A Systematic Survey of Industrial Internet of Things Security: Requirements and Fog Computing Opportunities

Tange, Koen Pieter; De Donno, Michele; Fafoutis, Xenofon; Dragoni, Nicola

Published in:
IEEE Communications Surveys and Tutorials

Link to article, DOI:
10.1109/COMST.2020.3011208

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
A Systematic Survey of Industrial Internet of Things Security: Requirements and Fog Computing Opportunities

Koen Tange, Graduate Student Member, IEEE, Michele De Donno, Student Member, IEEE, Xenofon Fafoutis, Senior Member, IEEE, and Nicola Dragoni

Abstract—A key application of the Internet of Things (IoT) paradigm lies within industrial contexts. Indeed, the emerging Industrial Internet of Things (IIoT), commonly referred to as Industry 4.0, promises to revolutionize production and manufacturing through the use of large numbers of networked embedded sensing devices, and the combination of emerging computing technologies, such as Fog/Cloud Computing and Artificial Intelligence. The IIoT is characterized by an increased degree of inter-connectivity, which not only creates opportunities for the industries that adopt it, but also for cyber-criminals. Indeed, IoT security currently represents one of the major obstacles that prevent the widespread adoption of IIoT technology. Unsurprisingly, such concerns led to an exponential growth of published research over the last few years. To get an overview of the field, we deem it important to systematically survey the academic literature so far, and distill from it various security requirements as well as their popularity. This paper consists of two contributions: our primary contribution is a systematic review of the literature over the period 2011-2019 on IIoT Security, focusing in particular on the security requirements of the IIoT. Our secondary contribution is a reflection on how the relatively new paradigm of Fog computing can be leveraged to address these requirements, and thus improve the security of the IIoT.

Index Terms—Industrial Internet of Things, cyber-security, security requirements, fog computing.

ACRONYMS

3GPP The 3rd Generation Partnership Project
5G Fifth generation cellular network technology
ABS Attribute Based Signatures
AC Access Control
ACL Access Control List

BYOK Bring Your Own Key
CIA Confidentiality, Integrity, Availability
CI Critical Infrastructure
DDoS Distributed Denial of Service
DHT Distributed Hash Table
DID Decentralized Identifier
DoS Denial of Service
DTLS Datagram Transport Layer Security
ENISA European Union Agency for Network and Information Security
GDPR General Data Protection Regulation
ICS Industrial Internet Consortium
IDS Intrusion Detection System
IETF Internet Engineering Task Force
IIoT Industrial Internet of Things
IoT Internet of Things
LLN Low-power Lossy network
LPWAN Low Power Wide Area Network
M2M Machine-to-Machine
MQTT Message Queue Telemetry Transport
NB-IoT Narrow Band IoT
NFC Near Field Communication
NFV Network Function Virtualization
OPC UA The OPC Unified Architecture
OT Operational Technology
OWASP Open Web Application Security Project
PKI Public Key Infrastructure
PLC Programmable Logic Controller
PUF Physically Uncloneable Function
SCADA Supervisory Control And Data Acquisition
SDN Software Defined Networking
SSI Self Sovereign Identity
TEE Trusted Execution Environment
TLS Transport Layer Security
TPM Trusted Platform Module
TSN Time Sensitive Networking
WPAN Wireless Personal Area Network
WSN Wireless Sensor Networks
ZTN Zero Trust Networking.

I. INTRODUCTION

Industry 4.0, also referred to as 4th industrial revolution, represents a new industrial era, whereby due to the
increasing availability, affordability, and capability of sensors, processors, and communication technologies, the number of embedded devices used in industrial applications is rapidly increasing. This leads to a growth in the interest for the IIoT: a large network of devices, systems, and applications communicating and sharing intelligence with each other, the external environment, and with humans [1]. According to Accenture [1], the IIoT could be worth 7.1 trillion U.S. dollars to the United States and more than 1.2 trillion to Europe by 2030.

In this wave of excitement, IoT security represents one of the biggest weak points holding back the adoption of the IIoT. As a matter of fact, IoT devices are often poorly secured [2] and thus easy targets for malware taking advantage of them to run devastating cyber-attacks, such as Distributed Denial of Service (DDoS) [3] (e.g., Mirai [4] affected consumer IoT) or sabotage attacks. Threats are not limited to the consumer IoT. In fact, traditional industrial environments have been subject to attacks in the past, sometimes with devastating results (e.g., StuxNet [5] or CrashOverride/Industroyer [6]). It is thus apparent that without security, IIoT will never be able to deliver its full potential. As a result, recent years have seen an unprecedented growth of research in IIoT security.

In this landscape, a relatively new computing paradigm has attracted attention: Fog computing [7]. Fog computing is a system-level architecture born from the necessity of bridging the gap between IoT and Cloud computing, by distributing resources and services along the continuum from Cloud to IoT [8]. Among others, one of the promises of Fog computing is to present a possible solution to the (I)IoT security problem.

A. Contribution

In this article, we present a systematic survey on the security requirements of the IIoT. As we quantitatively demonstrate in Section VI, the field of IIoT security has grown rapidly over the last few years, and this momentum motivates this article and the need for an up-to-date systematic survey.

In particular, as our primary contribution, we survey the literature on IIoT security over the period 2011-2019, which corresponds to more than 200 papers. In turn, we identify, categorize, and discuss the IIoT security requirements that have been identified by the research community, highlighting the research interest attracted by each of them over the target period. In addition, we provide statistics with regard to the geographical distribution and the publication venue of the surveyed papers.

As a secondary contribution, in the final part of the article, we discuss how the Fog computing paradigm can be used to address these requirements. Our reflection identifies numerous research opportunities at the intersection of Fog computing and IIoT security, along with open challenges and limitations still (partially) unsolved.

B. Outline

The paper is organized as follows. We first establish common ground by discussing the difference between IoT and IIoT, and providing a glimpse into recent IoT security surveys. Section III briefly mentions related work and motivates the need for a systematic literature review. Section IV describes the research method used in the review and formalizes the research questions. Section V surveys the security requirements resulting from the systematic review. Section VI presents a quantitative analysis of the results obtained during the review phase. Section VII discusses the role that Fog computing might play in meeting the IIoT security requirements. Finally, Section VIII concludes the paper.

II. IoT AND IIOT

Before we discuss the results of our systematic survey in depth, it is helpful to establish a common understanding of how IoT and IIoT differ. In this section, we first explore this difference, then, we provide an overview of recent IoT security surveys.

We find Table I, taken from the ENISA “Good practices for Security of Internet of Things in the context of Smart Manufacturing” [9] report, to be helpful in outlining the differences between IoT and IIoT, and use this as a guideline throughout our work. That said, the difference is not a precise, clear-cut one, and we sometimes do deviate from these guidelines, when it is abundantly clear that a scenario concerns the IIoT without meeting relevant criteria from that table.

In general, it is accepted that IIoT is a subset of IoT: IIoT typically covers consumer devices in retail and lifestyle, IIoT focuses mainly on Operational Technology (OT), the smart manufacturing process, smart logistics, and smart cities.

It should not be surprising that the safety and security requirements in IIoT are generally stricter than those found in a typical IoT scenario. Even so, we find significant overlap in used terminology in the literature, and IIoT having stricter requirements does not necessarily mean that any proposed
There exist ample surveys investigating the state of IoT security, and we will briefly look at several relatively recent surveys, discussing how their identified security requirements might relate to the IIoT.

In [11], the authors survey the literature for real IoT attacks and present a taxonomy. They also identify integrity, anonymity, confidentiality, privacy, access control and authorization, authentication, resilience, and self-organization as security requirements for IoT systems in general. These are all represented in the requirements collected in this work as well, and reiterate that generic IoT solutions can work for IIoT systems, if they do not violate scenario-specific constraints. Neshenko et al. [12] provide a much more thorough study of IoT vulnerabilities and attacks, but do not relate these to security requirements. Nevertheless, we can see that the familiar topics of authentication and access control, assurance, and confidentiality return implicitly throughout the text. The threats described by the authors include problems such as false data injection, improper patch management, and improper encryption. Many of these can be directly connected to the security requirements listed in this work.

In [13], the authors provide a top-down survey of IoT security. They discuss security requirements for healthcare, smart grids, manufacturing, smart homes, transport, and smart cities. Some of these are also considered to be in the IIoT domain [9], and indeed the security requirements identified in these sections overlap with the ones collected in this survey, albeit on a higher level of abstraction. In each investigated domain, they list a subset of these as requirements. For smart grids, they identify availability, confidentiality, integrity, non-repudiation, and privacy, and additionally list challenges we also identify in our work: heterogeneity, scalability, privacy, and so on. What is apparent through their work, is that the main way in which the requirements for the various domains differ is in their priority, for instance, privacy and confidentiality weigh higher in healthcare than in transport. Further, the authors make the insightful observation that one specific challenge for the IIoT that is not as apparent in general IoT networks, is that its crucial safety requirements often compete with security in terms of resources. It is perhaps the balance that must be found between these two aspects that sets the IIoT apart from normal IoT systems. Indeed, whenever resource constraints are not an issue, or when safety constraints are less strict, standard IoT solutions often suffice.

### III. Related Work

To the best of our knowledge, the most recent works focused on reviewing IIoT security are [14] and [15]. The former focuses primarily on threats characterization by looking at existing attacks, while the latter mainly reviews the differences between information technology and operational technology in an Industry 4.0 setting, and discusses the challenges. However, neither of these works explicitly discuss security requirements, opting to leave them as implied by the described threats and challenges. Another recent study [16] focuses on Industry 4.0 system architecture as a whole and observes that there is an increase in security-focused architectural proposals, but does not discuss security in depth. Some older surveys dated back to 2015 and 2016 mention IIoT security requirements [17], [18], but they also refrain from an in-depth discussion.

Recently, Hansch et al. [19] published a study identifying and mapping security requirements to an OPC UA model, allowing easier machine-based verification. While they provide many security requirements, they are based on a limited set of use cases, and no thorough explanation for their derivation is given. Moreover, they are of a less abstract level than the ones we attempt to derive in this work.

As a result, we deem it necessary to provide an up-to-date, systematic survey that specifically addresses IIoT security requirements.

### IV. Research Method

In this section, we present the research method that is used in this systematic literature review on security requirements for the IIoT.

We adopt the research method detailed by Petersen et al. [20] and utilize the suggested template for describing our approach. In the next subsections, we elaborate on research questions, search strategy, study selection, and validity concerns.

#### A. Research Questions

The main aim of this work is to identify security requirements for the IIoT. This can then guide us in identifying which of these show potential to be solved by Fog computing. In addition, we want to provide an overview of the research activity in the field: how research activity has developed throughout the years, how this research was published, and what its geographical distribution is.

Thus, our research questions can be formulated as follows:

- **RQ1**: what are the security requirements of the IIoT?
- **RQ2**: how are publications related to IIoT security spread throughout the years?
TABLE II

<table>
<thead>
<tr>
<th>Query</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td><strong>in title:</strong> IIoT OR &quot;Industrial Internet of Things&quot; OR &quot;Industry 4.0&quot;</td>
</tr>
<tr>
<td>Q2</td>
<td>(<strong>in title:</strong> IIoT OR &quot;Industrial Internet of Things&quot; OR &quot;Industry 4.0&quot;) AND <strong>in abstract:</strong> security</td>
</tr>
</tbody>
</table>

TABLE III

<table>
<thead>
<tr>
<th>Source</th>
<th>Q1</th>
<th>Q2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACM</td>
<td>60</td>
<td>12</td>
</tr>
<tr>
<td>IEEE Xplore</td>
<td>2702</td>
<td>323</td>
</tr>
<tr>
<td>ScienceDirect</td>
<td>369</td>
<td>21</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>3158</td>
<td>356</td>
</tr>
</tbody>
</table>

- **RQ3:** how is IIoT security research activity geographically distributed?
- **RQ4:** what are the most popular publication venues for IIoT security research?

Answering these questions will aid in getting a better understanding of the current security landscape for the IIoT, while at the same time identifying various concrete research opportunities related to Fog computing. Each of these can then be traced back to concrete security requirements relevant to the Industry 4.0 paradigm.

**B. Search Strategy**

We utilize the adjusted PICOC criteria for software engineering [21] in order to identify relevant keywords. In particular:

- **Population:** we consider the IIoT as the application area in which our research is conducted. However, this is a very broad population, therefore, we take into account only studies addressing IIoT security.
- **Intervention:** this criterion does not apply to our research questions, as we are interested in any work in the IIoT domain that describes security requirements.
- **Comparison:** we compare the security requirements identified by different studies by taking into account such factors as the number of studies that mention them, related threats, and proposed solutions.
- **Outcomes:** we present the identified security requirements as well as the properties of their mitigation, allowing us to discuss which requirements call for further research.
- **Context:** as we do not empirically compare the available works, this criterion does not apply to our study.

With these criteria in mind, we have formulated the following keywords: **IoT, Industrial Internet of Things, Industry 4.0, and Security.**

We considered as sources the following databases: ACM Digital Library, IEEE Xplore, Elsevier/ScienceDirect. In this domain, we believe that the combination of these three sources provides an accurate representation of the research that has been conducted globally.

We divided the search into two stages. First, we queried the databases for articles related to IIoT/Industry 4.0 in general, based on their titles. This provided an overview of the amount of research conducted in this field. After that, we narrowed down our search to only include works related to security, by excluding articles not containing the word “security” in their abstract. The queries are summarized in Table II. The search results for both queries are listed in Table III. The queries have been executed in March 2020.

**C. Study Selection**

Starting from 356 papers resulting from our queries, we further filtered the studies with multiple phases.

Firstly, the JabRef\(^1\) reference management software was used to identify and delete duplicates. Five duplicates were found, leaving the number of considered papers for the subsequent phases at 351.

Subsequently, we independently reviewed the titles and abstracts of each article in order to reduce selection bias. Each article was marked as being relevant, not relevant, or of doubtful relevance. Articles were voted for inclusion when the work covered cyber-security challenges and/or solutions for Industry 4.0, and it was published before 2020, since that is the year in which this study is conducted. We do not believe that filtering on a minimum publication date is necessary at this time, due to the relatively young age of this field. Articles were voted for exclusion when the work was not related to Industry 4.0 security, contained duplicate content, or was not presented in legible English.

