# Investigating the network traffic of Industry 4.0 applications – methodology and initial results

Gábor Soós, Dániel Ficzere, Pál Varga
T-Systems Magyarország Zrt.
*Dept. of Telecommunications and Media Informatics*
*Budapest University of Technology and Economics*
Budapest, Hungary
{soos, ficzere, pvarga}@tmit.bme.hu

***Abstract*—Industrial networks – both wired and wireless – have been used by manufacturing plants and factories for many years. These networks are not always considered as primary elements of industrial architecture, although they should be. With the advent of Industry 4.0, there are already many innovation initiatives, industry recommendations, and standards in the topic of industrial network infrastructure – but integration guidelines on Industry 4.0 and the networking architectures are incomplete. In the absence of a better methodology, the networks are usually over-designed, but still rigid and are not ready to cover future needs.**

**This paper aims to provide methods on determining some key parameters of existing industrial traffic. Building upon the state-of-the-art, the paper presents a method for testing and analyzing an existing data transfer network traversing industrial traffic. Furthermore, we provide insights on how it is worth to examine industrial networks, and find out whether a given transmission technology can meet the demands – based on the analysis of existing industrial traffic.**

***Keywords*—Industry 4.0, Private LTE, Communication Systems, Cellular Mobile Network, Cybersecurity, Digital Twin, Network profiling**

## I. INTRODUCTION

The objectives of Industry 4.0 cover both the economic and the technological aspects of the transition towards digitalized and automated industry. Primary goals include fully automated manufacturing technologies and developing systems that can execute industrial and manufacturing processes cost-effectively and flexibly with IIoT (Industrial IoT) and 5G. In modular, structured intelligent factories, processes are monitored in real-time. Also, industrial devices – sensors, actuators, production line elements, moving devices – communicate, collaborate, and interact with each other and humans in real-time as islands. It is essential to develop and scale well-designed networks to serve these communication needs.

To properly understand the future challenges in the field, current research directions in the existing factories, including even the low-level communication solutions must get revealed. From the communication service provider point of view, the most valuable, most critical resources and parameters have to be determined and fixed at the beginning of defining the network dimensions. The various data transmission solutions have different advantages and drawbacks – see comparisons in [1] [2] – especially for latency [3], throughput and coverage characteristics. This paper addresses the challenge of investigating the network traffic characteristics' requirements for Industry 4.0 and turning these into engineering best practice methodology.

The rest of this paper is structured as follows: Section II discusses the related work, presenting the main perspectives on industrial networks and the relevance of 5G technology in this area. We introduce some key parameters to characterize a network and the fundamental issues with the limitations of 5G as a supplemental industrial technology in Section III. Section IV describes our measurement methodology and presents an analytical and a graphical presentation method that can highlight the strong and weak points of a given network traffic capabilities. Initial results of the methodology applied in a real scenario are presented and discussed in Section V. Section VI concludes the paper.

## II. RELATED WORK, OVERVIEW OF INDUSTRIAL NETWORKS

There are fundamental differences between wired and wireless network solutions. In the case of industrial networks, some aspects of network characteristics become more relevant, such as stability, determinism, security, error rate, and latency. Usually, wired solutions provide a more reliable and stable transmission. A wired network are considered more secure as it has physical protection and hence it is harder to get unauthorized access to it. Also, it provides a lower latency and bit error rate (BER) than wireless networks. Wired networks can be inflexible in terms of mobility, and installations can take longer to set up because more components are required to complete the process. Maintenance and scaling in a wireless network can be much cheaper and more manageable.

There are several protocols for industrial networks, and in almost all cases, they are wired solutions. The market for industrial networks is broadly fragmented, as shown in Table 1. Just like fieldbuses, Industrial Ethernet has not standardized on a single protocol. Instead, each network type serves different applications in the industrial market. EtherNet/IP [4] is the most popular Industrial Ethernet protocol with a 15% market share. Amongst fieldbuses,

978-3-903176-31-7

PROFIBUS DP [5] was the most commonly installed network in 2019. Wireless technologies have gained impressive growth rates of 30% or more in each of the past three years, although they still represent a relatively small portion of the overall market of newly installed nodes. Moreover, there are further, new wireless implementations coming. Cellular technologies such as Private LTE and 5G networks can be enablers for Smart and Flexible Manufacturing in factories.

