A Study about the Different Categories of IoT in Scientific Publications

Sebastian Fischer
Secure Systems Engineering
Fraunhofer AISEC
Berlin, Germany
e-mail: sebastian.fischer@aisec.fraunhofer.de

Katrin Neubauer
Dept. Computer Science and Mathematics
Ostbayerische Technische Hochschule
Regensburg, Germany
e-mail: katrin.1.neubauer@oth-regensburg.de

Rudolf Hackenberg
Dept. Computer Science and Mathematics
Ostbayerische Technische Hochschule
Regensburg, Germany
e-mail: rudolf.hackenberg@oth-regensburg.de

Abstract—The Internet of Things (IoT) is widely used as a synonym for nearly every connected device. This makes it really difficult to find the right kind of scientific publication for the intended category of IoT. Conferences and other events for IoT are confusing about the target group (consumer, enterprise, industrial, etc.) and standardisation organisations suffer from the same problem. To demonstrate these problems, this paper shows the results of an analyses over IoT publications in different research libraries. The number of results for IoT, consumer, enterprise and industrial search queries were evaluated and a manual study about 100 publications was done. According to the research library or search engine, different results about the distribution of consumer-, enterprise- and industrial- IoT are visible. The comparison with the results of the manual evaluation shows that some search queries do not show all desired publications or that considerably more, unwanted results are returned. Most researchers do not use the keywords right and the exact category of IoT can only be accessed via the abstract. This shows major problems with the use of the term IoT and its minor limitations.

Keywords—Internet of Things; IoT; publications; consumer; industrial; enterprise; categorization.

I. INTRODUCTION

The Internet of Things is defined in ISO/IEC 20924:2018 page 9 as “[...] infrastructure of interconnected entities, people, systems and information resources together with services, which processes and reacts to information from the physical world and virtual world.” [1] This definition is very broad and includes all possible devices that are connected to other devices via a network (not necessarily the Internet), like smartphones, personal computers, connected vehicles, airplanes, smart grid components, smart home devices, connected environment sensors, eHealth hardware, wearables and many more. The ISO/IEC definition is not the only one using this range of devices, also researchers are using IoT to describe all kind of products and prototypes. This leads to difficult situations where conferences or other events focus on IoT and the attendees do not know if the presentations are in their field of interest.

Searching for IoT scientific publications can be difficult as well. With only IoT, a too wide range of topics are returned. Restrictions, such as “consumer” or “enterprise” can help, but a lot of researchers do not use it. For example, the publication “Smart Charger Based on IoT Concept” [2] is about a consumer product, but the title and the keywords (Smart Charger, Arduino, Phone Charger, Battery Charger) are only containing “IoT” and “Smart Charger”. A search for “IoT” and “consumer” will not include the publication.

In this study, we want to show the different problems of IoT as a general term. We start with some related work in Section II and the first part of research (Section III) consists of the different numbers of IoT publications in selected research libraries. The second part (Section IV) shows the results of a manual review of 100 publications according to their IoT category. In Section V, the results were then compared and evaluated to show the problems with the term IoT in research. At the end, a short conclusion and our future work are given in Section VII.

II. RELATED WORK

There is no recent study about current research on IoT publications, which includes the different categories “consumer”, “industrial” and “enterprise”. Some publications, like a study from Mishra et.al. [3] are covering the years from 2000 to 2015 or another study about the IoT trends reaches from 1992 to 2015 [4].

Some newer bibliometric studies from 2019 and 2020 are restricted to Blockchain [5] or Industrial 4.0 [6]. They are both showing the increasing amount of IoT publications, but no current overview of the whole situation of the last two years.

This study was inspired by the approach of the publications mentioned above, although the focus is different. The used academic search libraries differ in many point. For example, the target group and the type of search are different. IEEE Xplore targets technical publications, while Google Scholar and Semantic Scholar are universal. A 2018 paper examined the sizes of different libraries and identified Google Scholar as the largest [7]. Semantic Scholar, on the other hand, uses an algorithm that is based on artificial intelligence and is therefore supposed to provide very precise results [8]. In the course of this paper, the differences with respect to IoT will become clear again.

III. IOT PUBLICATIONS IN RESEARCH LIBRARIES

The aim of this study is to find out whether it is possible to find publications on specific areas of IoT without getting too many results and limit the great diversity of IoT, but also
without overlooking relevant publications. For this goal, we started with “IoT” as a search query in our manual study (Section IV) and after analysing the publications, we came up with three categories “industrial”, “consumer”, “enterprise”, as most of the devices can be classified into these (see Table III).

To find research about used encryption methods in consumer IoT devices, for example, the first search approach would be “consumer IoT encryption”. However, some researchers are not restrict their publications about encryption and just use the term IoT. The previous query will not find this work. If we just use “IoT encryption”, there are too many results (compared to the restricted). Research about encryption in vehicles, industrial environment, etc. are included as well.

To prove this statement we started with different research libraries and different queries and collected the numbers of results.