The following rules were used for filtering out articles based on title and abstract review (this has been done jointly by two authors of the paper):

- when both authors considered an article relevant, the article was included for the next phase;
- when one author expressed doubt and the other author considered an article relevant, the article was included for the next phase;
- when both authors expressed doubt, a joint review was done considering also other sections of the article (e.g., introduction, outline, conclusion) to determine its relevance. If this review did not clear up doubts for either of the authors, the article was given the benefit of the doubt and included for the next phase;
- when one author considered an article relevant, while the other considered it to not be relevant, the article was marked for joint review as described in the previous rule;
- when one author considered an article not relevant, while the other considered it to not be relevant, the article was marked for joint review as described in the previous rule;
- when both authors considered an article not relevant, the article was excluded.

\(^1\)https://www.jabref.org
After the individual title and abstract reviews, 68 articles were excluded and 69 were marked as doubtful entries requiring a joint review. These were then jointly reviewed, leading to an additional 35 exclusions. The remaining 248 papers were considered for full-text reading, overall reducing the number of papers to analyse by 92% compared to results of Q1 and 30% compared to Q2.

In the full-reading phase, we extracted information relevant to the stated research questions, as well as identifying the challenges discussed in the papers. We then used this data to provide a comprehensive picture of the security challenges and corresponding requirements for the IIoT. In this phase, it became clear that a number of papers were not relevant to our work, resulting in the discarding of other 43 papers. Additionally, we identified 13 papers of interest through reverse snowball sampling and added these to the selection. This brings the final number of papers considered in this survey to 218.

The entire study selection process and related numbers are summarized in Figure 1.

D. Validity Evaluation

Every study that is subject to manual selection is vulnerable to researcher bias in the filtering process. In order to reduce this issue, we performed the filtering process twice: two authors of this paper selected studies independently, and the results of the filtering process were based on a systematic approach combining the selections of both authors, and in some cases a joint review.

Also, to mitigate possible selection bias, we have performed reverse snowball sampling, allowing for the introduction of papers originally not considered due to not being captured by our search queries.

Furthermore, we have described our research process in detail, and have taken care to list the criteria by which we filtered studies. This is done to increase the repeatability of this work.

Finally, it is worth mentioning that our approach does not suffer from the Matthew’s effect, as opposed to querying databases that rank papers based on citation count [22].

V. IIoT Security Requirements (RQ1)

In this section, we present the security requirements that we found to have been discussed in the selected literature. We describe why these requirements are deemed relevant and summarize some of the proposed solutions. We also discuss why these requirements are difficult to satisfy for Industry 4.0 applications, which gives the insight needed to see why the discussed security requirements are hard to meet with conventional solutions. Furthermore, they provide a set of motivational factors for why the research discussed in this section is necessary.

We observed that the focus of the investigated literature is mainly on Industry 4.0, even if in this field highly varying scenarios are considered. For example, some articles discuss petrochemical plant management [246], while others focus on drones [177], [192], [199], and so on. Each of these scenarios has its own threat model and will thus also differ in terms of security requirements from the others, to a certain degree. However, we note that the majority of them show considerable overlap, and that even the ones that are unique to one particular scenario, might still translate into a research opportunity, or might be possibly addressed with Fog computing. Therefore, we have attempted to include all such requirements in this section, and mention their relevance to particular scenarios, to provide context.

In the rest of this section, we discuss all IIoT security requirements found in this study, grouped by the overarching categories to which they belong. Figure 2 depicts a hierarchical structure of the various subsections, together with all the works related to each subsection. References were picked and positioned using the following heuristics: firstly, if a work is mentioned in a subsection (be it in a table or the text itself), it is included in the level 1 node representing that subsection (e.g., Authentication); secondly, if a work is mentioned in a topic within a subsection (e.g., Key Distribution), it is included in the level 2 node representing that subsubsection in the mind-map. Additionally, in order to minimize redundancy in the mind-map, the following rule was followed: when a reference is included for both a subsection (e.g., Network Security) and

2In electronic versions of this work, nodes and references in this map are clickable, allowing for easier navigation through the document.
Fig. 2. A clickable mind-map giving an overview of the categories (subsections) and specific topics (subsubsections) discussed in Section V. References in this mind-map were chosen for inclusion when explicitly mentioned in the portion of text represented by each node, or when deemed relevant to the category, based on a full-text review.

one or more of its subsubsections (e.g., Wireless), then preference is given to the latter, and the reference is removed from the subsection itself. This does not eliminate redundancy between nodes of the same level (e.g., a reference can still be included for both Key Distribution and Mutual Authentication), but it does allow for a representative overview of works relevant to any topic.

Finally, in Section V-J, we close this section with a summary and an analysis of the obtained results.

Except for Section V-A and Section V-J, every section contains a table relating the most important security requirements of that category to a collection of works that we deemed the most relevant to these topics. Additionally, every table shows the research interest (low, medium, high, very high) of the scientific community for each security requirement in that category. This interest is inferred from the percentage of works identifying or addressing the specific security requirement compared to all the (unique) papers related to that category. The number of papers addressing a specific category is taken from Figure 2 as the number of papers appearing in the corresponding level 1 (i.e., subsection) and all level 2 (i.e., sub-subsections) nodes, but removing duplicates. For instance, the total number of papers discussing Network Security is given by the count of the references appearing in Figure 2 for the nodes Network Security, Latency and timeliness, Availability, and Wireless, without duplicates. It is important to note that a number of works identify multiple security requirements, thus, appear in multiple subsections; as such, the calculated percentages do not represent disjoint partitions of the set of investigated works, thus their sum will not result in 100%. The range of percentages assigned to each interest level are shown in Table IV and have been chosen based on the distribution of
TABLE IV
INTEREST LEVELS ASSIGNED TO EACH SECURITY REQUIREMENT IN RELATION TO ITS CATEGORY

<table>
<thead>
<tr>
<th>Relative interest</th>
<th>Range (x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0% ≤ x ≤ 7%</td>
</tr>
<tr>
<td>Medium</td>
<td>7% &lt; x ≤ 25%</td>
</tr>
<tr>
<td>High</td>
<td>25% &lt; x ≤ 45%</td>
</tr>
<tr>
<td>Very High</td>
<td>45% &lt; x ≤ 100%</td>
</tr>
</tbody>
</table>

Fig. 3. The CIA triad, with some examples for each property.

percentages assigned to security requirements across all categories. As an example of the interest level, consider a category AB discussed by 50 papers, and a security requirement AB-01 identified by 5 of these 50 papers, the research interest for the requirement AB-01 will be medium, with a percentage of 10%.

The aim of these tables is to give the interested reader a stepping stone to more in-depth works for each requirement, but also the topics itself.

A. The CIA Triad

The Confidentiality, Integrity, Availability (CIA) triad is a well-known information security model, and can be considered as a set of extremely abstract security goals or requirements. A subset of these lie at the root of every other security requirement. Figure 3 provides a graphical representation of the triad and shows some examples of solutions related to each property. We briefly describe the three properties as they are described by [130] below:

- **Confidentiality** pertains to protecting information in all its forms. This includes data encryption, access control, network isolation, but also privacy aspects.
- **Integrity** concerns consistency, accuracy, authenticity, and more generally the overall trustworthiness of entities.
- **Availability** concerns operational guarantees of the system. This covers topics such as redundancy and decentralization, but also guarantees that tasks will be performed within hard deadlines.

Typically, the CIA triad is used in information security, meaning that the three properties relate to information only. However, it is equally applicable in other domains, such as cyber-physical systems [132]. Indeed, many of the works we investigated explicitly mention the triad (e.g., [25], [63], [89], [97]). Traditionally the focus in industrial environments has been first on availability, second on integrity, and last on confidentiality. However, with Internet-connected systems, this requires reconsideration, and all three aspects should be brought up to an acceptable level. Thus, with the development of new IIoT and Industry 4.0 solutions, confidentiality and integrity should be weighed equally to availability.

While these three aspects are a very good starting point and are certainly important to keep in mind when specifying the security goals for any system, it is not always useful to reduce concrete requirements back to elements of the CIA triad, if one already has more (e.g., contextual) information that might help with deriving an unambiguous security goal. For example, it is easy to state that data at rest should be kept confidential, but such a requirement does not convey the conditions that a confidentiality mechanism should satisfy. Moreover, it leaves a lot of room for interpretation (e.g., confidential to which parties?). On the other end of the spectrum, very fine-grained requirements are only possible if one is developing for a specific scenario.

In the next subsections, we strive to find a middle ground where we describe security requirements at a high enough level to see where the challenges in achieving them lie, but at the same time refrain from going too deep into any scenario, although we might refer to them as anecdotal evidence supporting the legitimacy of a requirement.

B. Authentication

Authentication of remote entities (both humans and machines, or even applications) is a key concern for many forms of IoT communication [31]. Within the context provided by IoT and IIoT applications, this brings some extra challenges [15], [17], [49]. There is a need for extremely lightweight authentication schemes, with little overhead in terms of computation time and transfer size, among other things.

A second but very important concern is verifying the integrity and authenticity of data, e.g., to ensure that a configuration file was created by an authorized party, and not modified since. Also here, the IIoT domain has special requirements that prevent the adoption of commonly used authentication mechanisms. Many topics in this section therefore also concern integrity, albeit not explicitly mentioned in every instance.

Wang and Wang [82] name some other typical challenges (mainly aimed at wireless industrial communication) that need to be taken into account when investigating authentication and integrity methods. They consider extreme resource constraints, the open broadcast nature of wireless communications (i.e., anyone can read and send messages on certain frequencies), extremely large network sizes, and lack of infrastructure support.

As an example of authentication challenges in existing systems, we consider the Message Queue Telemetry Transport (MQTT) protocol. This is a widely deployed protocol for data
exchange in the industrial domain, and features some very basic and insecure authentication methods [57]. According to Katsikeas et al. [34], the protocol allows authentication through a simple username and password combination, which are communicated in plaintext. A second authentication method sometimes used is through a unique client identifier, which is easily spoofable. While it is possible to secure these methods by complementing MQTT with Transport Layer Security (TLS) or IPSec, those two protocols are too resource-intensive for many IIoT applications, and lighter alternatives are necessary, such as TinyTLS [56] or DTLS [58]. Now that industrial networks are becoming increasingly connected to the Internet, this becomes more and more important.

The importance of sufficiently secure authentication mechanisms is reflected by the fact that positions 2 and 3 in the Open Web Application Security Project (OWASP) IoT Top 10 [47] on IoT vulnerabilities concern attack vectors where (a lack of) authentication is an important aspect. Its importance is also underlined by the popularity of this topic with recent research efforts, with many papers addressing or identifying the above issues (see Table V). These works identify several authenticity properties that can be considered requirements in various use-cases in the IIoT domain. We describe these in more detail in the following subsections.

For a comprehensive survey on IoT authentication algorithms, we refer the interested reader to Ferrag et al. [31]. The authors cover many authentication algorithms and compare them based on computational efficiency, threat protection, and more. Kail et al. [33] provide another survey covering multiple industrial protocols aimed specifically at Low Power Wide Area Network (LPWAN) technologies.

Next, we discuss a number of authentication-related topics in the following subsections. First, we look at key distribution, after which we discuss mutual authentication and multi-factor authentication. Then, we address non-repudiation as a requirement, followed by anonymous authentication and privacy preservation in authentication algorithms. As a final topic, we discuss attestation techniques through trusted hardware.

### 1) Key Distribution

Key distribution is a challenging requirement for many applications in the IoT [68], and naturally extends to the IIoT. With devices being set up and used in hostile environments, possibly being very mobile, dynamically joining and leaving networks, and possibly being very constrained in resources, there is a pressing need for efficient, flexible, and dynamic key management mechanisms.

Airehrou et al. [59] argue that traditional Public Key Infrastructure (PKI) is outdated, stating that “it was at no time designed to handle the complications of managing industrial-scale networks of 50 billion devices that IoT promises to usher in.” This raises the question of whether all IoT devices should exist in the same authentication domain, and if centralized authentication authorities such as PKIs are even a sensible choice for that many devices. We will not attempt to answer these questions here.

In order to deal with dynamic environments, some naturally implied requirements for key management solutions are that they can handle node addition, revocation, as well as rekeying [63]. Resource-constrained devices will have issues with key generation, computationally intensive algorithms, and transmission of large/many messages. Moreover, in an industrial setting, device owners might not trust the manufacturer to generate keys for them, and will want to do this themselves [67]. Availability can be an issue as well. In Critical Infrastructure (CI) environments, an authentication authority has to be reachable at all times. Because of this, Blanch-Torne et al. [62] state that it is not sufficient to rely on one central authority for authentication. Additionally, they also identify transitive authentication (if $A$ knows $B$ and $B$ knows $C$, $B$ can introduce $A$ to $C$) as a requirement in some scenarios.

In [63], the authors propose a key management solution that aims for little transmission overhead by requiring only one transmitted message for one-way-authentication. While this makes for an energy-efficient protocol, it appears to not be very scalable or dynamic, since all nodes need to be known beforehand, and addition, revocation and rekeying are not thoroughly discussed.

Ulz et al. [67] propose a Bring Your Own Key (BYOK) approach, to address the trust issue between device owners and manufacturers. However, it does require devices to have Near Field Communication (NFC) capabilities, and key distribution requires a human to physically move between a central server and the device.

Another approach is suggested by [62], where there is no centralized authority, but a Distributed Hash Table (DHT) that takes care of identity propagation and lookups. Their solution is a distributed one, and also provides transitive authentication. It is scalable and dynamic, but the protocol is not designed with energy-efficiency in mind, and can require a considerable number of messages at times.
While the above-mentioned sources address the identified requirements to some extent, none of them addresses multiple at once. Clearly, there is still plenty of opportunity for novel research in this area. One potential solution to several key distribution challenges that might become viable in the future, is quantum key distribution [65], [66]. In such a system, it is impossible to eavesdrop on a transmission without altering its payload, meaning that any eavesdropping attempt can be detected.

Blockchain technologies are another promising candidate, showing potential to overcome several key challenges. Bartolomeu et al. [61] discuss Self Sovereign Identity (SSI) techniques for IoT, which build on top of blockchains to provide Decentralized Identifiers (DIDs). These systems have as a property that all entities carry their own identification data, eliminating the need for a centralized root of trust. They discuss the challenges faced by several frameworks capable of providing DIDs, some of the most prevalent being the need for a common data model for interaction between parties, and a lack of research in their application to Machine-to-Machine (M2M) authentication. A different approach is taken in [60], where the blockchain-based BCTrust protocol is extended with key management functionality. One challenge with blockchain is that due to the immutability of blockchains, revocation or alteration of data is impossible. The standard solution is to add append modifications at the end of the chain, but there exist some early results showing that small scale changes are possible using Chameleon hashing schemes. This comes at the cost of some security [64], but further research is needed.