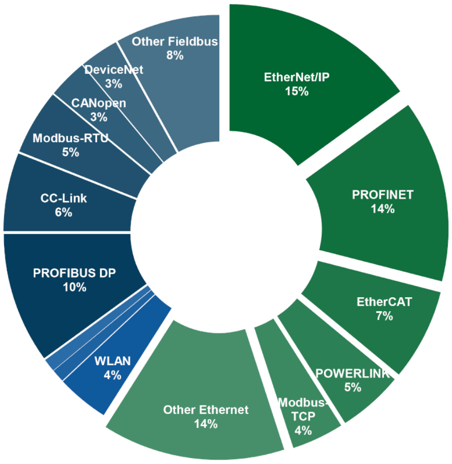

Fig. 1. Market shares of newly installed nodes for fieldbuses, Industrial Ethernet and wireless technologies in 2019 [6]

When examining the future directions of industrial use-cases – such as integration of production lines and increased collaboration – the architecture changes, and communication alters towards different characteristics [7] to cover requirements (e.g., physical placement flexibility). Furthermore, there may be diverse communication needs between different workflow processes. Hence, these networks also called as Industrial or Non-Public Network (NPN) [8] services shall be designed [9] to meet current and future needs of emerging 5G use-cases.

When turning towards 5G network solutions, it is clear that the features of the network (throughput, latency, reliability) and its general usability by different use-cases will be greatly influenced by the architecture of the built 5G network. 3GPP [10] further develops various solutions for the NPN, providing more options in terms of network architecture. These network solutions have numerous advantages, including reduced latency, dedicated resources, the possibility of unique network management, or keeping traffic in the industrial area. Also, the enterprises can utilize complete 5G network segments as they would own it and make it possible to cover new type of use-cases such as Time Sensitive Networking (TSN). With the presented measurement method, architectural effects can also be measured through end-to-end measurements. The presented measurement method can help to define what type of NPN architecture is the most suitable for given use-cases or industrial traffic.

## III. DESIGNING AND MANAGING INDUSTRIAL NETWORK

### A. *Defining the main KPIs*

The network architecture should be designed according to the main characteristics of the communication – which first have to be determined. The most critical questions are which kind of endpoints wish to transmit from where, how often, and how much data. When further detailing the traffic characteristics, the packet size and the sending frequency are important factors. Based on [11], let us highlight the key parameters and aspects to consider for planning the network resources:

- Delay – the delay between the request and the response messages of the network devices determines how fast the system must transfer the packets. However, we can put an artificial delay in the measurement, which can be used to test the device sensitivity on latency.
- Bandwidth – metering this provides insights into the current bandwidth utilization, so the desired network presumably need to improve that.
- Packet size – the traffic handling capability of the existing network is dependent on the packet size and the sending rate. Beside measuring these, it is also worth to examine that, in case of increasing packet sending rate the transmission bandwidth performance would reduce or possibly increase.
- Client mobility – Unique identifiers of the endpoints can be found in different network segments from time to time. For example, in the case of an Ethernet network, the same MAC address receives IP addresses from multiple domains within a specific time.
- Retransmission – we must observe whether packet retransmission causes the system to increase or decrease its transmission capacity. In some cases, sending compressed packets are better for the overall system than experiencing retransmission due to lost data fragments.
- Endpoint (or user) traffic density – determines how much a particular client is transmitting relative to the average within a network segment. If customers in one area generate more than other areas, they have different network needs, so we might want to plan for that.

### B. *Finding the capabilities and limitations of the network*

A wide range of Industry 4.0 use-cases and testbeds already exists [12] but they are still not entirely well-defined yet. There are several publications on what protocols [13] and what architecture [14], [15] will be used, but in practice, there still are just a few concrete production cases. When planning for the data transmission infrastructure, it is essential to know what amount of data, how many packets, between whom, at what distance, and at what frequency we may wish to transmit. Another critical issue is the design of network architecture and its scale. To properly implement a network solution, we need to analyze the traffic requirements between the devices, classify and even prioritize them, and define which is more suitable for p2p (peer-to-peer) connection rather than server-client architecture. On the other hand, in most cases, 5G or private LTE will not be the only transmission technology rather than a supplemental solution to the existing industrial protocols such as PROFINET or EtherNet/IP. As equipment manufacturers can not redesign every product, the supplemental transmission technology should be designed to serve the needs of different kinds of industrial protocols and support additional features. These features can be native L2 forwarding, low latency (1 ms), or extra-low jitter (100 us) [16]. For instance, every PROFINET device sequentially executes its program within a specific cycle time. The inputs are read at the beginning and the outputs are set at the end of each cycle. It means the device's input data must arrive before the next cycle begins; any kind of delay is not tolerated. Furthermore, in time-sensitive applications, PROFINET operates with reduced OSI stack where the Real-Time channel skips the encapsulation steps in the Network, Transport and Session layers [17], [18]. LTE and 5G do not support native L2 transmission in this kind of way.

