015). Operating systems for low-end devices in the internet of things: A survey. *IEEE Internet of Things Journal*, *3*(5), 720–734.

He, G., Si, Y., Xiao, X., Wei, Q., Zhu, H., & Xu, B. (2021). Preventing iot ddos attacks using blockchain and ip address obfuscation. *2021 13th International Conference on Wireless Communications and Signal Processing (WCSP)*, 1–5.

Hosseinzadeh, S., Hyrynsalmi, S., & Leppänen, V. (2016). Obfuscation and diversification for securing the internet of things (iot). In *Internet of things* (pp. 259–274). Elsevier.

Hosseinzadeh, S., Rauti, S., Laurén, S., Mäkelä, J.-M., Holvitie, J., Hyrynsalmi, S., & Leppänen, V. (2018). Diversification and obfuscation techniques for software security: A systematic literature review. *Information and Software Technology*, *104*, 72–93.

Khan, K. M., Shaheen, M., & Wang, Y. (2017). Data confidentiality in cloud-based pervasive system. *Proceedings of the Second International Conference on Internet of things, Data and Cloud Computing*, 1–6.

Koivunen, L., Rauti, S., & Leppänen, V. (2016). Applying internal interface diversification to iot operating systems. *2016 International Conference on Software Security and Assurance (ICSSA)*, 1–5.

Kolias, C., Kambourakis, G., Stavrou, A., & Voas, J. (2017). Ddos in the iot: Mirai and other botnets. *Computer*, *50*(7), 80–84.

Kumar, S., Tiwari, P., & Zymbler, M. (2019). Internet of things is a revolutionary approach for future

technology enhancement: A review. *Journal of Big data*, *6*(1), 1–21.

Kuzlu, M., Fair, C., & Guler, O. (2021). Role of artificial intelligence in the internet of things (iot) cybersecurity. *Discover Internet of things*, *1*(1), 1–14.

Laato, S., Farooq, A., Vilppu, H., Airola, A., & Murtonen, M. (2022). Higher education during lockdown: Literature review and implications on technology design. *Education Research International*, *2022*.

Lee, I. (2020). Internet of things (iot) cybersecurity: Literature review and iot cyber risk management. *Future Internet*, *12*(9), 157.

Li, S., Tryfonas, T., & Li, H. (2016). The internet of things: A security point of view. *Internet Research*.

Li, S., Xu, L. D., & Zhao, S. (2015). The internet of things: A survey. *Information systems frontiers*, *17*(2), 243–259.

Li, X., Liu, S., Wu, F., Kumari, S., & Rodrigues, J. J. (2018). Privacy preserving data aggregation scheme for mobile edge computing assisted iot applications. *IEEE Internet of Things Journal*, *6*(3), 4755–4763.

Lu, Y., & Da Xu, L. (2018). Internet of things (iot) cybersecurity research: A review of current research topics. *IEEE Internet of Things Journal*, *6*(2), 2103–2115.

Mahood, Q., Van Eerd, D., & Irvin, E. (2014). Searching for grey literature for systematic reviews: Challenges and benefits. *Research synthesis methods*, *5*(3), 221–234.

Mäki, P., Rauti, S., Hosseinzadeh, S., Koivunen, L., & Leppänen, V. (2016). Interface diversification in iot operating systems. *Proceedings of the 9th International Conference on Utility and Cloud Computing*, 304–309.

Malanski, P. D., Dedieu, B., & Schiavi, S. (2021). Mapping the research domains on work in agriculture. a bibliometric review from scopus database. *Journal of Rural Studies*, *81*, 305–314.

Nausheen, F., & Begum, S. H. (2018). Healthcare iot: Benefits, vulnerabilities and solutions. *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, 517–522.

Nord, J. H., Koohang, A., & Paliszkiewicz, J. (2019). The internet of things: Review and theoretical framework. *Expert Systems with Applications*, *133*, 97–108.

Pastrana, S., Tapiador, J., Suarez-Tangil, G., & Peris-López, P. (2016). Avrand: A software-based defense against code reuse attacks for avr embedded devices. *International conference on detection of intrusions and malware, and vulnerability assessment*, 58–77.

Rauti, S., Laato, S., & Pitkämäki, T. (2020). Man-in-the-browser attacks against iot devices: A study of smart homes. *International Conference on Soft Computing and Pattern Recognition*, 727–737.

Rauti, S., Laurén, S., Mäki, P., Uitto, J., Laato, S., & Leppänen, V. (2021). Internal interface diversification as a method against malware. *Journal of Cyber Security Technology*, *5*(1), 15–40.

Remesh, A., Muralidharan, D., Raj, N., Gopika, J., & Binu, P. (2020). Intrusion detection system for iot devices. *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 826–830.

Van Eck, N., & Waltman, L. (2010). Software survey: Vosviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538.

Venkatraman, S., & Alazab, M. (2017). Classification of malware using visualisation of similarity matrices. *2017 Cybersecurity and Cyberforensics Conference (CCC)*, 3–8.

Vignau, B., Khoury, R., & Hallé, S. (2019). 10 years of iot malware: A feature-based taxonomy. *2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, 458–465.

Weber, R. H. (2015). Internet of things: Privacy issues revisited. *Computer Law & Security Review*, *31*(5), 618–627.

Xu, X., He, C., Xu, Z., Qi, L., Wan, S., & Bhuiyan, M. Z. A. (2019). Joint optimization of offloading utility and privacy for edge computing enabled iot. *IEEE Internet of Things Journal*, *7*(4), 2622–2629.

Yavari, A., Panah, A. S., Georgakopoulos, D., Jayaraman, P. P., & van Schyndel, R. (2017). Scalable role-based data disclosure control for the internet of things. *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, 2226–2233.

Zikria, Y. B., Yu, H., Afzal, M. K., Rehmani, M. H., & Hahm, O. (2018). Internet of things (iot): Operating system, applications and protocols design, and validation techniques.