2) Mutual Authentication: In [82], mutual authentication is identified as one of the requirements for any practical authentication scheme, and Kolluru et al. [76] state that mutual authentication between any two IoT devices is necessary, as many of them are exposed to external environments. Moreover, because of this many-to-many requirement, a user/password system is neither user-friendly nor flexible enough. It is also difficult to handle in dynamic environments. They thus identify the need for authentication mechanisms that can be used between any pair of devices, with minimal user interaction. Autenrieth et al. [71] even state that fully automated mutual authentication is a requirement. Some recent work that aims to facilitate this is done in [74], and uses trusted components such as Physically Uncloneable Functions (PUFs). PUFs are functions implemented in hardware in a way that aims to make them very hard to copy, thus being able to act as a device “fingerprint”. Another way to facilitate M2M authentication in settings where the participating devices are geographically nearby, is by using physical context such as luminosity or temperature. Loske et al. [79] survey the available literature on this so-called context-aware authentication. If the transmissions are wireless, devices can also be identified through their radio frequency fingerprint [81].

One way of minimizing interaction is by relying on biometrics for identification and authentication (although one should be careful to not use biometrics for authorization). One property of biometric-based authentication schemes is that they cannot be used for M2M authentication, as biometrics are always derived from living beings. Therefore, these types of protocols might not be feasible in every industrial context, although they adapt well to some (e.g., smart healthcare [73]). In [72], a two-factor mutual authentication method is proposed, combining smart cards and biometrics, although recent work shows that their protocol is not secure against various attacks [75]. Li et al. [77], [78] use a combination of user/password and biometrics instead as a two-factor approach, while Deebak et al. [73] combine smart cards, passwords, and biometrics. The proposed methods claim to be very lightweight – but reliance on biometrics by itself requires specialized hardware (or some non-trivial computational capacity to process, e.g., audio or video signals), which might not always be an option. Further, it typically requires physical proximity, although recent work [80] shows that remote biometric authentication is a possibility.

Another way to minimize user interaction is by deriving identities through analysis of behavioral patterns. This shares the property that it cannot be used for M2M authentication with biometric-based methods. The Fifth generation cellular network technology (5G) authentication scheme for smart devices proposed in [70] uses Cloud-based learning to dynamically identify and authenticate users based on behavioral patterns, showing another approach for minimizing user interaction. This concept has also been used in the field of intelligent vehicles whereby drivers are identified by their driving behavior [69].

3) Non-Repudiation: Non-repudiation is a message property, ensuring that the author of a message is not able to later repudiate (i.e., deny) their authorship of that message. Non-repudiation can also extend to concepts other than messages (e.g., an entity cannot repudiate their accountability for an action that was started/requested by them).

Fraile et al. [84] provide some concrete examples showing why non-repudiation can be considered a security requirement. Firstly, users might perform illegal actions, and the system needs a way to track these actions. If these actions are reproducible, the system becomes susceptible to log injection attacks, an observation echoed by Ankele et al. [83]. Another example mentioned is the situation where a manufacturer finds out that their configuration files on some hardware have been deleted, after the hardware vendor has performed updates to this system. Without a non-repudiation mechanism in place, the deletion of these configuration files cannot be unambiguously traced back to the software update. Another example can be found in [89], where the challenges in applying the Assurance Case methodology for the IoT are laid out. Assurance Cases are structured arguments, for use during, e.g., software development, that show that certain properties of a system hold. The authors of this work identify non-repudiation as a requirement for the assurance of security properties of a system. The blockchain-based authentication and access control scheme described in [87] also states that non-repudiation is an essential property.

Li et al. [86] propose a certificateless authentication scheme for Wireless Sensor Networks (WSN) environments. The advantage of their approach is that, because some of the heavier computations can be moved to third parties (e.g., a gateway), the computational requirements on sensor nodes
themselves can remain low. Their protocol achieves non-repudiation by ensuring that messages are publicly verifiable. Certificateless schemes are a fairly popular topic in this domain. More recent examples of work on similar schemes for IIoT are [85], [88], [92] (broken in [91]), and [90].

4) Anonymity and Privacy: Anonymous authentication is verifying the authenticity of an entity without disclosing that entity’s identity. This is necessary in situations where one wants to protect the privacy of users. Lin et al. [87] identify the need to protect users from being identifiable when an adversary has access to the authentication service. Cui et al. [93] also mention privacy-preserving access control. One example of a threat due to lack of anonymous authentication is provided by [72]: an adversary could conduct traffic analysis to create profiles on sensitive assets in an industrial environment, and possibly derive sensitive data from those profiles. Paliwal [94] proposes a hash-based privacy preserving authentication scheme specialized for WSNs scenarios. In this work, a variety of requirements are identified for schemes for WSNs, although these mostly relate to low-level properties that generalize to any secure authentication scheme, such as resistance against replay attacks. Because of this, we consider these to be too low-level to be included in our analysis as is, but rather as implied by other requirements.

In [87], a public blockchain-backed authentication mechanism is proposed, thereby turning user anonymity into a hard requirement. The work in [93] does not rely on a blockchain, but relies on a server to provide computational aid (in a secure manner). While both proposed schemes use Attribute Based Signatures (ABS) as cryptographic constructs, the two approaches cater to different goals: blockchains are widely considered to be resilient and highly available systems, which can be useful in scenarios that require these aspects, while server-aided encryption schemes target low-power devices with very limited computational ability or battery life.

5) Attestation: Attestation is a method for detection of unintended and malicious changes to software [17]. Doing this remotely can provide guarantees on the integrity and authenticity of a piece of software that is being run on a remote system, and therefore allows one to place more trust in a remote system than is possible in a scenario without remote attestation.

Because attestation aims to enable these higher levels of trust, it poses very strong security requirements on hardware. At the same time, remote attestation methods implemented purely in software typically have to rely on very strong assumptions that are hard to achieve in practice [17]. Attestation can be done in a practical setting through the use of Trusted Execution Environment (TEE) provided by trusted hardware, such as ARM TrustZone [247], Intel SGX [248], or implementations of the Trusted Platform Module (TPM) standards [249]. Not all of these might run on low-end hardware, but some recent embedded controllers contain trusted hardware components [250] that also enable attestation to some extent.

References [17], [95], and [97] all identify the need for remote attestation, in order to increase the system’s resilience against intruders. Especially in contexts where parts of an overall system are deployed in hostile environments, where it is important that the correct functioning of the software is continuously verified. Additionally, Laaki et al. [96] also identify the possibility for hardware attestation to protect the digital twin representation of proprietary hardware setups.

As mentioned in [17], there has not been a lot of activity on trusted hardware in this domain as of yet, with most of the available attestation protocols proposed so far aiming for a more general-purpose scenario, not taking into account aspects that make integrity and authentication protocols for the IIoT a challenging domain.

C. Access Control

Access Control (AC) is necessary in a wide variety of situations; already when a device allows for two modes of interaction, one for normal user behavior and one for system administrators to deploy updates, a rudimentary form of access control is needed. Furthermore, a lack of adequate privilege separation has been identified as one of the most severe shortcomings in existing systems, such as the Supervisory Control And Data Acquisition (SCADA) protocol [100].

AC invariably relies on authentication methods, as one needs to authenticate users in order to enforce access policies. It is therefore not surprising that AC mechanisms inherit many of the authentication requirements described in Section V-B. The challenges in access control relate to resource consumption, but also availability. In highly distributed scenarios, it should not happen that AC policies are unavailable due to a connection failure.

Aiming to minimize energy consumption for lightweight devices, Li et al. [86] propose a certificateless signature scheme as well as an AC framework for WSNs. This is made possible by relying on a (collection of) trusted systems in the network that are powerful enough to perform a part of needed cryptographic operations. The lightweight devices then cooperate with the trusted systems to create cryptographic signatures. Some natural security requirements are mentioned, such as the CIA triad and non-repudiation. Beltrán et al. [24] also target low-power devices, but they explore a setting in which these resource-constrained systems interact with Cloud services. In this scenario, they identify the need for identification, authentication, authorization, and accounting mechanisms. Furthermore, they state that depending on the particular application, fine-grained authorization control might be needed, or the ability to handle dynamically changing privileges. In some other situations, they state it is useful to manage access policies centrally. However, in order to be compatible with many systems from different developers, some form of federation is needed too. In order to address these issues, they propose a token-based federated authentication scheme that makes use of PUFs to meet the energy constraints of low-power devices. The resulting authentication scheme is flexible enough to act as a building block for many types of authorization mechanisms.

The blockchain-based authentication protocol proposed in [87] also contains an AC framework, and tackles the availability and single point of failure challenges through use of a blockchain, and a DHT containing AC policies. An additional feature of this work is that it respects the privacy of
users through the use of ABS techniques. A different approach is taken by He et al. [102]. In this work, ring signatures are used to construct a distributed lightweight AC framework. This framework specifically targets WSNs and achieves user anonymity by grouping users with similar rights, ensuring that AC authorities cannot differentiate between signatures from users in the same group. Lahbib et al. [104] also propose a blockchain system, identifying the need for dynamic access control and distributed governance. They utilize smart contracts and leverage the non-repudiation and integrity inherent to blockchain systems to propose a resource management framework, with fine-grained AC built in. Yao et al. [109] share the sentiment that distributed AC is needed, but propose a Fog solution based on attribute credentials.

Kim et al. [103] consider a scenario where nodes in multihop Low-power Lossy network (LLN)s want to communicate with each other. They also identify the need for federation, but from a reliability perspective. In order to guarantee the availability of a system, it cannot rely on a single point of failure for access control enforcement. At the same time, they identify the need for a transparent scheme, that is also scalable. Decentralized protocols such as the one proposed in their work, can increase scalability, as changes are propagated much more organically through the network, than with a centralized structure, avoiding congestion issues.

In the work presented by Chen and Ng [99], AC and authorization are also identified as one of the major challenges for the IIoT. They propose an access control framework for a scenario where the owner of an IIoT device has the right to control the AC policies of their device, and wants to set up fine-grained policies. At the same time, a large number of IIoT devices are shared by multiple entities that can interact with them based on these policies.

Preuveneers et al. [107] argue that identity management is crucial for AC purposes, and propose a framework handling identities, authentication, and authorization in a networked production scenario. They also raise the point of compatibility with legacy devices, which is worth considering in any IoT environment.

Vanickis et al. [108] make the observation that due to the increase in frequency and sophistication of security attacks in recent years, there is a need to include risk assessment in the process of specifying AC policies, and that as a result of these trends there is a growing interest in Zero Trust Networking (ZTN) protocols as opposed to perimeter-based security. The principle behind ZTN is to treat the intranet with the same level of trust as the Internet. Their proposed policy enforcement framework is built upon this principle, and is able to provision firewalls across different segments of a network.

D. Maintainability

Maintainability concerns the ability to configure, reconfigure, and update (parts of) a system. In Industry 4.0, these concepts become crucial as the software and configuration of IIoT systems must have the ability to be changed, in order to provide protection against previously unknown security threats [116]. Updateability can be considered a countermeasure against security attacks, since it allows for continuous changes to firewall configurations as threats are identified, as well as software patches for newly discovered software vulnerabilities. As we will see in this section, the challenges related to maintenance are again related to resource constraints and the dynamism of IIoT environments, making traditional maintenance solutions insufficient to adequately address the needs in this domain.

In [110], George and Thampi state that the availability of security updates is a critical concern for IIoT devices, but that due to some IIoT systems being so lightweight and the infrastructure not being fixed, it is extremely difficult to always patch all devices in a network. To mitigate this, they describe an approach that ensures update deployment on high-risk vulnerabilities, to reduce the risk of serious attacks on the infrastructure. For this, they propose a number of risk mitigation strategies that can be used to help identify the devices most in need of updates. Yadav and Paul [114] also identify the timely application of patches to all vulnerable systems in a network as a problem, and propose a patch prioritization method to mitigate this.

In addition, some IIoT systems require the ability to be updated without any disturbance to the service they provide. Mugarza et al. [111] propose a secure updating mechanism for mixed-criticality systems. However, their approach requires the ability to run and monitor updated binaries in a sandboxed mode. Not every device has the resources for this. They follow up on this research with an application of their system to a smart city scenario [112]. The proposed update process is in accordance with several safety standards, a requirement identified in Section V-E.

According to Seitz et al. [113], updating IIoT systems is often complex and cumbersome, and requires an expert technician to perform the update, which can be a lengthy process. This does not scale with the increase in connected devices.
and therefore the update process must be streamlined and simplified, with minimal possibilities for errors due to human behavior. Their suggestion is a marketplace, not unlike those seen on smartphones. In addition to usability, they state that update management of devices should be possible both on-site and remotely, and updates and installations must be logged so that they are traceable, for transparency and in case of problems. While their proposed solution appears as a global and decentralized marketplace, this might not be a good fit for every type of device, especially when the functionality of such a device is secret. Moreover, it raises questions about how much power can be given to app developers and where the trust in a system should lie, which are topics that can be highly dependant on a specific scenario, and are worthy of further investigation on their own.

Another problem is mentioned by Ulz et al. in [67], wherein they state that cryptographic keys also require the possibility to be updated securely. This can be interpreted as a requirement relating to the maintainability of a system’s configuration, and is an argument against the deployment of hardcoded keys at manufacturing time, which sometimes happens in production environments. In a later work, they take this notion further, and propose a hardware device, that can be temporarily attached to a system to allow for secure updating and reconfiguration [52]. The updates are verified and installed in an isolated environment provided by the special hardware, for increased security and traceability, but still allows for remote queuing and deployment of updates, to some extent. However, this approach might not be practical in environments where it is hard to physically reach all deployed systems.

1) Smart Maintenance: Industry 4.0 enables smart maintenance, which is essentially predictive maintenance of (parts of) devices based on remote data collection about their usage. This allows for a more streamlined production line where system downtime and maintenance costs are reduced to a minimum. Its relevancy is underlined by the inclusion of continuous maintenance and maintenance frequency being used as measurable safety indicators in a meta-model proposal for automated security dependability detection within IIoT systems [25]. Priller et al. [117] provide a case study on this subject detailing a number of smart maintenance security requirements, notably the ability to update as well as secure communication channels themselves.