In packet-switched networks, packet size and packet rate can characterize a transmission medium. The product of the two provides a good indication to characterize the main traffic attributes. Standards and White Papers [1] define the state space related to the limitations of the transmission with a given packet number. However, the question is whether the various vendor implementations fill this space. To verify such questions, a complex measurement methodology is needed to be detailed. However, measuring all of the above parameters is quite a complicated, time-consuming, and costly task.

In this paper, we provide a method to validate whether a given transmission technology can serve the traffic needs; for instance, 5G is capable of serving the traffic needs of a specific PROFINET use-case. Thus, the packet size, packet rate, and packet delay are the most critical parameters to be measured. Mobility – or the change of service area – is a parameter which is needed to be examined individually. In the following sections, we provide a detailed measurement procedure for the three most important parameters and examine how the given transmission technology can serve and handle these requirements.

## IV. METHODOLOGY OF EVALUATING NETWORK CAPABILITIES VERSUS USE-CASE NEEDS

Our goal is to determine whether the network is able to meet user demands and what scaling is required, based on measurements of existing use-cases and networks. Figure 2 gives an overview of the evaluation process, which have the following Passive and Active measurement steps:

- A1: Evaluation of the current local use case with passive network measurements;
- A2: Adding epsilon random deviation to the traffic pattern;
- A3: Add modifications to represent the new use case needs;
- B1: Active measurement to examine the actual features of the transmission channels on site;
- B2: Complex Active transmission measurement to investigate how the network capabilities fit the new use case needs.

The planned evaluation procedure consists of 5 steps. The purpose of type (A) steps are to quantify the communication needs of existing use-cases. We try to find out what sort of network needs of these use-cases will have within a certain time-frame. At the end of the process, the distribution of the traffic KPIs get determined. Type (B) steps are there to actively test an existing or under construction network by measuring their KPIs with artificially generated traffic patterns (that are based on type A steps), as shown by Figure 2.

The state space of and applications' traffic can be described with (1) where PL is Payload Length in Byte and IAT is Inter Arrival Time in milliseconds.

$$A_{n,m} = \sum_{n=1}^{j} \sum_{m=1}^{k} ((PL_n \times IAT_k) \tag{1}$$

Where $A_{n,m}$ is a traffic matrix, describing the state space of the application traffic needs. Furthermore, the traffic of a subscriber can represented as the sum of the different applications, where $S_{n,m}$ is the traffic matrix of an individual subscriber traffic:

$$S_{n,m} = \sum_{i=1}^{l} (A_i) \tag{2}$$

Finally, $\theta^{\varrho}_{n,m}$ represents the required aggregated network traffic matrix as a sum of individual subscribers' traffic:

$$\theta^{\varrho}_{n,m} = \sum_{i=1}^{l} (S_i) \tag{3}$$

Of course this is true only in case of deterministic traffic [19], which can be a fair assumption in case of machines.

### A. Passive measurement steps

*1) A1 - evaluation of the existing network needs:* During the measurement, we examine the discussed two parameters: the PL and the IAT. Since the mean and standard deviation calculation would provide a very rough approximation, we make the following statements. The measurement should determine how often packets of the same length are following each other (IAT). So, in a 3-dimensional figure, the X-axis will be the packet size, the Y-axis will be the inter-arrival period for packets of a given length (IAT), and the Z-axis will determine the throughput of the traffic in the given measurement time-window. Based on this examination method, the traffic pattern can be determined to a given network segment and the scaling characteristics for that type of traffic. Then, we define a descriptive formula, where the two main parameters are packet size and packet rate.