Overall, six libraries / search engines were used:
- Springer Link
- IEEE Xplore
- ScienceDirect
- ACM digital library
- Google Scholar
- Semantic Scholar

These libraries / search engines are the most common ones and widely used in computer science. Because of their different search algorithms (as seen in the results), data from all of them are shown. For example, IEEE Xplore finds a lot of results for “IoT” alone, but not much with “IoT” and other words combined. The words are all combined the same way over all search engines with the “AND” operator to find only publications with both words in it (e.g., “IoT AND consumer”).

The search was done with some word combinations to investigate the different areas of IoT. However, only a few words yielded many results. A precise search for a specific area is thus very well possible (e.g., automotive), as can be seen in Table I. However, the abbreviation IoT for industrial IoT is not very common. All the results in this paper are only with new publications from the years 2019 and 2020, to show a current overview of the research in the field of IoT.

To get a better separation, for example of the whole 44,800 IoT results of Google Scholar, we used the three search terms in addition to “IoT”: “industrial”, “consumer” and “enterprise”. The results are shown in Table II. In our example from Google Scholar, we get about 45.5 % for “industrial”, 38.2 % for “consumer” and 31.7 % for “enterprise”. The sum is over 100 % because some of the publications can include more than one of them. This shows (in the case of Google Scholar) a good idea of how to find the right IoT category for a research (see Table II).

<table>
<thead>
<tr>
<th>Search term</th>
<th>Springer Link</th>
<th>IEEE Xplore</th>
<th>ScienceDirect</th>
<th>ACM digital library</th>
<th>Google Scholar</th>
<th>Semantic Scholar</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOT</td>
<td>16,545</td>
<td>10,996</td>
<td>7,203</td>
<td>3,027</td>
<td>44,800</td>
<td>56,000</td>
</tr>
<tr>
<td>IOT</td>
<td>529</td>
<td>398</td>
<td>359</td>
<td>74</td>
<td>4,730</td>
<td>2,230</td>
</tr>
<tr>
<td>smart home IOT</td>
<td>4,096</td>
<td>615</td>
<td>1,954</td>
<td>814</td>
<td>20,000</td>
<td>11,500</td>
</tr>
<tr>
<td>automotive IOT</td>
<td>1,277</td>
<td>117</td>
<td>639</td>
<td>155</td>
<td>8,270</td>
<td>2,830</td>
</tr>
</tbody>
</table>

IV. IoT PUBLICATIONS STUDY

Because of the big differences in the search results and therefore in the search type, we made a manual study with 100 publications about their category of IoT. We want to know exactly, which publication belongs to industrial, consumer, enterprise or is not related to IoT at all. For this study, we needed 100 full publications most random as possible. Because we do not know the algorithms behind the different search engines, we decided to use Semantic Scholar with the option “has PDF”. This adds a bit randomness and makes it easier to get the full text. All the search parameters are:

- Keyword: IoT
- Language: English
- Publication date: 2019 and 2020
- Option: “has PDF”
- Sort by Relevance

This search leads to 11,800 results. We downloaded the first 100 publications [2], [9]–[107] and determined the categories. For a better evaluation of the results, it was also noted whether the category of the IoT devices in the publication can already be identified in the title, the abstract or only in the text. Additionally, it was evaluated whether the category can already be extracted from the keywords.

Table III shows the result of the manual review. First, the total number of publications. Not specified publications are
referring to IoT in general. For example, the publication about “Security on IoT Devices with Secure Elements” [30] can be applied to consumer, enterprise and industrial IoT devices. The category “consumer” consists of devices, which are meant to be used by consumers, not professional people. “Enterprise” describes the category for devices used by companies or installed / assembled by a professional service. The last category “industrial” are IoT devices for production. Overall, the different areas for each category were assigned as follows:

**Consumer**
- Smart Home devices
- Wearables
- Connected home automation and alarm systems

**Enterprise**
- Smart city devices
- Environment sensors (for big buildings or fields)
- Medical devices
- Vehicles (transportation)
- Sensors for bigger buildings
- Alarm systems (for business)

**Industrial**
- Machine sensors
- Machine control systems
- Industrial sensors
- Industrial devices with network connection

The lists above are not exhaustive. Medical and transportation devices can be used by consumer, but they have to be installed by a professional. Therefore, they are assigned to enterprise.

The remaining columns in Table III are showing the difficulty of assigning the publications to the categories. If the category can be be determined by the title, the publication is added to column t. If it is only in part possible, it is added to column (t). For example, the title “IoT based home automation using Raspberry Pi” [23] is clearly for consumer, because home automation is one of the consumer parts. In this case the publications is added to column t. Another title “IoT-Enabled Door Lock System” [28] is not clear, because a door lock system can be for the smart home market or just for business buildings. In this case the publication is added to column (t) as the product is in the title, but the main category can only be recognized in the abstract. Therefore, the publication is added to column a as well. The procedure is the same for the columns a and (a). If it is not possible to recognize the category from the title or abstract at all, the publication is added to the text column. If the category is already determined by the title, it will not be counted to the abstract or text, but it can be added to the keywords.