Lesjak et al. [36] reason that smart maintenance requires secure communication channels, as status information of machines is sensitive data. Moreover, the maintainer needs the ability to verify the validity of this data. They argue that there are systems for which it is essential that they are exposed to the Internet as little as possible, and propose solutions using NFC to permit secure transmission of data to the maintainer, as well as identity provisioning over NFC [37]. The specific requirements identified in this work are the need to support legacy devices, prevent data leakage, protect against Internet access, protect the validity of the maintenance data (towards the maintainer), and protect transparency of the communicated data (towards the customer). In a later study, Lesjak et al. [38] propose an MQTT-based approach where they add a further requirement that data transmission must not cause safety-critical interference, so that operational functionality remains unaffected.

E. Resilience

The Industrial Internet Consortium (ICS) has published an IIoT security framework [123] in which they define resilience as “the emergent property of a system that behaves in a manner to avoid, absorb and manage dynamic adversarial conditions while completing the assigned missions, and reconstitute the operational capabilities after causalities”. This definition overlaps with several aspects of system trustworthiness such as safety and reliability, but also security. Indeed, [15] and [118] identify resilience as an important security challenge for the IIoT. The implication that resilience requirements bring to the security domain are that security technologies should provide the capability to continue normal system operations if parts of the system are considered compromised. This could for example be done by rerouting tasks to other capable components, or through other means, often belonging to one of three canonical approaches identified by Laszka et al. [126]: redundancy, diversity, and hardening.

The manner in which this requirement should be satisfied, depends heavily on the scenario. In a WSN, it might be acceptable to simply deploy enough sensors to guarantee some redundancy, meaning that a small number of compromised sensors can be kept contained and their output discarded until the issue has been addressed. In a power plant however, it might be catastrophic to disable one generator entirely if one of its components has been compromised. Instead, it might be possible to provide the compromised components’ functionality in some other way, or temporarily reroute energy from other generators to guarantee some level of operations.
Fraile et al. discuss device driver security in a connected virtualized factory environment [84]. They identify multiple resilience-related issues, one being that intermittent connectivity might cause loss of history if status information should be continuously sent to a centralized database or the Cloud. Their proposed solution is to keep local databases that keep a short-term history that can be synchronized with a back-end once connectivity is restored. Another identified issue is to avoid system failure, in case of a compromised device driver. The authors propose redundancy and smart fallback mechanisms to adapt to possible threats. The difficulty in a fallback mechanism is that it requires the exact same configuration and as much as possible of the current system state of the normal system, in order to allow for rapid recovery. This is not only difficult because state replication can introduce considerable overhead, but it also means that the fallback system is vulnerable to the same threats as the normal system. To mitigate this issue, the authors propose introducing some diversity in the fallback system. The proposed solutions in this work are all rather specific to the considered scenario and architecture, but use elements that are common in resiliency mechanisms in general.

When looking at low-energy devices, WSNs have been identified as a way to increase the robustness of SCADA systems against network failures, due to their distributed and self-organizing nature [125]. However, major concerns exist regarding their ability to communicate securely, and the ability to interface with some proprietary SCADA protocols. The authors also identify a number of challenges relating to the security of WSNs and propose a decentralized multi-agent architecture to remedy a number of these.

In [25] and [120], a number of measurable indicator points are identified, among which those relating to resiliency. In the latter, they then use these indicator points to propose a method for automated standard compliance testing in the Industry 4.0 domain. Standard compliance is a powerful aid in verifying the resilience, reliability, and safety of a system, and can be applied to a wide spectrum of devices. Related to standards compliance, Bauer et al. [119] investigate European Union Agency for Network and Information Security (ENISA) guidelines on secure Cloud services, and extract a number of measurable security metrics that relate service level agreement objectives between Cloud providers and their (industrial) customers to concrete responsibilities. These metrics could also be used in compliance testing. In this work, reliability and redundancy are also identified as measurable indicators. Similarly, Leander et al. [127] investigate the applicability of the IEC 62443 cybersecurity standards [124] in Industry 4.0 applications. For a short survey on the security standards relevant to Industry 4.0, we refer to [122].

F. Data Security and Data Sharing

In today’s world, data security is critical in nearly any digital environment, and the IIoT is no different. Many of the works investigated in this survey identify confidentiality of data as a security requirement in some form (e.g., [17], [36], [49], [71], [133], [136]). Traditionally however, availability and integrity are considered more favorable than confidentiality for industrial environments [129], [132], as they have a measurable economic impact. This is not a sustainable viewpoint in an era of connected devices, and is changing fast now that companies seek to connect their systems to the Internet.

In a survey among companies, Autenrieth et al. [71] found that they too consider data security to be one of the critical factors for migration to Industry 4.0, a finding confirmed by another study conducted by Moyne et al. [136]. In this work, the authors additionally state that companies are hesitant to adopt data-sharing based technologies (Cloud, smart security, and maintenance, fault detection and prevention, etc.) as there is no evidence of these technologies being safe or secure when it comes to protecting intellectual property, as a result of which they identify the need for a standardized way to achieve intellectual property protection in the presence of data sharing mechanisms. The sentiment that companies are reluctant to rely on Cloud providers for data storage and sharing is shared by Esposito et al. [29]. However, they also note that most data breaches come from within companies, and not Cloud providers. They propose a cloud storage solution that aims to minimize the attack surface both in the Cloud and within the company. They identify data loss mitigation as another requirement, identifying four key elements for an effective solution: prevention, identification, notification, documentation.

The challenges in this domain relate to three colliding factors: Firstly, due to the heterogeneity of devices, data security mechanisms need to be able to operate with extremely few resources. Secondly, due to the criticality of some IIoT applications, the data security requirements are very high. Thirdly, many smart capabilities are enabled by the sharing of data, but in industrial contexts, data is often sensitive and confidentiality is of utmost concern, which poses a dilemma.

Data security covers a wider area than just encryption techniques. One of the vital aspects of the Industry 4.0 paradigm is making smart use of available data. This inevitably involves sharing data with other entities, that can be anywhere from a part of the system to being outside the organization boundaries. As an example, consider the discussion on the sharing of device usage metrics in Section V-D1. Even if no other data
is shared, usage metrics will have to be sent to the device manufacturer to enable smart maintenance, but might also be used to deduce sensitive information such as production volume. A similar example would be data analysis for anomaly detection (Section V-G). While encryption techniques do offer ways to aid with partial sharing of data, we will also discuss other ways of keeping data confidential.

1) Data Transport: The MQTT protocol is widely used for data sharing between industrial systems, but by itself only supports user/password authentication, and provides no security measures on the network or application layer. This becomes problematic especially in the context of the IIoT. In order to remedy this, Lesjak et al. [38] propose using TLS as a secure layer upon which MQTT can function. While this provides all the security benefits of TLS, it does add considerable overhead to the edge devices that will now have to manage TLS contexts. In their work, the authors propose using a trusted hardware extension at the edge devices that can store keys and also manage the TLS context. While modern devices might have access to cheap trusted hardware, this is not always possible with legacy devices, therefore, other solutions will need to be investigated. Katsikeas et al. [34] also observe that TLS can be used to secure MQTT communication, but note that this will not work well in WSNs due to severe resource constraints. Therefore, they try to minimize the overhead by encrypting messages at the link layer. In a later work, Lesjak et al. [40] observe that an often-needed requirement is communication with other stakeholders, e.g., equipment manufacturers (for smart maintenance) or nearby links in a supply-chain. To enable authenticated, secure data communication between these, the authors propose a hybrid multi-stakeholder protocol on top of MQTT that allows end-to-end encryption of payloads that need to be transmitted to external parties.

Alternatively, more modern protocols such as The OPC Unified Architecture (OPC UA) [145] have authentication and encryption support [34], [143], and hardware acceleration for the cryptographic primitives used in these is starting to appear in lightweight products [251]. Adoption of the OPC UA could thus help in meeting some of these constraints. One recent experimental deployment combines this with trusted hardware to facilitate secure connections [142], but acknowledges that further research is needed. Finally, it is worth noting that regardless of the security protocol used, from an energy and efficiency standpoint, there is a case to be made for selectively encrypting only those messages that might harm the system if tampered with. In [144], the authors propose a symbolic analysis model that can identify such messages.

2) External Parties: Data confidentiality when at rest or in transit, is often realized through cryptographic means. The challenges in finding suitable ciphers for the very diverse IIoT environment are described by Zhou et al. in [55]. Again, the main challenges appear to concern energy and other resource constraints. Irrespective of the cipher used, the authors also identify the key distribution and management problem, as previously discussed in Section V-B1. More generic challenges are described by Yu et al. [160]. They argue that Reliable storage, convenient usage, efficient search, and trustworthy data deletion are some of the major issues for Cloud and Fog scenarios.

As the Cloud promises a large amount of storage and computational resources, Cloud connectivity is often necessary for Industry 4.0 applications. With a suitable encryption scheme, data might be stored securely in the Cloud [98], but even then it is not possible to interact with it in any way other than retrieving it for decryption. Seeking to remedy this, there has been a recent increase in research efforts in modern cryptographic techniques such as homomorphic encryption, allowing for computation on encrypted data ([65], [66]), and searchable encryption, enabling search operations on encrypted data ([146], [161]). Specific to the IIoT data sharing scenario, Deng [147] proposes an anonymous aggregate encryption system that allows IIoT devices to encrypt data into one ciphertext that can be decrypted by multiple recipients with their individual keys, while retaining their relative anonymity.

Fu et al. [149] propose one way of ensuring confidentiality in the Cloud, while maintaining the ability to search through data sets, through a privacy-preserving encryption scheme. Deployed IIoT devices transmit their data to special (on-site) servers which aggregate the data, remove redundant entries and prepare it for storage in a Cloud-backed database by indexing and encrypting it. Users can then search this database through trapdoor queries, meaning that the search process can be performed on the encrypted data. In order to obtain the searched data, users can download the encrypted results, and use their private keys to decrypt them. As a result, the Cloud environment will never have any access to the unencrypted

<table>
<thead>
<tr>
<th>ID</th>
<th>Security requirement</th>
<th>Related sources</th>
<th>Relative interest</th>
<th>% within category</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSS-01</td>
<td>data loss mitigation</td>
<td>[29]</td>
<td>Low</td>
<td>2%</td>
</tr>
<tr>
<td>DSS-02</td>
<td>data confidentiality</td>
<td>[36], [30], [52], [53], [55], [98], [105], [130], [135], [139], [149]</td>
<td>Medium</td>
<td>19%</td>
</tr>
<tr>
<td>DSS-03</td>
<td>standardization</td>
<td>[136]</td>
<td>Low</td>
<td>2%</td>
</tr>
<tr>
<td>DSS-04</td>
<td>secure data transport</td>
<td>[34], [38], [40], [137], [142], [143], [145]</td>
<td>Low</td>
<td>7%</td>
</tr>
<tr>
<td>DSS-05</td>
<td>secure external data storage</td>
<td>[29], [65], [66], [98], [131], [139], [141], [146], [156], [159], [161]</td>
<td>High</td>
<td>34%</td>
</tr>
<tr>
<td>DSS-06</td>
<td>data flow control</td>
<td>[162]–[164]</td>
<td>Low</td>
<td>5%</td>
</tr>
<tr>
<td>DSS-07</td>
<td>data protection legislation compliance</td>
<td>[50], [98], [165]</td>
<td>Low</td>
<td>5%</td>
</tr>
</tbody>
</table>

TABLE IX
DATA SECURITY AND DATA SHARING RELATED SECURITY REQUIREMENTS, SOURCES THAT IDENTIFY THESE, AND THEIR INTEREST LEVEL RELATIVE TO THE CATEGORY. THE RELATIVE INTEREST LEVEL IS BASED ON THE PERCENTAGE OF WORKS ADDRESSING THE SPECIFIC SECURITY REQUIREMENT COMPARED TO THE TOTAL NUMBER OF PAPERS FOR THAT CATEGORY.
data. Xu et al. [159] propose a similar solution, also relying on trapdoors to perform search queries on encrypted data sets in the Cloud. The difference is that in this solution, the used encryption techniques aim to be lightweight enough to allow for decryption by the IIoT devices (specifically sensor nodes) themselves, without requiring an intermediate server. This approach only targets data storage, search, and retrieval. Other use cases for Cloud environments, such as big data analysis in the Cloud itself, cannot be solved using this method. Miao et al. [155] also propose a Cloud-assisted method in the context of an e-health scenario, while attempting to minimize intensive tasks such as decryption and decryption at the Edge side, to computation requirements and power consumption.

With the advent of blockchain technology, there has been an increase in interest in data sharing solutions based on decentralized ledgers. Sani et al. [157] propose a privacy preserving blockchain using mutually authenticated encryption for confidential data exchange, while others propose things such as energy trading [150] and big data markets [152]. Huang et al. [151] list three main challenges in blockchain technology: the trade-off between efficiency and security, coexistence of transparency and privacy, and conflicts between concurrency and throughput. These concerns are shared by Nikander et al. [156], who discuss throughput, latency, and resource requirements more in-depth. Further, they identify four models of operation for lightweight devices to participate in blockchains. Another proposed solution is to integrate devices with multiple ledgers, although the authors state that this is an active field of research. The aforementioned concerns are also identified in [148], where the authors further state that while blockchain promises enhanced data security and availability, for the IIoT domain there remain challenges regarding data privacy, integrity, and identity certification. They also list interoperability, standardization, and regulatory aspects as more general blockchain challenges. Other blockchain-based proposals in this domain are [153] and [154]. For a more thorough discussion on security requirements and challenges for blockchain in the IIoT, we refer the interested reader to [35], and for a discussion on risky characteristics common to blockchain technologies we point to [158].

3) Data Flow-Control: Through data flow control, data access policies can be enforced on a higher-level than encryption techniques, which provides a way to address security- and privacy requirements relating to the processing of data as it moves in a system.

Al-Ali et al. [162] describe a real-world use case for data flow monitoring, where certain data on machine error rates is shared within the company itself, and across organization boundaries based on a set of privacy policies. Some of these policies cannot be statically enforced because they depend on dynamically changing processes or coordinated interaction between different entities. They conclude that the ability to capture dynamic situations is a challenge that has yet to be overcome.

Identifying data security as a design requirement, Bloom et al. [163] investigated input-output patterns in existing IIoT applications in order to gain a better understanding of ways to secure information related to IIoT operations. Based on their observations, they propose some design patterns that can help protect data flow already in the design stage. Schütte and Brost [164] state that data flow enforcement is a requirement in certain contexts, and propose a policy-controlled data flow control framework capable of monitoring messages between entities both statically and at run-time. This allows users to not just specify access policies, but also to state how data elements are allowed to be processed by the system. Whether dynamic monitoring with this solution is possible in time-critical systems, is still a subject for further study.