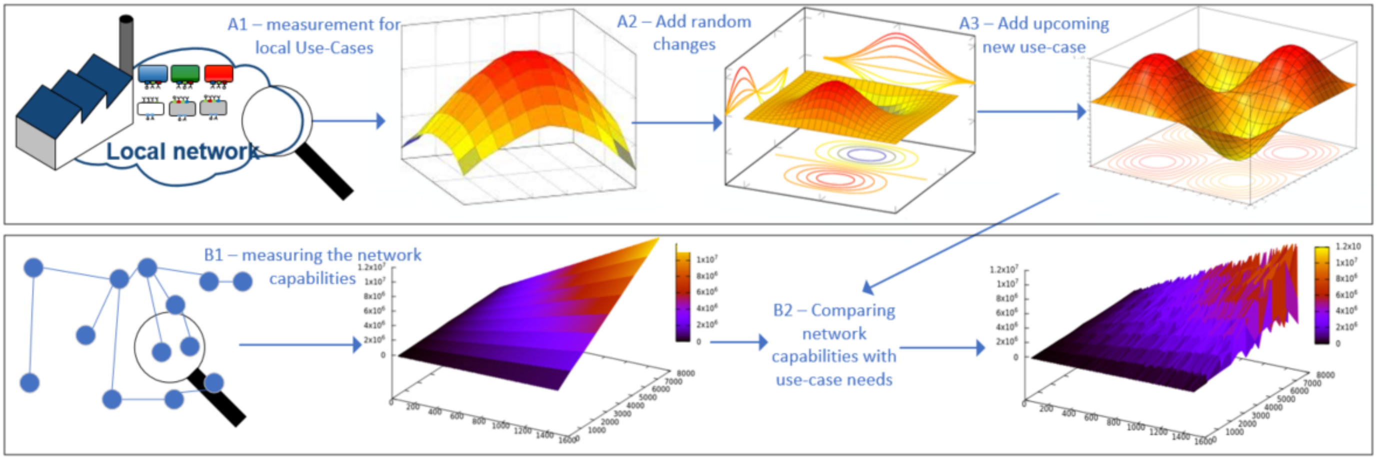

Fig. 2. Mehodology for mapping network capabilities with the industrial use case needs

*2) A2 - Considering epsilon deviation of the traffic pattern:* We apply an epsilon deviation to the traffic pattern resulted in A1. It is assumed that future use-cases will have similar data transfer needs, but they will differ a little bit. During the measurements, an epsilon is defined, which can influence the measurement range to be covered, but still, provide realistic and non-biased demands about the network.

*3) A3 - Adding the effect of the future use-cases:* In A3, we add pre-defined, use-case specific traffic pattern deviations to the existing A2 results. The various A1 traffic patterns that have been epsilon-deviated gets aggregated with new use case patterns, and describes as one common traffic distribution model. Based on the theoretical considerations, Figure 2 shows how the common need for multiple use-cases manifests for the network.

### *B. Active measurement steps*

In this part of the investigation methodology, we use active measurement methods to examine how network capabilities actually fit the various use case traffic needs. In practice, we send packets with different sizes and inter-arrival times through a given DoNUT (Device or Network Under Test). Then, we examine how many packets arrive and how much of them can be handled by the DoNUT. The input parameters are the packet size and the inter-arrival time. The bandwidth, so the output of the transmission is presented in Byte/sec. In our work so far, we are not dealing with real-life measurements, although we have prepared initial input patterns and network, and gathered initial results through simulation.

In our work, we arbitrarily ignored the fact that one data flow could affect other data flows. So, there may be a case where the calculated sum of data transmission capacity is less than the channel capacity, but the channel is still unable to transmit the required amount of data. With this criteria, the capacity of the data transmission channel can be easily described with different size and length of PL multiplied by IAT buckets.

$$\theta^{\gamma}_{n,m} = \sum_{n=1}^{j}\sum_{m=1}^{k}((PL_n \times IAT_k) \tag{4}$$

*1) B1 - Basic active measurements:* It is necessary to determine the transmission characteristics for the measured channel under investigation. As shown in Figure 3, in the case of a theoretical result, a channel with linear and infinite capacity is able to traverse all the theoretical packets. In contrast, the reality is likely to be different. DoNuT will show that certain channels and equipment have limitations for specific packet sizes. Our goal is to construct a function from the measured results that determine the transmission characteristics of the channel. Besides the graphical representation, we can also create approximation functions with multiple variables and then give a Root-mean-square error from the measurement result.