There are only 9 publications in the keywords column, because only clear keywords like “industrial” count. If the keywords are not clearly about the category, like “door lock”, they do not count.

There are 30 publications for IoT in general, 22 for consumer devices, 33 for enterprise, 14 for industrial and one publication, which is not related to IoT, but has some serial number with iot in the title. Most of the time, the publications for enterprise can be categorized with the title alone, 10 directly and 15 not clearly with related words. Overall, the most publications can be categorized without reading the whole text (but not without reading the abstract), in only 15 cases, further reading is needed. The keywords usage is not good, as only 9 are clearly categorizable.

### V. Results

All results are from the previous research in early April 2020 as described in Sections III and IV. Figure 1 shows the percentage of the different categories according to the search results for only the term “IoT” in the different research libraries, compared with the manual study.

In the manual study, about 65 percent of all publications can be categorized. Semantic Scholar and Springer Link are near to this number with 58 and 77 percent. But with different weightings of the categories. This may be due to the limited number of samples in the manual study of 100.

IEEE Xplore shows a significantly lower number of results if the search term is expanded with the categories. This is due to the search method of IEEE Xplore, since only the metadata (title, abstract and keywords) are searched by default. This procedure has advantages and disadvantages, as will be shown in Section VI.

The other three libraries, ScienceDirect, ACM digital library and Google Scholar are over 82 percent (Google with

### TABLE III. RESULT OF THE MANUAL REVIEWED PUBLICATIONS

<table>
<thead>
<tr>
<th></th>
<th>total</th>
<th>title</th>
<th>abstract</th>
<th>text</th>
<th>keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>not specific</td>
<td>30</td>
<td>2</td>
<td>6</td>
<td>17</td>
<td>6</td>
</tr>
<tr>
<td>consumer</td>
<td>14</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>enterprise</td>
<td>22</td>
<td>4</td>
<td>8</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>not IoT</td>
<td>33</td>
<td>10</td>
<td>15</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>17</td>
<td>32</td>
<td>39</td>
<td>15</td>
</tr>
</tbody>
</table>

Figure 1. Result of the search in research libraries
VI. PROBLEMS

Since IoT is a comprehensive term, some problems arise when searching for scientific publications. Some of them are described in more detail in this section on the basis of the previous study.

We use the same example from Section III: searching for an encryption method for consumer IoT devices, like a smart home sensor. If we use “IoT AND consumer AND encryption”, we get a lower number of search results, but missing general IoT solutions for encryption, which do not include “consumer” in their text. If we change the search term to “IoT AND encryption AND NOT enterprise AND NOT industrial” we might miss some general research, too, but not as much as before. But also publications about production line encryption will be included, because they often miss the term “industrial IoT” or IIoT. Therefore, all unwanted terms must be excluded.

It takes less effort, to search for more specific term like “smart home” instead of IoT to get fewer results. However, by doing that, one misses a lot of publications or has to search for a lot of specific words. A Keyword search would be the best solution, but only a small subset would be returned. A restriction to categories is almost impossible, regardless of the fact that the keywords exists exactly for this purpose.

One of the biggest problems, with the large amount of search results is the difficulty to determine, if the publication is relevant. The results of the manual publications study shows, most of the time the abstract is necessary to get the information. This should be easier if the title or the keywords are better.

Another problem are the different ways in which the search engines work. Depending on the library, a restriction of IoT is useful or not (fewer results from IEEE Xplore with the category).

As a last issue, it is not clear how many publications in total from one category have been published in 2019 and 2020 because every search engine differs in the number of results and some are showing publications in more than one category. Therefore, this research question cannot be answered by this study.

VII. CONCLUSION AND FUTURE WORK

IoT is a too broad term. Nearly every device can be counted as an Internet of Things devices. Therefore, a scientific search about IoT returns thousand of results. No categorization or other distinction is used by many researchers. In this study we only presented results about the big three categories “consumer”, “enterprise” and “industrial”. The more detailed results are not necessary for the biggest problems with IoT and not shown in this paper.

Some weak points about this study are the limitation of 100 papers from only one research library and no further research about the quality of the publications. Nevertheless, the study shows the need of clear categories and a strict use of them. The best way is to include them into the keywords and avoid using words from other categories in the whole publication, as the most search engines including the whole text. In some publications, the term IoT is not necessary at all (e.g., smart home or smart vehicles).
As future work, we are trying to find suitable categories and additional characteristics to build a categorization for every IoT device. Because not only researchers are struggling with the term IoT, standardisation organisations have the same problem, too. They have to decide, which product should be included in a new standard and which restrictions can be applied to all the included ones. They use very broad definitions like in ETSI EN 303 645, consumer devices are defined to be used typical in the home or as wearables, but they can be included in enterprise IoT environments as well: “Consumer IoT devices are commonly also used in business contexts. These devices remain classified as consumer IoT devices.” [110]

REFERENCES