4) Data Privacy: Data privacy and ownership is an important topic for many companies and governments, and with the recent popularity of Cloud storage services, these issues require careful consideration [98]. With the amount of data that is generated by modern devices, it becomes possible to create detailed profiles of users, putting their privacy at risk [168]. In an attempt to mitigate this, an anonymous data collection framework is proposed in [169].

With recent legislation in the European Union (the General Data Protection Regulation (GDPR) [166]) effectively requiring privacy-by-design for all products, data privacy should be taken seriously by manufacturers as well. Preuveneers et al. [50] discuss the implications of the GDPR in Industry 4.0 and smart factory environments. For example, some requirements derived from this legislation are that (in general) customers of a service have the right to retrieve their personal information, the right to be forgotten, and the right to erasure of their personal information. This should be taken into account when designing systems that interact with humans and might collect such information. Acknowledging this need for integration, Conzon et al. [165] describe a model-based framework for IoT, the security and privacy principles of which are derived from the GDPR.

Privacy does not only concern data collection and Cloud storage, but also requires the obfuscation or omission of metadata and other properties that can be leveraged by adversaries. For example, in WSN networks, sensor nodes are often spread over a geographically wide area, and an adversary might attempt to locate the source node of specific traffic based on message flow. To remedy this, source location privacy schemes should be deployed such as the one proposed in [167].

G. Security Monitoring

Dynamic monitoring of behavior in a system is an effective way to detect and respond to malicious activity, and systems that provide these capabilities are commonly known as Intrusion Detection Systems (IDSs). In the IIoT domain, two commonly identified security requirements are the ability to monitor infrastructure, and respond to known and unknown threats when necessary [55], [193], [198], [206]. The reason these are deemed particularly important for the IIoT comes from the fact that older, less secure devices are likely to be connected to the network as well [208]. These devices cannot always be patched to protect against known vulnerabilities, and therefore require continuous monitoring. An example of this is the IDS proposed by Kim and Kang [189], which specifically targets the Modbus protocol, a widely used industrial
control protocol, and a good example of an existing protocol severely lacking in security mechanisms. Similarly, the MQTT protocol has been covered like this [178]. A second reason can be found in providing protection against Denial of Service (DoS) attacks [115] and improving congestion control in general [187].

Hasan and Moustafa [184] state that latency is one of the major challenges for security monitoring systems, due to the geographical distance between devices in certain Industry 4.0 networks, network latency can become too high for acceptable response times to intrusions, especially when using Cloud security services.

Another identified challenge for security monitoring in the IIoT is the imbalance of data sets. Due to the sheer amount of data generated by IIoT devices and the low attack frequency, obtained data sets that can be used for machine learning approaches to intrusion detection tend to be very imbalanced [206].

Many proposed IDS solutions exist that are designed to work in the general IT domain. However, it becomes harder to monitor threats when taking into account the extreme environments in which some IIoT appliances are deployed, resource constraints, and data privacy requirements. On the other hand, as IIoT system activity is largely the result of automated processes, the traffic patterns tend to be fairly static and periodic, making it easier to perform accurate anomaly detection [180], [208]. Additionally, this predictability introduces the possibility for utilizing these patterns against the system through stealthy injection attacks [196], or to establish covert communication channels, as demonstrated in [172], and should be monitored against. In [176], Bernieri et al. show that this predictability can be used against attackers by developing a honeypot for a water distribution system. It simulates physical processes, and is able to detect attacks that aim to modify the system’s behavior. A machine learning based IDS capable of detecting these types of attacks proposed in [181]. However, Genge et al. [183] note that when monitoring the output of physical processes, care has to be taken to take the gradual decay of processes (e.g., the wear on equipment) into account. They show that this can be done through statistical analysis. As the authors observed, there is very little work done in this area, and more research is needed to develop sophisticated measures that incorporate for process aging.

Settanni et al. [198] propose a self-adapting IDS that detects anomalies in the range of certain control values. Their solution requires the continuous collection of logs of all connected devices to a central control system, which is acceptable in environments with reasonably powerful machines, but not in WSNs or other sparse environments with lightweight nodes. The anomaly detection algorithm for physical quantities proposed by Zugasti et al. [208] similarly looks at observed quantities. However, in this work, no attention is given to the resource overhead of this approach, nor where it should be deployed in an IIoT system.

Very recently, there has been a surge in interest in machine learning techniques for anomaly detection in IIoT. For example, [200] and [203] provide a performance comparison of various machine learning algorithms for detecting anomalies in IIoT, in [171], Al-Hawawreh et al. propose a deep neural network approach for use in brownfield installations, in [170] the authors propose a similar system for ransomware detection, and in [197], the authors employ machine learning to detect time synchronization attacks. In [173], Alem et al. acknowledge the power of machine learning, but warn against high potentially false-positive rates. To mitigate this, they propose a hybrid system, that derives a semantic model from the ISA95 standard. Then, using a neural network for anomaly detection, they can filter out false positives and categorize anomalies based as being malicious or just dysfunction. Deep learning IDSs do not come without risk. In [186], the authors show that one can reliably create adversarial samples that defeat deep learning based systems. The findings in [188] agree with this, as also there the authors manage to bypass machine learning systems. Additionally, they show two methods of increasing resilience through retraining of the networks. Robustness against adversarial samples is something that needs to be taken into account when using machine learning for security monitoring. For a more thorough overview of the state machine learning for industrial IDSs, we refer the interested reader to [207].

Moustafa et al. [193] identify the requirement for IIoT monitoring services to handle a large amount of heterogeneous data sources. Their proposed solution uses Markov models and a central processing system (with parts running both in the Cloud and Fog). The data collection itself happens through middleware, thereby minimizing the overhead on resource-constrained devices.

Threat response is a requirement identified by many works in this area, e.g., [55], [198], [205]. While this is usually in the form of notifying security personnel and mitigating the threat by stopping the service, Babiceanu and Seker [175] use the flexibility provided by Software Defined Networkings (SDNs)
to let the network operate in multiple modes, increasingly trading quality of service for security.

Whereas some approaches focus on intrusion detection in one layer in the Edge-Cloud spectrum, Yan et al. [115] propose a monitoring framework that contains systems operating in the Edge, Fog, and Cloud layers. This way, resource overhead for extremely lightweight Edge devices is kept to a minimum, while at the same time allowing localized management and response through the Fog layer. The Cloud layer uses data analysis approaches to intelligently detect attacks. This is similar to the DDoS mitigation approach proposed by Zhou et al. [205], where local virtual network functions, Fog, and Cloud work together to respond to DDoS attacks.

Another aspect of security monitoring concerns the continuous monitoring of network traffic ensuring that network security policies are not violated. This type of monitoring is to help maintain the integrity required of Industry 4.0 network infrastructure, and as such does not target devices themselves, but rather SDN controllers and routing devices. Melis et al. [191] propose a live monitoring solution of flow permission controls, as well as a proactive formal verification mechanism of the security policies in SDN systems.

That security monitoring can also be proactive, can be seen by looking at the fuzzing frameworks proposed by Flores and Mugarza [182] and Niedermaier et al. [45]. The authors of the latter propose a fuzzing framework, that continuously tries to “attack” networked services with randomized data streams. It is lightweight, and is able to identify vulnerabilities due to common software bugs such as buffer overflows. However, with an approach such as this, care has to be taken that system performance is not affected, and that critical services remain available. As such, fuzzers might mainly be a tool for security researchers, and developers aiming to create a highly secure product. But when deployed carefully, production systems can also utilize them to detect configuration errors and vulnerabilities.

That IIoT environments can benefit from specialized monitoring approaches can also be seen when looking at drone scenarios. In their behavior and vulnerability assessment, Sharma et al. [199] identify a number of security requirements that are specific to this scenario, as well as several requirements that are more generally applicable. Specifically, they identify the need for: identification mechanisms; continuous monitoring; predictive and highly accurate vulnerability assessments; and the ability to identify anomalous drones to be marked by the monitoring service, so that this information can be shared with all drones in a swarm. Their solution utilizes Petri Nets to monitor behavior. Some other proposed monitoring solutions aimed at drone scenarios are based on behavior rule specifications [192] and recursive parameter estimation [177]. Another example of specialized security monitoring is provided by Deshpande et al. [179] propose a heartbeat protocol catering specifically to WSNs, ensuring that overhead on the sensor level is minimal.

H. Network Security

Achieving adequate network security consists of many things, including authentication, secure transport, reliable and secure routing, and more. In previous sections we already discussed some of these, and will therefore focus on network infrastructure security.

With industrial networks becoming increasingly complex due to a large number of connected devices, we are faced with problems similar to those that occurred during the rapid expansion of the World Wide Web [187]. Because of this, many performance and scalability issues need to be addressed, such as bandwidth and latency contention. According to [212], many configuration, traffic control, and security systems rely on proprietary software which make integration in generic management frameworks impossible. At the same time, they state that network infrastructure is required to be flexible, to handle dynamic environments. To solve this challenge, two paradigms aimed at separating configuration and control from data transfer itself have been gaining traction: SDN and Network Function Virtualization (NFV). SDN concerns configuration and management, while NFV concerns virtual environments to run network and security functions on a layer that is abstracted the devices on which it runs. The authors propose an architecture using these paradigms to enforce security policies on switches with SDN and NFV capabilities, and move away from, e.g., firewalls. They essentially attempt to address four security requirements through this approach: the ability to specify and enforce network security policies, to minimize management and configuration overhead, to allow for dynamic reconfiguration of the network and its security policies, and to minimize the overhead caused by enforcement of security policies. Other points where SDNs can improve system security are discussed in [217].

1) Latency and Timeliness: Marchetto et al. [221] state that additionally connectivity and isolation between endpoints are network security requirements, although these can possibly be interpreted as security policies by themselves. While they identify these security requirements, their work addresses a slightly different matter: the Virtual Network Embedding problem, which concerns the placement of virtual network functions so that they are optimized and can be verified to correctly enforce the desired security policies. This can potentially be utilized by other works to keep overhead to a minimum and minimize network latency. Hu [218] also identifies latency as a challenge and states that the controllability and configurability of network architectures and applications are key elements in reducing latency. The implied requirement is thus that IIoT environments must be controllable and configurable at every level. This network latency issue is also relevant to security monitoring (previously discussed in Section V-G), as keeping latency to a minimum is a large issue in network monitoring services, and possible solutions include alteration of the network architecture [184].

For time-critical applications, there exist specialized standards such as the Time Sensitive Networking (TSN) standards to provide deterministic and timely networking capabilities between systems. These applications often require remote access to sensors, actuators, and Programmable Logic Controller (PLC)s driving industrial devices. These connections must fulfill the same requirements as when those devices would be directly connected on the machine level [219]. For
this, safety and security measures must be present in the network architecture to correctly prioritize such traffic.

With a gradual movement towards an IIoT enabled industrial process, it is expected that many legacy devices will remain operational for some time in parallel with new technologies, in a sunset phase. These legacy devices must thus be isolated from the Internet, but care must be taken in the isolation technologies, as to not provide too much overhead in time-critical processes. To that end, Lackorzynski et al. [220] compare multiple readily available VPN solutions on metrics important to industrial appliances.

2) Availability: Latency is not the only issue. From a dependability perspective, single points of failure should be eliminated. However, with modern Cloud infrastructures, often the network virtualization solutions proposed by Cloud providers constrain customers to that one Cloud service provider [222]. To allow critical applications to utilize the Cloud for enhanced functionality, without sacrificing availability, the authors propose a platform to allow virtualized networks spanning multiple Cloud providers as well as private networks, while also solving the Virtual Network Embedding problem. This way, they are able to explore the flexibility of combining on-premises systems with Cloud systems, and satisfy privacy requirements by creating security policies limiting the mapping of sensitive NFV applications to specific classes of networks.

3) Wireless: Many smart devices make use of wireless technologies for data transmission. These wireless communication standards work on a lower level than the data transport technologies discussed in Section V-F1. However, the security requirements for wireless transmission that we found in the investigated literature largely overlap with those of data transport security. A common type of wireless communication technologies aimed at long-range low-power IoT devices are LPWAN technologies [229].

Chen et al. [223] list a number of security requirements in a review of the Narrow Band IoT (NB-IoT) standard. This standard was developed by the 3rd Generation Partnership Project (3GPP) [252] and focuses on extremely low-power devices and indoor connections. The authors identify DoS as a much more apparent threat than in traditional networks, as low-power mobile devices will be easily drained from battery power. Another requirement is to prevent eavesdropping of transmissions, as information leakage can lead to devastating results. The authors also identify the need for devices to sign and encrypt their transmissions, in order to mitigate the potential impact of a compromised base station (they identify this as more likely than with traditional wireless technologies). Mutual authentication between devices and the base station is also mentioned, in order to prevent spoofing attacks. Recently, an exploratory investigation has shown that properties derived from the relative distance and direction between transmitters can help in identifying these types of attacks [32]. As the NB-IoT standard supports a large number of devices (100,000) being connected to one terminal, it is challenging to create sufficiently lightweight and efficient authentication and access control mechanisms for these.

Kail et al. [33] compare the security properties of several LPWAN technologies in the unlicensed bands. This comparison is done through the inspection of a number of capabilities that are to be expected of a secure standard, and therefore we consider them as sensible security requirements for wireless technologies: authentication, message integrity, confidentiality, Over-the-Air firmware upgrade capabilities, reliable communication, and key exchange capabilities. Note that these requirements are also already covered in other sections, so we do not list them in Table XI. Additionally, they identify the need for protection against common attacks against wireless technologies, such as wide-band jamming, selective jamming, eavesdropping, traffic analysis, replay attacks, and wormhole attacks. Their conclusion is that further research on security and privacy-related features for low-power wireless communication standards is needed. Wang et al. [226] argue that in order to satisfy the confidentiality requirement, encryption techniques are not sufficient, and propose a friendly jamming scheme, making it harder for eavesdroppers to distinguish communication from noise.