*2) B2 - A complex transmission measurement:* The distribution of KPIs for the specific use-cases is calculated in A3, and the function describing the network in B1 will show what kind of traffic the network can handle and what can not. To double-check the results, we must carry out further measurements. We examine the DoNuT with 1-1 traffic patterns, then we proceed the B1 measurement again to see if there is any bandwidth reserve in the transmission channel. The essence of the measurement is first to take a sample from the given network and then perform a simple statistical analysis on it, analyzing its state space.

### *C. Presenting the results*

We are providing an analytical and a graphical representation of methodology's results. The analytical results are based on the equation (3) and (4), we created an algorithm (1) to validate whether a given network is capable to serve the traffic needs of a use-case. The input of the algorithm are $\theta^{\gamma}$ as the network capacity for traffic matrix and $\theta^{\varrho}$ as the required aggregated network traffic matrix. The output is $\theta_{\phi}$ boolean matrix, which describes $\theta^{\gamma}$ is able to serve the needs of $\theta^{\varrho}$.

```
function Network Traffic calculation $(a, b)$;
Input :  $\theta^{\varrho}$as required aggregated network traffic matrix, $\theta^{\varrho} \in \mathbb{N}^{0 n\times m}$;
         $\theta^{\gamma}$ as network capacity for traffic matrix, $\theta^{\gamma} \in \mathbb{N}^{0 n\times m}$;
         where any $\theta^{\zeta}_{ij}$ corresponds to the $NRT_{ij} \in \mathbb{N}^{0}$ with parameters $IAT_j = j[ms]$; $PL_i = i[bytes]$
Output:  $\theta_{\phi}$ as network capacity able to fill, where $\theta^{\phi} \in \beta^{n\times m}$ is a boolean matrix
$i \leftarrow 1$ $j \leftarrow 1$
while $i <= n$ do
    while $j <= m$ do
        if $\theta^{\varrho}_{ij} \leq \theta^{\gamma}_{ij}$ then
            $\theta^{\phi}_{ij} \leftarrow True$
        else
            $\theta^{\phi}_{ij} \leftarrow False$
        end
        $j \leftarrow j + 1$
    end
    $i \leftarrow i + 1$
end
```

Illustrating the described parameter space requires a 3D representation, where one axis shows the packet size and the other shows the inter-arrival time. The third axis represents the total bandwidth. In a duplex transmission environment where the outgoing direction does not interfere with the incoming, there is only one figure needed, but for half-duplex connections, there should be two. Figure 3 shows the generic transmission graph of the Ethernet 802.3u standard. The data transfer rate varies according to the packet size and transmission rate. This is a calculated figure with the theoretical properties of the Ethernet 802.3u standard, so the bandwidth has a linear connection with the packet size and the IAT.

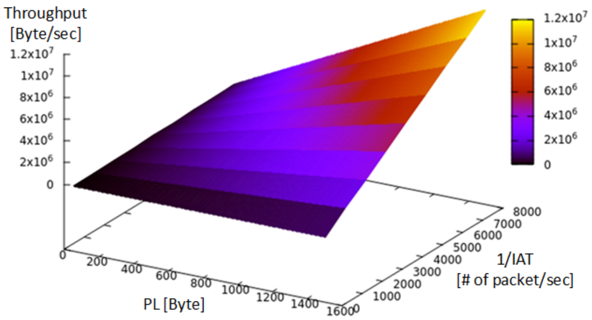

Fig. 3. Theoretical throughput of 100 Mbps Ethernet cable

We can examine the problem from another, practical point of view, where the channel has a theoretical maximum bandwidth of 100 Mbps for Fast Ethernet. It can be determined what number of packets have to be transmitted on it within a given time-frame. It is worth checking if this number of packets can actually be transmitted over the given channel during that time.

## V. INITIAL MEASUREMENT RESULTS AND VISUALIZATION EXAMPLES

To initially validate our methodology, we merely used freely available, open-source traffic patterns, and tested these through the type A, passive measurement method steps. For type B measurements, the method was modified with predefined distribution functions.

### A. Evaluating the traffic of an existing industrial scenario

There is a considerable amount of open-source .pcap capture-files available from industrial networks. The traffic of a typical SCADA-representing .pcap file was analyzed [20], and the result is plotted in Figure 4. It shows that the packet size distribution in the network is quite simple. There are mostly relatively many small packets with low frequency. Larger packets are barely included in the sample; around 300 Byte-long packets with low IAT are represented.