6TiSCH [228] is a standardization effort by the Internet Engineering Task Force (IETF), aimed at low-power deterministic IPv6 communication for WSN technologies and industrial IIoT networks, by building on the IEEE 802.15.4 standard for low-rate Wireless Personal Area Network (WPAN) devices, and supports a different category of devices than LPWAN technologies. Although timeliness is one of the major goals of 6TiSCH, it also aims to incorporate a variety of security properties. For example, the authors state that support for Datagram Transport Layer Security (DTLS) and TLS is taken into consideration. A further discussion of 6TiSCH security is given in [227]. Related to WPAN technology, Ulz et al. [225] propose a secure communication framework utilizing NFC, aimed at providing a reliable solution for mobile robots that need to communicate with machines. Due to the short-range nature,

<table>
<thead>
<tr>
<th>ID</th>
<th>Security requirement</th>
<th>Related sources</th>
<th>Relative interest</th>
<th>% within category</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS-01</td>
<td>dynamicity of configuration</td>
<td>[212], [218], [219]</td>
<td>Medium</td>
<td>10%</td>
</tr>
<tr>
<td>NS-02</td>
<td>security policy enforcement</td>
<td>[221]</td>
<td>Low</td>
<td>5%</td>
</tr>
<tr>
<td>NS-03</td>
<td>management overhead minimization</td>
<td>[212], [224]</td>
<td>Low</td>
<td>7%</td>
</tr>
<tr>
<td>NS-04</td>
<td>network isolation</td>
<td>[129], [210], [211], [217], [220], [221], [224]</td>
<td>Medium</td>
<td>24%</td>
</tr>
<tr>
<td>NS-05</td>
<td>timeliness</td>
<td>[187], [213], [218]</td>
<td>Medium</td>
<td>1%</td>
</tr>
<tr>
<td>NS-06</td>
<td>availability (DoS, jamming, etc.)</td>
<td>[187], [214], [219], [221], [222], [226]</td>
<td>Medium</td>
<td>21%</td>
</tr>
<tr>
<td>NS-07</td>
<td>wireless transmission security</td>
<td>[32], [33], [209], [215], [216], [225]</td>
<td>High</td>
<td>34%</td>
</tr>
</tbody>
</table>
this naturally helps remedy eavesdropping and interference issues.

The 5G standard also addresses IoT scenarios, and provides support for virtualization of network resources. This enables the creation of isolated network partitions with different demand profiles. Two key scenarios that 5G targets are massively deployed low-bandwidth IoT devices, and critical latency-sensitive applications. Both of these map very well to common IoT and Industry 4.0 scenarios. Kurtz et al. [224] elaborate on network slicing, and how it can be realized through use of SDN and NFV technologies. The security requirements identified in their work concern strict isolation of network traffic, and the ability to provide hard service guarantees, such as on latency, data rate, and reliability. Additionally, they mention the need for manageability in this environment, as misconfiguration of systems can have a negative impact on the capabilities of the overall network. Their results show that 5G technologies can be used for real-time, critical applications.

I. Models and Methodologies

In this subsection, we discuss proposed security models and methodologies in the investigated literature. As the security issues that these address are relatively high-level, the security requirements are relatively abstract and encompass multiple aspects of IIoT systems. Therefore, the security requirements listed in table XII are to be interpreted as recommendations and tools to improve the degree to which other security requirements can be satisfied, as well as easing the process of doing so.

Shaabany and Anderl [51] states that software and hardware should be designed carefully, with security in mind, in order to reduce the attack surface as much as possible during design time. Among some less-security related requirements, they argue that specialized functions should be standardized for reuse as much as possible (across manufacturers as well), that all components should be uniquely identifiable and that this identifier should be used in communication with other components, and that security guarantees should be given on every hierarchical level. To aid in addressing these needs, the authors propose a security-by-design approach encompassing both hardware and software. It is thus clear that security should be considered at every step of the development lifecycle of a system, and in [234], Eckhart et al. propose 14 security activities spread across multiple phases in the development process that have shown to be effective for cyber-physical systems. Maksut et al. [42] take a more flexible stance than Shaabany, observing that security and business process performance will always come at the cost of each other. They state that one possible solution is to create a self-adapting system that can flexibly provide end-to-end security. To this end, they propose the investigation of self-adapting models and describe a relevant meta-model. As an example, they suggest that TLS sessions can be re-used for intermittent communication in situations where the threat is deemed to be low, but the rate at which they should be renewed can be dynamically scaled up and down to accommodate for differences in threat levels. Another security-by-design approach recommends the usage of security control assignment matrices to determine the types of security controls that should be present in various parts of a system [132].

It is often easier and more effective to create more specific architectural frameworks rather than generic ones, and the investigated literature contains specialized models and methods for various scenarios. The security-by-design approach in [165] specializes in actuating and sensing scenarios, while in [231], the authors introduce an integrated model aimed specifically at mobile e-health applications. Their approach also considers security issues at design time and can be integrated into more generic architectures. Craggs et al. [233] target research scenarios, and describe a reference architecture for research testbeds, making the accurate observation that real IIoT scenarios are likely to have a mixture of legacy and new technologies and that security solutions should account for this. In [237], a method for arriving at a security capability-model for IIoT supply-chains is described, as the authors identified that businesses generally lack insight in their own supply chains, which is a security liability. In a comprehensive work, McGinthy and Michaels [242] describe secure architectural frameworks for IIoT and WSN sensor nodes, with security features grouped by energy class. They address many security requirements that should be satisfied for these classes, including data confidentiality, attestable boot procedures, and key management. Bécue et al. [23] state that it is necessary to improve the prevention, detection, investigation, and response to adversarial machine learning attacks on AI-powered modules. At the same time, humans and machines should aid in the surveillance of each other; if a human behaves anomalously, machines should be able to detect and report this, and vice versa. They propose using a “cyber-range” approach where digital twins of physical devices are modeled by a team of engineers using feedback from the operators, as well as common design techniques such as risk assessments. These digital twins are then used to simulate more optimized usage scenarios, and red/blue teams perform attack and response scenarios, that help the digital twin learn about how to protect and respond to attacks by itself. Once a digital twin is deemed sufficiently secure it can be used in production settings. This approach requires decisions that steer towards such a model early on in the architectural design process. This is also necessary for the model described by Condry and Nielson [26]. In this model, the authors leverage capabilities of gateways between control systems and the Internet to allow for direct communication between control systems and client devices. Kondeva et al. [240] observe that the fields of safety and security engineering are closely related but have their own techniques and methods. They consider that safety and security requirements should not clash with each other and that these should be integrated more tightly. To this end, they introduce a method to generate attack trees from fault tree analysis.

Risk assessment for the IIoT is another field that has seen activity in recent years. In [44] and [126] two risk assessment models for the IIoT are presented. The first is mainly focused on water sewage systems, but has aspects that can be generalized, while the second aims to be general, and utilizes use cases as its input. The authors of [44] state that it
is not possible to protect against threats without a proper risk assessment. The reason that traditional risk assessment methods are not adequate is due to the complexity of integrating all the aspects of an IIoT system, and due to the increased impact factor in IoT environments because of the increased amount of physical assets and ways it can affect human lives. To this end, they propose a 10-phase comprehensive risk assessment method, that is able to capture many relevant aspects. Mouratidis and Diamantopoulou [43] take things even further by proposing a more formal security analysis method for the IIoT. In their method they build on the Secure Tropos language to allow for precise modeling of industrial environments, their security constraints, and relevant threats. They then use graph analysis to trace possible attack paths and identify which devices should satisfy certain security requirements. A more manual approach is taken by Boyes et al. [232]. They propose a multidimensional categorization framework, that can help with a better analysis of threats, aside from being useful as a more general categorization framework. They envision that a proper categorization of devices will help with identifying similar threats across different aspects of the IIoT domain.

As resource constraints are often a bottleneck for IIoT systems, it is perhaps surprising that there has not been a lot of work on modeling the overhead these bring. The only such work that was found in the literature is by Ivkic et al. [238], and describes an onion layer model that enables one to sum all overhead introduced by security functions.

### J. Summary and Discussion

In our survey of the literature on security in the IIoT domain, we have extracted 49 security requirements covered by the investigated works, spread across 8 categories: Authentication, Access Control, Maintainability, Resilience, Data security and data sharing, Security Monitoring, Network Security, and Models and Methodologies. Additionally, we have made an effort to summarize the literature in our discussion.

In this subsection, we summarize the findings discussed in this section in two ways. Firstly, in Table XIII, we lay out the number of works per category, providing a measure of the distribution of the papers across categories. As detailed in Section V, the number of papers addressing each category is taken from Figure 2 as the number of papers appearing in the corresponding level 1 (i.e., subsection) and all level 2 (i.e., subsections) nodes, but removing duplicates. Secondly, we summarize all the identified research requirements in Table XIV, listed in reverse order by their popularity based on the total number of investigated works. Note that the overall interest for this table is computed based on the total number of works covered in this survey, and is thus different from earlier tables in this section where it was computed based on the numbers within each category. Table XV lists these new thresholds.

A few observations can be made when looking at the popularity of the categories, which are laid out in Table XIII. Firstly, research interest in Authentication, together with Data Security and Data Sharing appears significantly higher than the other categories. This is interesting because these intuitively also have the most in common with standard IoT scenarios. At the same time, the very IIoT-centered categories of Maintainability and Resilience are some of the least active. We believe that this exposes a promising area for new research.

Access Control has seemingly been of the least interest, perhaps because many of its security requirements and works are already implicitly treated in the Authentication section, and various works present frameworks that provide both, but are discussed in the Authentication category.

Security Monitoring is also fairly popular, with 41 works discussing it in various ways. What stands out about this category is that considering its popularity, there are relatively few (4) different requirements covered in the literature. This stands out even more when looking at Table XIV, where requirement SM-01 is the most popular of all. Further, both SM-01 and DSS-05 have seen significantly more interest than any other requirement. This is perhaps because these requirements are the most open-ended out of all identified requirements, thereby collecting a large variety of works that discuss them.

Finally, the observant eye might notice that in Table XIV the percentages sum up to 92.6%. This is because, throughout the study, roughly 7.4% of the investigated works identify some categories as requirements, meaning they have been included in this work, but do not identify any of the specific security requirements. Therefore, they are included in the category count, but not in the requirement count.

---

**TABLE XII**

<table>
<thead>
<tr>
<th>ID</th>
<th>Security requirement</th>
<th>Related sources</th>
<th>Relative interest</th>
<th>% within category</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM-01</td>
<td>adequate risk/threat assessment</td>
<td>[43], [44], [83], [100], [126], [232], [235], [236], [241], [244]</td>
<td>High</td>
<td>33%</td>
</tr>
<tr>
<td>MM-02</td>
<td>minimization of overall attack surface</td>
<td>[51]</td>
<td>Low</td>
<td>3%</td>
</tr>
<tr>
<td>MM-03</td>
<td>security by design</td>
<td>[23], [26], [41], [42], [51], [132], [230], [231], [234], [240], [242], [243], [245]</td>
<td>High</td>
<td>43%</td>
</tr>
</tbody>
</table>

**TABLE XIII**

<table>
<thead>
<tr>
<th>ID</th>
<th>Category</th>
<th>Papers (N’ * )</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Authentication</td>
<td>77</td>
<td>27%</td>
</tr>
<tr>
<td>AC</td>
<td>Access Control</td>
<td>16</td>
<td>6%</td>
</tr>
<tr>
<td>M</td>
<td>Maintainability</td>
<td>15</td>
<td>5%</td>
</tr>
<tr>
<td>R</td>
<td>Resilience</td>
<td>16</td>
<td>6%</td>
</tr>
<tr>
<td>DSS</td>
<td>Data Security and Data Sharing</td>
<td>58</td>
<td>20%</td>
</tr>
<tr>
<td>SM</td>
<td>Security Monitoring</td>
<td>42</td>
<td>15%</td>
</tr>
<tr>
<td>NS</td>
<td>Network Security</td>
<td>29</td>
<td>10%</td>
</tr>
<tr>
<td>MM</td>
<td>Models and Methodologies</td>
<td>30</td>
<td>11%</td>
</tr>
</tbody>
</table>
VI. QUANTITATIVE RESULTS

In this section, we provide a quantitative analysis of the set of studies resulting from the presented research.

In particular, we address research questions (RQ2)-(RQ4) by analyzing the number of publications related to IIoT security over the years, the geographical distribution of these studies, and the favorite publication venues.

A. Spread of Publications Throughout the Years (RQ2)

Figure 4 shows the number of publications between 2011 and 2019. Security research for the IIoT starts first appearing around 2011, being initially dormant but slowly growing from 2013 onward. In 2017, a drastic increase in activity can be seen. While it is tempting to attribute this growth to the fact that 2016 saw several serious IoT and...

<table>
<thead>
<tr>
<th>Overall interest</th>
<th>ID</th>
<th>Security Requirement</th>
<th>Category</th>
<th>Overall %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very High</td>
<td>SM-01</td>
<td>infrastructure monitoring</td>
<td>Security Monitoring</td>
<td>9.5%</td>
</tr>
<tr>
<td></td>
<td>DSS-05</td>
<td>secure external data storage</td>
<td>Data Security and Data Sharing</td>
<td>7.1%</td>
</tr>
<tr>
<td></td>
<td>A-06</td>
<td>mutual authentication</td>
<td>Authentication</td>
<td>4.6%</td>
</tr>
<tr>
<td></td>
<td>MM-03</td>
<td>security by design</td>
<td>Models and Methodologies</td>
<td>4.6%</td>
</tr>
<tr>
<td>High</td>
<td>DSS-02</td>
<td>data confidentiality</td>
<td>Data Security and Data Sharing</td>
<td>3.9%</td>
</tr>
<tr>
<td></td>
<td>A-02</td>
<td>key distribution</td>
<td>Authentication</td>
<td>3.5%</td>
</tr>
<tr>
<td></td>
<td>SM-02</td>
<td>threat response</td>
<td>Security Monitoring</td>
<td>3.5%</td>
</tr>
<tr>
<td></td>
<td>NS-07</td>
<td>wireless transmission security</td>
<td>Network Security</td>
<td>3.5%</td>
</tr>
<tr>
<td></td>
<td>MM-01</td>
<td>adequate risk/threat assessment</td>
<td>Network Security</td>
<td>3.5%</td>
</tr>
<tr>
<td></td>
<td>A-08</td>
<td>minimization of user interaction</td>
<td>Authentication</td>
<td>2.8%</td>
</tr>
<tr>
<td></td>
<td>AC-04</td>
<td>decentralized AC</td>
<td>Access Control</td>
<td>2.8%</td>
</tr>
<tr>
<td>Medium</td>
<td>A-01</td>
<td>multi-factor authentication</td>
<td>Authentication</td>
<td>2.5%</td>
</tr>
<tr>
<td></td>
<td>NS-04</td>
<td>network isolation</td>
<td>Network Security</td>
<td>2.5%</td>
</tr>
<tr>
<td></td>
<td>A-07</td>
<td>privacy-preserving authentication</td>
<td>Authentication</td>
<td>2.1%</td>
</tr>
<tr>
<td></td>
<td>NS-05</td>
<td>timeliness</td>
<td>Network Security</td>
<td>2.1%</td>
</tr>
<tr>
<td></td>
<td>NS-06</td>
<td>availability (DoS, jamming, etc.)</td>
<td>Network Security</td>
<td>2.1%</td>
</tr>
<tr>
<td></td>
<td>A-03</td>
<td>node addition, revocation, revoking</td>
<td>Authentication</td>
<td>1.8%</td>
</tr>
<tr>
<td></td>
<td>A-04</td>
<td>decentralized key management</td>
<td>Authentication</td>
<td>1.8%</td>
</tr>
<tr>
<td></td>
<td>AC-02</td>
<td>fine-grained AC</td>
<td>Access Control</td>
<td>1.8%</td>
</tr>
<tr>
<td></td>
<td>R-01</td>
<td>continuation of operation with compromised subsystems</td>
<td>Resilience</td>
<td>1.8%</td>
</tr>
<tr>
<td></td>
<td>R-03</td>
<td>standards compliance</td>
<td>Resilience</td>
<td>1.8%</td>
</tr>
<tr>
<td></td>
<td>A-10</td>
<td>attestation</td>
<td>Authentication</td>
<td>1.4%</td>
</tr>
<tr>
<td></td>
<td>AC-01</td>
<td>handle dynamic changes</td>
<td>Access Control</td>
<td>1.4%</td>
</tr>
<tr>
<td></td>
<td>M-01</td>
<td>software updateability</td>
<td>Maintainability</td>
<td>1.4%</td>
</tr>
<tr>
<td></td>
<td>M-08</td>
<td>secure status transfer</td>
<td>Maintainability</td>
<td>1.4%</td>
</tr>
<tr>
<td></td>
<td>DSS-04</td>
<td>secure data transport</td>
<td>Data Security and Data Sharing</td>
<td>1.4%</td>
</tr>
<tr>
<td></td>
<td>A-09</td>
<td>non-reputation</td>
<td>Authentication</td>
<td>1.1%</td>
</tr>
<tr>
<td></td>
<td>AC-06</td>
<td>transparency</td>
<td>Access Control</td>
<td>1.1%</td>
</tr>
<tr>
<td></td>
<td>M-02</td>
<td>configuration updateability</td>
<td>Maintainability</td>
<td>1.1%</td>
</tr>
<tr>
<td></td>
<td>M-03</td>
<td>disturbance-free updates</td>
<td>Maintainability</td>
<td>1.1%</td>
</tr>
<tr>
<td></td>
<td>DSS-06</td>
<td>data flow control</td>
<td>Data Security and Data Sharing</td>
<td>1.1%</td>
</tr>
<tr>
<td></td>
<td>DSS-07</td>
<td>data protection legislation compliance</td>
<td>Data Security and Data Sharing</td>
<td>1.1%</td>
</tr>
<tr>
<td></td>
<td>SM-04</td>
<td>security policy enforcement</td>
<td>Security Monitoring</td>
<td>1.1%</td>
</tr>
<tr>
<td></td>
<td>NS-01</td>
<td>dynamicity of configuration</td>
<td>Network Security</td>
<td>1.1%</td>
</tr>
<tr>
<td>Low</td>
<td>AC-03</td>
<td>centralized AC</td>
<td>Access Control</td>
<td>0.7%</td>
</tr>
<tr>
<td></td>
<td>AC-05</td>
<td>privacy-preserving AC</td>
<td>Access Control</td>
<td>0.7%</td>
</tr>
<tr>
<td></td>
<td>M-05</td>
<td>traceability</td>
<td>Maintainability</td>
<td>0.7%</td>
</tr>
<tr>
<td></td>
<td>M-06</td>
<td>compatibility</td>
<td>Maintainability</td>
<td>0.7%</td>
</tr>
<tr>
<td></td>
<td>R-02</td>
<td>operation with intermittent connectivity</td>
<td>Resilience</td>
<td>0.7%</td>
</tr>
<tr>
<td></td>
<td>NS-03</td>
<td>management overhead minimization</td>
<td>Network Security</td>
<td>0.7%</td>
</tr>
<tr>
<td></td>
<td>A-05</td>
<td>transitive authentication</td>
<td>Authentication</td>
<td>0.3%</td>
</tr>
<tr>
<td></td>
<td>AC-07</td>
<td>compatibility</td>
<td>Access Control</td>
<td>0.3%</td>
</tr>
<tr>
<td></td>
<td>M-04</td>
<td>usability of update process</td>
<td>Maintainability</td>
<td>0.3%</td>
</tr>
<tr>
<td></td>
<td>M-07</td>
<td>transparency</td>
<td>Maintainability</td>
<td>0.3%</td>
</tr>
<tr>
<td></td>
<td>DSS-01</td>
<td>data loss mitigation</td>
<td>Data Security and Data Sharing</td>
<td>0.3%</td>
</tr>
<tr>
<td></td>
<td>DSS-03</td>
<td>standardization</td>
<td>Data Security and Data Sharing</td>
<td>0.3%</td>
</tr>
<tr>
<td></td>
<td>SM-03</td>
<td>handle heterogeneous sources</td>
<td>Security Monitoring</td>
<td>0.3%</td>
</tr>
<tr>
<td></td>
<td>NS-02</td>
<td>security policy enforcement</td>
<td>Network Security</td>
<td>0.3%</td>
</tr>
<tr>
<td></td>
<td>MM-02</td>
<td>minimization of overall attack surface</td>
<td>Models and Methodologies</td>
<td>0.3%</td>
</tr>
</tbody>
</table>

TABLE XV

<table>
<thead>
<tr>
<th>Interest level</th>
<th>Weighted interest</th>
<th>Range (x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0% ≤ x ≤ 1.0%</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>1.0% &lt; x ≤ 2.3%</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>2.3% &lt; x ≤ 4.0%</td>
<td></td>
</tr>
<tr>
<td>Very High</td>
<td>4.0% &lt; x ≤ 100%</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4. Number of publications per year.
industry related security incidents (such as Mirai [4] and Crashoverride/Industroyer [6]), which served to illustrate the importance of security on these devices, it should be noted that this is in line with the overall growth of IoT as a research area. In 2018 and 2019, the growth in activity continued, showing that the research community deems IIoT security to be of high importance.

B. Geographical Distribution of IIoT Security Research (RQ3)

The geographical distribution of research activity is shown in Figure 5. The data for this was obtained by extracting the country of affiliation of the first author of the considered studies.

German-speaking countries are strongly represented, making for a total of 22% of contributions. One possible explanation is that one of our search terms, Industry 4.0, was originally coined by the German government [253], thus, it might have seen higher adoption in German-speaking countries. This raises the question of whether our search terms were successful in providing a good global sample of studies in this field. We believe they were, since the field we are considering is very narrow; we specifically searched for Industrial challenges in order to be able to extract security requirements unique to this field. Furthermore, we have conducted reverse snowball sampling to ensure a fair research scope.

China and the United States of America are the two other major contributors. This can be attributed to the size of their industries and thus the relevance of research in this area. However, interestingly, 54% of the studies originate from Europe, showing that this topic is also regarded as highly relevant in countries with smaller industries.

The ‘others’ group consists of the 23 countries that have 3 or fewer publications in this field: Algeria, Belgium, Brazil, Czech Republic, Finland, Greece, Hungary, Iran, Ireland, Japan, Malaysia, Morocco, New Zealand, Norway, Pakistan, Qatar, Romania, Russia, Saudi Arabia, Serbia, Taiwan, Turkey, Ukraine.

C. Venue Types for Publication (RQ4)

We have grouped the studies based on the venue type of their publication, which is shown in Figure 6. As can be seen, conference proceedings are the most popular dissemination method, followed by journals. The ‘others’ category consists of venue types in which 4 or fewer publications were published: congresses, summits, and forums.

Looking at the specific venues of publication (Figure 7), it can be seen that the IEEE Transactions on Industrial Informatics journal is by far the most popular venue, with 25 publications. One noteworthy observation here is that, out of all considered studies, only 16 were published in venues focused on security. The vast majority of IIoT security-related works appears to be published in venues targeting industrial systems or IoT instead.

VII. OPPORTUNITIES ENABLED BY FOG COMPUTING

In Section V, we have extracted security requirements for the IIoT from the investigated literature and discussed a number of challenges that stand in the way of the adoption of conventional solutions to address these requirements. In this section, we reflect on the challenges and discuss how Fog computing shows promise as a remedy to a number of those.

It is important to note that Fog computing is a relatively new paradigm the exact definition of which is still being debated.
in the scientific community and often intersects with similar paradigms, such as Edge computing, Mobile Edge computing, and Mobile Cloud computing. To maintain consistency with earlier work, we use the definition of Fog computing as used in [254]; a paradigm that extends the Cloud and integrates Edge and IoT, while providing a new, horizontally scalable highly virtualized layer that distributes computing, storage, control, and networking capabilities across the Cloud-To-Things spectrum [8]. For a more detailed treatise on the differences between Fog, Edge, and other paradigms we refer the interested reader to [254].

Also, we are aware that a comprehensive and thorough discussion on how Fog computing could tackle the IIoT security requirements would require a dedicated treatment that would result in an entire paper itself, which is out of the scope of this work (for instance, in [255] we focus on how Cloud requirements can impact IoT). Thus, the aim of this section is to provide food for thought on the topic and a source of inspiration for future research, rather than an exhaustive analysis.

In detail, we first give the definition of Fog computing assumed in this work. Then, we revisit the majority of topics covered in Section V and depicted in Figure 8: authentication, access control, maintainability, resilience, data security and data sharing, security monitoring, and network security. For each of these, we discuss how we envision what Fog-enabled solutions might look like and suggest potential research opportunities, but we leave confirmation of the validity of these ideas as a topic for further research. We close the section with a discussion on limitations and open challenges for Fog computing.

A. Fog Computing

Fog computing is a relatively recent computing paradigm born from the necessity to provide the missing link in the Cloud-to-Thing continuum [8].

According to the IEEE standard 1934-2018 [256], Fog computing is “a horizontal, system-level architecture that distributes computing, storage, control, and networking functions closer to the users along a cloud-to-thing continuum”. Thus, Fog computing can be considered as an extension of Cloud computing that distributes the benefits of the Cloud closer to the IIoT and across multiple layers of the network topology.

Any system that wants to be compliant with the aforementioned definition of Fog computing needs to present the following attributes, also referred to as pillars: security, scalability, openness, autonomy, reliability, availability, serviceability, agility, hierarchy, and programmability. A thorough discussion of these pillars can be found in [8], [256].

In this setting, the fog node is “the physical and logical network element that implements fog computing services” [8]. Since Fog nodes can be placed on-premises, they can be accessed by devices even when the connection to the outside world is failing. This helps us in identifying research opportunities for issues arising from intermittent connectivity. Note that this can be generalized: if there is a connection failure anywhere on the route from the (local) Fog node to the (remote) Cloud, then all Fog nodes that are positioned before the unreachable hop are still reachable and thus able to provide the local system with their services.

B. Fog-Enabled Authentication

When looking at the authentication challenges discussed in Section V-B, it can be observed that there are a number of points where a Fog node can be helpful in addressing them.

A first intuitive way of applying Fog computing to these challenges can be found by considering existing authentication solutions that require third-party servers in their setup or execution, such as [63], [86], [93]. A Fog node fits the requirements for these servers perfectly, as it is not severely restrained by computational or energy resources, is on-premises, and has very low response times. If Fog computing nodes are considered as part of the infrastructure, many of the issues with relying on a third-party server are thus addressed “for free”.

Secondly, Fog nodes can serve to enhance traditional PKI infrastructures, where Fog nodes can act as “certificate authorities” for local devices or help establish a federated and robust
key infrastructure through, e.g., peer-to-peer networking capabilities with other Fog nodes. To our knowledge, no work investigating this currently exists.

In dynamic environments, Fog nodes can potentially help alleviate issues relating to node addition, removal, and rekeying as well. For example, it could serve as a trusted “gateway” to which Edge devices are paired, preventing them from communicating directly with any other system. This is not unlike how Bluetooth devices can be paired with smartphones and other devices. Node addition, removal, and rekeying can then be handled from the Fog node.

As we have seen in Section V-B2, some proposed solutions require biometric features ([72], [77]), smart cards ([72]) or NFC tags ([36], [37], [67], [225]), in the authentication process. Also here there is potential for Fog nodes: not every lightweight system might be equipped with the necessary sensors for this. However, it might be possible to equip Fog nodes with sensors and use them as proxies for sensor readings. This would increase scalability, as a Fog node can be positioned so that it is more easily accessible than the Edge devices connected to it. Thus, if maintenance engineers would want to, e.g., authenticate updates for the devices by using NFC keys or biometrics, they will only need to seek out the Fog node and present the relevant keys to it, as opposed to seeking out every relevant device separately.

Fog nodes might also enable the possibility of bringing TPM and/or TEE capabilities to Edge devices that do not contain these modules themselves. For that to be possible, the Edge devices need to set up a trusted channel between the Fog node’s TPM/TEE module, which could be possible through some form of a key setup protocol that involves a one-time pairing step. Fog nodes could be equipped with multiple TPM or TEE modules to serve more than one Edge device (or itself) at the same time, such as the recently introduced Intel SGX cards [257]. Trusted hardware capabilities in Fog nodes can also be used for attestation purposes in various settings (against remote Fog nodes, against Edge devices, and so on). We expect that there are a lot of fruitful research directions for the combination of Fog nodes and trusted hardware.

**C. Fog-Enabled Access Control**

As with authentication, Fog nodes have the potential for enhancing AC challenges in industrial scenarios.

Firstly, some AC policies could be outsourced from extremely resource-constrained devices to a Fog node (e.g., accessing sensitive files from a central repository), or if the scenario is suited for it, AC can be managed completely by a Fog node. Another identified challenge for AC frameworks is that while managing policies centrally gives more flexibility, it introduces new risks due to the central server now being a single point of failure. Fog nodes could provide a “hybrid” middle ground where AC is federated between various Fog nodes on-premises, and that Edge devices can then query these Fog nodes, thus increasing the overall reliability and scalability. To the best of our knowledge, this is still an open research area.