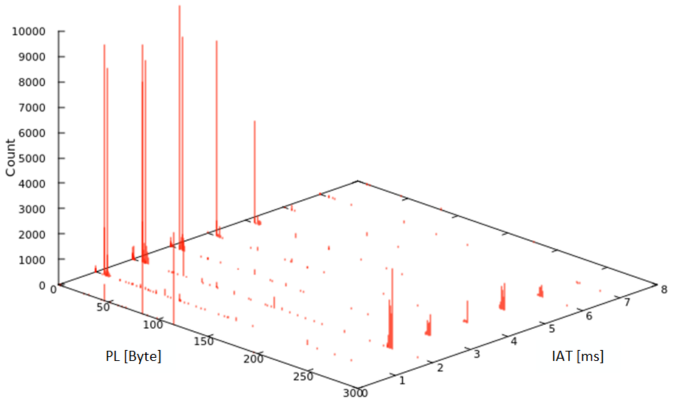

Fig. 4. SCADA-type industrial traffic

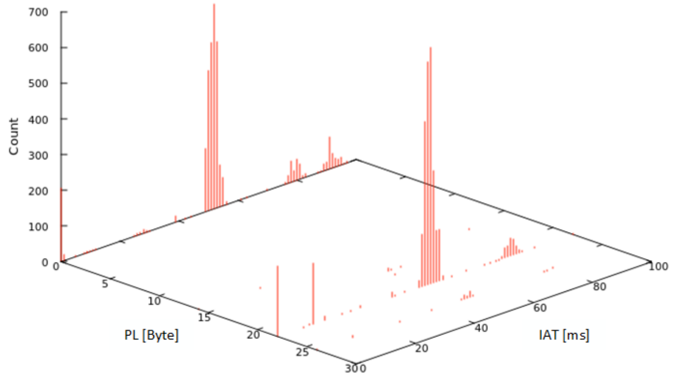

Fig. 5. Skype calls traffic

Figure 5 shows the packet distribution of Skype traffic from a publicly available data source. The packet size and packet-rate distributions are significantly different from the industrial traffic. This traffic pattern contains mainly small, few-Byte-long packets, which can be some sort of keep-alive messages, and there are significant number of 20 Byte packets (typical segment size of the codec), where the IAT has low variance as well. The Figure shows just a few exceptions; there is hardly any other packet size in the network. Based on these two simple examples, we can already make simple conclusions. If Figure 5 traffic was transferred over a network where the packet size is limited to the range between 0-30 Bytes, then the network was able to handle it correctly. On the other hand, SCADA traffic of Figure 4 would not be traverse the network with the same transmission strategy limitations.

In the current paper we simply calculated the network capacity on a theoretical level with the assumption of the Payload Length bucket part follows a uniform distribution. Using the presented investigation method and artificially generated traffic, the next step will be to examine the capabilities of a DoNUT in a wider range of PLs.

## VI. CONCLUSION

In our current work, we provided a simple method that shows whether the planned future industrial network will be suitable to handle the traffic needs if a given use-case or not and which traffic characteristics can be problematic. The contribution of this paper is twofold.

First, for the analysis of existing traffic, we examined a large number of existing researches, and selected a method to quantify the traffic passing through a given network, but without loss of information. The graphical representation methods have been partially improved, making the patterns of traffic more suggestive. In the final part of the traffic modeling, as future work, we intend to add new use-cases to the existing ones, thus providing a more complex picture of the future industrial networks.

Second, we proposed a methodology for examining existing industry traffic described with modified use-cases. As part of this, the paper introduced the measurement and representation method for examining transmission networks both in analytical and graphical way. We presented a simple method for measuring and determining the potential weaknesses of a transmission media or device and then plotting the results. Through this, we can compare the use-case needs with the capabilities of different networks and can define that the network will be suitable for the use-case. It can be an excellent to design and scale future 5G NPNs.

## ACKNOWLEDGEMENTS

This work has been performed in the framework of the H2020 project 5G-SMART co-funded by the EU. The authors would like to acknowledge the contributions of their colleagues. The consortium is not liable for any use that may be made of any of the information contained therein.