As mentioned in Section V-C, compatibility with legacy devices is another issue in the IIoT. Fog nodes could act as a bridge between newer devices and legacy devices with poor security, keeping them sufficiently isolated from the wider network and providing security measures where necessary in exposed interfaces, possibly through a ZTN approach.

**D. Fog-Enabled Maintainability**

Fog computing can bring large benefits to the maintainability of industrial systems.

By their very nature, industrial systems are connected to the Internet, and thus enable the possibility of managing software and configuration updates for attached Edge devices. Fog nodes are perfectly situated to verify the validity of such updates and perform in-depth tests such as performing the updates in a sandboxed environment and then observing for anomalies before deploying them on real devices, while at the same time allowing for the application of updates with minimal disturbance to the services themselves. In practice, this would turn the solution proposed by [111] into a Fog application.

Fog nodes also provide an ideal target platform for an “industrial app marketplace” such as proposed in [113]. It is not difficult to envision a system where a Fog node would allow users to view software packages together with their version number and update information for all connected devices, in an ordered and user-friendly way. Moreover, Fog nodes could go further and allow for management of configuration files for connected Edge devices as well. For example, one could think of an application where configuration files are retrieved from a Cloud service, verified by the Fog node and subsequently delivered to specified Edge devices, filling in sensitive information fields as necessary so as to prevent the Cloud from requiring access to this information.

As Fog nodes could provide an easily accessible location for the reading of NFC tags or other hardware authentication modules, one could easily extend maintenance processes with those extra authentication factors without requiring engineers to physically attend to each affected device individually.

The ideas described here are merely speculative, and there is plenty of room for research in any of these areas. We expect a variety of maintainability-enhancing applications of Fog computing will be identified and researched in the future.

**E. Fog-Enabled Resilience**

Fog nodes could act as reactive security agents, isolating or disabling connected devices when they appear compromised. This allows security personnel to then further investigate the issue, while the system itself can continue operations. This is also discussed in [258], where a number of Fog use-cases and research challenges are listed. The authors state that automatic fault detection and reconfiguration is essential, and identify the potential for Fog nodes to do this autonomously, but state that this is a challenging topic that requires addressing. However, a solution to this challenge would enable resiliency as it is defined by the ICS.

A second challenge that can be overcome through Fog computing, is maintaining normal operation through intermittent Internet connectivity. To an extent, a Fog node can take over
processes normally executed in the Cloud. Thus, when the connection to the Cloud fails, the operational capability of Edge devices is not affected. Related to this, some devices continuously or periodically need to transmit data to the cloud, where it can then be processed. If this were done directly, data loss is a risk in case of intermittent connectivity. As an alternative to introducing some data storage capabilities on the Edge devices themselves, a Fog node could collect data from the devices, and forward it to the Cloud. Then, when there is no connection to the Internet, the Fog node can act as a buffer and send the buffered information upwards to the Cloud once the connection is restored. This way, Edge devices do not need to worry about failing Internet connectivity at all.

Finally, Fog nodes and their application-independent software can be developed to satisfy resiliency-related indicator points, which in turn can aid in providing contractual service guarantees as is currently often seen in Cloud service agreements.

F. Fog-Enabled Data Security and Data Sharing

Whenever it is necessary for a device to access sensitive data that should be stored securely, this requires the device to firstly have the storage capacity, and secondly the means to secure this data at rest. For lightweight systems that do not have the capacity to store and secure data securely, Fog nodes can provide a solution; they are not tied to severe resource constraints and can be equipped with ample storage and computational capacity for common encryption methods. Additionally, Fog nodes can be deployed on the local network, meaning data will never have to leave the premises. Even for extremely large amounts of data, Fog nodes could act as middleware between external Cloud storage, and encrypt/decrypt data stored in the Cloud transparently, e.g., using the techniques described in [149], [159]. To the Edge devices, it can be presented as originating from the Fog node, and they do not need to be aware of the underlying storage and security mechanisms.

As Fog nodes can be positioned between Edge devices and external parties as gateways, this also unlocks the opportunity to secure and control data flow to these external parties. A Fog node can set up and maintain highly secure, authenticated channels with remote parties, potentially alleviating some of the challenges involved in designing lightweight Edge devices that need to interact with these parties, as they only need to concern themselves with secure communication with the Fog node. If the Fog node additionally has the ability to access the message content of traffic passing through it, it can enforce data flow policies, e.g., as described in [164], allowing fine-grained data security mechanisms on top of encryption techniques.

In Section V-F4 we stated that the protection of sensitive data is in many cases now a legal requirement in the European Union. Fog nodes present a very natural way of meeting these requirements, as they can store data locally, while at the same time allowing for fine-grained data sharing with third parties, should a user allow this. Moreover, it can become easier to manage user rights such as the right to be forgotten.

G. Fog-Enabled Security Monitoring

Because Fog nodes can take on central positions in Industrial networks, they provide a great platform for security monitoring solutions.

For example, a Fog node could run IDS software to detect anomalies or attack signatures. This also provides an opportunity for the Fog and Cloud to augment each other. Intrusion detection models could be trained in a Cloud environment, while executed on a Fog node, thereby addressing the latency issues normally apparent in Cloud solutions. Examples of this can be found in [115], [193]. Because Fog nodes stand in direct connection to sensor devices, they can also perform simple anomaly detection techniques such as ensuring that sensor values are within a certain value range, without adding overhead to the sensors themselves.

Another use of Fog nodes as a security monitoring tool could be the deployment of an anti-malware for IoT devices that is supported by the Fog infrastructure [259]. Indeed, De Donno et al. [260], [261] propose an anti-malware software for IoT and they discuss how the use of Fog computing helps to solve some of the challenges intrinsic in the deployment.

Fog nodes can also potentially take action based on incoming traffic patterns, enabling the mitigation of DoS attacks aimed at very specific devices, even when those devices are not able to protect themselves against those attacks. This also presents the opportunity for dynamic traffic shaping, and other techniques that might help reduce battery consumption on lightweight IoT devices connected to the Fog node.

H. Fog-Enabled Network Security

Also in network infrastructure, Fog computing can potentially help in overcoming current challenges.

With the rise of SDN and NFV technologies, Fog nodes can possibly play a role as a platform for some of these. For example, they can create isolated network environments between themselves and each connected device.

Fog nodes could also be equipped to handle TSN standards when there is a need for deterministic and timely delivery of network traffic between two connected devices. By moving the management of these interfaces to a Fog node, opportunities are created for easier (remote) management and reconfiguration of time-critical systems, even going so far as to move entire control applications to Fog nodes. As an extreme manifestation of this vision, one could imagine “plug-and-play” industrial hardware that can be connected to a Fog node which will then autonomously configure and use it.

We also see opportunities for Fog nodes to improve the availability of critical services, in two ways. Firstly, Fog nodes could run critical applications in a federated fashion, allowing migration or load balancing of tasks between them. This way, the application only becomes unavailable when all participating Fog nodes fail. Secondly, a Fog node can act as a middleware for a critical service running in the Cloud. By deploying this service on multiple Cloud providers, Edge applications relying on it will not be affected by the outage of any one cloud provider; the Fog node can automatically route requests to the remaining available providers.
Finally, Fog nodes could potentially aid in securing wireless infrastructure, by incorporating wireless technologies in security monitoring solutions. This way, jamming attacks or other anomalies in the wireless spectrum can be detected.

I. Challenges and Limitations

Fog computing is not a panacea capable of filling any Cloud-IoT gap without much issue. The paradigm is very much in its early stages, and deployment so far has been extremely limited. Open challenges include practical federation frameworks, resource offloading, and resilience [262]. While we believe that solutions to these challenges are capable of satisfying the security requirements collected in this work, we acknowledge that every solution comes with its own trade-offs, and a thorough analysis of the benefits and drawbacks of Fog computing can only be done once enough Fog-based systems exist to investigate. Nevertheless, one can attempt to make an analysis based on the current state-of-the-art. Thus, in this section, we briefly discuss what we consider some of the biggest potential drawbacks.

Firstly, Fog systems add extra workload to maintenance personnel, and will likely require special training, making it more costly than the Cloud. Whereas Cloud infrastructure is maintained by a specialized team on the Cloud service provider’s end, the Fog paradigm shifts this responsibility to users of the system. The spread of functionality across the Cloud-to-Things continuum potentially complicates this even more. If a security issue is found in a well-known piece of Fog infrastructural software, it is the responsibility of maintainers at every point in the continuum to update their software, as opposed to having to update just the Cloud infrastructure, which is managed by one entity. If one maintainer of a Fog node fails to do this within an appropriate time-window, this can put all entities making use of that node at risk.

Secondly, incident response might be hampered by the distributed nature of Fog systems. We believe this might manifest itself in multiple ways: necessary security expertise might not be available on-site, and specialized incident response teams will have to be called in from external parties. Further, complex incidents might require cooperation between multiple entities along the continuum for forensic analysis, which might not always be possible or add a lot of overhead.

Finally, compatibility between Fog nodes can potentially be a huge issue. If standards are not well-defined or not followed rigorously, it will be very hard to meet the harsh requirements set by industrial environments with nodes from different providers that cannot interoperate efficiently and accurately. This, in turn, can negatively impact the ability to federate and offload tasks to other nodes in the local network, as well as potentially violate security policies if some nodes in the system are unable to uphold the necessary requirements.

VIII. Conclusion

In this work, we have performed a systematic literature review about security for the IoT.

As in any mapping study, it is challenging to take all studies of the field into account, but it is more important to have a good representation of studies rather than a high number of studies [20]. To achieve a good representation, we have methodologically constructed the search queries and queried multiple literature repositories. After that, we utilized reverse snowball sampling to further increase the quality, and to mitigate any possible selection bias. Our initial search queries resulted in 356 possibly relevant papers, which we brought down to a selection of 218 papers through the use of a systematic approach comprised of several phases. These papers were fully read and analyzed for the purposes of this study.

At glance, the work has elaborated around four main research questions: (RQ1) what security requirements exist for the IoT, (RQ2) how scientific publications about IoT security are spread during the years, (RQ3) how IoT security research activity is geographically distributed, and (RQ4) what publication venues are the most popular for IoT security.

First, we have answered question RQ1 by extracting security requirements for the IoT from the investigated works and exploring them, along with the related challenges that make these requirements hard to meet with existing solutions and a measure of their interest in the research community. Then, we have addressed questions (RQ2)-(RQ4) by providing a quantitative analysis of the investigated IoT security research. Finally, we provided a discussion on how Fog computing can play a role in meeting the requirements posed by industrial environments, by taking a Fog computing perspective and revisiting the requirements that were extracted during our investigation, as well as pointing out what limitation and challenges still need to be faced to achieve massive Fog computing deployment.

This work identifies an abundance of research opportunities in the IoT security area and shows that Fog computing, as a rising computing paradigm, can become a powerful tool in securing a variety of connected industrial environments, once its limitations and challenges are overcome.

REFERENCES


U. Hunkeler, H. L. Truong, and A. Stanford-Clark, “MQTT-S—A pub-
O. Bergmann, S. Gerdes, and C. Bormann, “Simple keys for sim-
P. Autenrieth, C. Lörcher, C. Pfeiffer, T. Winkens, and L. Martin,
F. Al-Turjman and S. Alturjman, “Context-sensitive access in industrial
D. Wang and P. Wang, “Understanding security failures of two-
R. Ankele, S. Marksteiner, K. Nahrgang, and H. Vallant, “Requirements


E. Zugasti, M. Iturbe, I. Garitano, and U. Zurutuza, “Null is not always
M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan, and R. Jain,
T. Yu, V. Sekar, S. Seshan, Y. Agarwal, and C. Xu, “Handling a tril-
N. Moustafa, E. Adi, B. Turnbull, and I. You, “Behavior and vulnerabil-
G. Settanni, F. Skopik, A. Karaj, M. Wurzenberger, and R. Fiedler,
S. Tamy, H. Belhadaoui, M. A. Rabbah, N. Rabbah, and M. Rif, “An
evaluation of machine learning algorithms to detect attacks in SCADA
A. Wadsworth, M. I. Thanoon, C. McCurry, and S. Z. Sabatto,
T. Wang, P. Wang, S. Cai, Y. Ma, A. Liu, and M. Xie, “A unified trust-
detection system for edge-based IoT relying on machine-
learning-aided detection,” IEEE Netw., vol. 33, no. 5, pp. 75–81,
S. Yu, V. Sekar, S. Seshan, Y. Agarwal, and C. Xu, “Handling a trillion
unfixable flaws on a billion devices: Rethinking network security for the
L. Zhou, H. Guo, and G. Deng, “A fog computing based approach to
M. Zolanvari, M. A. Teixeira, and R. Jain, “Effect of imbalanced
datasets on security of industrial IoT using machine learning,” in Proc.
M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan, and R. Jain,
“Machine learning-based network vulnerability analysis of indus-
trial Internet of Things,” IEEE Internet Things J., vol. 6, no. 4,
S. Jeong, W. Na, J. Kim, and S. Cho, “Internet of Things for smart manufac-


Nicola Dragoni received the M.Sc. (cum laude) and Ph.D. degrees in computer science from the University of Bologna, Italy. He is a Professor in secure pervasive computing with DTU Compute, Technical University of Denmark, where he also serves as the Deputy Head of the Ph.D. School. He is also part-time Professor in computer engineering with the Centre for Applied Autonomous Sensor Systems, Örebro University, Sweden, and he is affiliated with the Copenhagen Center for Health Technology (CACHET) and the Nordic IoT Hub. He has coauthored over 110 peer-reviewed articles and he has edited three journal special issues and one book. He is active in a number of national and international projects. His main research interests include pervasive computing and cybersecurity, with current focus on Internet-of-Things, fog computing, and mobile systems.

Xenofon Fafoutis (Senior Member, IEEE) received the B.Sc. degree in information technology from the University of Athens, Greece, in 2007, the M.Sc. degree in computer science from the University of Crete, Greece, in 2010, and the Ph.D. degree in embedded systems engineering from the Technical University of Denmark in 2014. From 2014 to 2018, he held various researcher positions with the University of Bristol, U.K., and he was a core member of SPHERE: U.K.’s flagship Interdisciplinary Research Collaboration on Healthcare Technology. He is currently an Associate Professor with the Embedded Systems Engineering Section, Department of Applied Mathematics and Computer Science, Technical University of Denmark. His research interests primarily lie in wireless embedded systems as an enabling technology for digital health, smart cities, and the (Industrial) Internet of Things.