## REFERENCES

[1] 5G-ACIA. (2019) 5G Traffic Model for Industrial Use Cases. [Online]. Available: https://www.5g-acia.org/fileadmin/5G-ACIA/Publikationen/5G-ACIA_White_Paper_Traffic_Model/WP_5G_5G_Traffic_Model_for_Industrial_Use_Cases_22.10.19.pdf

[2] Varga, P.; Peto, J.; Franko, A.; Balla, D.; Haja, D.; Janky, F.; Soos, G.; Ficzere, D.; Maliosz, M.; Toka, L., "5G support for Industrial IoT Applications — Challenges, Solutions, and Research gaps," vol. 20, 2020, p. 828.

[3] IEC/IEEE. (2018) TSN Profile for Industrial Automation. [Online]. Available: http://www.ieee802.org/1/files/public/docs2018/60802-industrial-use-cases-0918-v13.pdf

[4] Z. Gong, B. Liu, S. Yang, and X. Gui, "Analysis of industrial ethernet's reliability and realtime performance," in *2009 8th International Conference on Reliability, Maintainability and Safety*, July 2009, pp. 1133–1136.

[5] W. Bao, H. Zhang, H. Li, W. Huang, and D. Peng, "Analysis and research on the real-time performance of profibus fieldbus," in *2009 WRI World Congress on Software Engineering*, vol. 1, May 2009, pp. 136–140.

[6] Engineering360. Best NetFlow Analyzer Software 2020. [Online]. Available: https://insights.globalspec.com/article/11936/what-is-the-most-popular-industrial-network-protocol

[7] A. Neumann, L. Wisniewski, R. S. Ganesan, P. Rost, and J. Jasperneite, "Towards integration of industrial ethernet with 5g mobile networks," in *2018 14th IEEE International Workshop on Factory Communication Systems (WFCS)*, June 2018, pp. 1–4.

[8] 5G-ACIA. (2019) 5G Non-Public Networks for Industrial Scenarios. [Online]. Available: https://www.5g-acia.org/publications/5g-non-public-networks-for-industrial-scenarios-white-paper/

[9] NGMN Alliance. (2019) 5G E2E Technology to Support Verticals URLLC Requirements. [Online]. Available: https://www.ngmn.org/wp-content/uploads/Publications/2019/200210-NGMN_Verticals_URLLC_Requirements_v16.pdf

[10] 3GPP TS 23.501; 5.30.2.1. (2020-03) System architecture for the 5G System (5GS); Stage 2 .

[11] S. Gangakhedkar, H. Cao, A. R. Ali, K. Ganesan, M. Gharba, and J. Eichinger, "Use cases, requirements and challenges of 5g communication for industrial automation," in *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, May 2018, pp. 1–6.

[12] Petr Novak, Jirı Vyskocil, Petr Kadera, Lukas Kathreiny, Kristof Meixnery, Dietmar Winklery, Stefan Biffl, "Engineering Roles and Information Modeling for Industry 4.0 Production System Engineering," 2019.

[13] M. Li, P. Zeng, F. Zhang, H. Wang, and R. Zhong, "Structure and protocols of wireless industrial control network," in *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, vol. 6, Oct 2010, pp. V6–140–V6–143.

[14] F. M. for Economic Affairs and Energy, "Network-based Communication for Industrie 4.0 – Proposal for an Administration Shell," *Plattform Industrie 4.0*, 2016.

[15] B. Kada, A. Alzubairi, and A. Tameem, "Industrial communication networks and the future of industrial automation," in *2019 Industrial Systems Engineering Conference (ISEC)*, Jan 2019, pp. 1–5.

[16] P. Ferrari, A. Flammini, D. Marioli, and A. Taroni, "Experimental evaluation of profinet performance," in *IEEE International Workshop on Factory Communication Systems, 2004. Proceedings.*, 2004, pp. 331–334.

[17] M. Yang and G. Li, "Analysis of profinet io communication protocol," in *2014 Fourth International Conference on Instrumentation and Measurement, Computer, Communication and Control*, 2014, pp. 945–949.

[18] PROFIBUS PROFINET International. (2019) PROFINET Design Guideline. [Online]. Available: https://www.profibus.com/download/profinet-installation-guidelines/

[19] A. Kalør, R. Guillaume, J. Nielsen, A. Mueller, and P. Popovski, "Network slicing for ultra-reliable low latency communication in industry 4.0 scenarios," 08 2017.

[20] Capture files from 4SICS Geek Lounge. (2020) Capture files from 4SICS Geek Lounge. [Online]. Available: https://www.netresec.com/?page=PCAP4SICS