A Modular Framework for Evaluating Smart Grid Communication Protocols over Mobile Networks

Lucas Lutze*, Christian Sorgatz*‡, Christian Freudenmann§ and Michael Rademacher‡*
*Hochschule Bonn Rhein-Sieg, Sankt Augustin, Germany (lucas.lutze@smail.inf.h-brs.de, christian.sorgatz@inf.h-brs.de)
‡Fraunhofer FKIE, Bonn, Germany (michael.rademacher@fkie.fraunhofer.de)
§450connect GmbH, Cologne, Germany (christian.freudenmann@450connect.de)

***Abstract*—This work introduces a modular framework for analyzing the transmission overhead in mobile networks that arises from typical protocols used in the energy industry. Our framework can emulate these protocols using a real LTE-M cell. In particular, the LTE band 72 in the 450 MHz frequency range is evaluated which is currently deployed in Germany for smart grid applications. Compared to previous work, we focus our analysis on the average LTE resource block utilization in the uplink using up to five real devices. Two traffic profiles were provided by different energy industry companies as typical use cases in a smart grid. Our results reveal that the utilization of the uplink resource blocks increases linearly with an increasing number of devices. These results can be used to better understand the scalability of smart grid applications over mobile networks.**

***Index Terms*—Smart Grid, IEC 104, IEC 61850, MQTT, LTE-M, 450 MHz**

I. Introduction

The gathering of information from heterogeneous devices using well-defined communication protocols is imperative for a successful operation of a smart grid. Protocols defined by the IEC 104 [1] and IEC 61850 [2] standards are widely considered for this task. One key challenge for smart grid communication is the rapid increase of distributed measurement devices. The connection of these devices using fixed line network technologies (i.e. fibre) is an economic challenge for grid operators due to the deployment costs. Therefore, mobile networks (Long Term Evolution (LTE), 5G and 6G) are widely considered as a key technology for the communication in smart grids, but provide less bandwidth compared to fixed line network technologies. Therefore, it is important that the smart grid end devices use the available bandwidth very efficiently.

In Germany, the licence for the 450 MHz frequency band expired end of December 2020. In March 2021, the German Federal Network Agency (Bundesnetzagentur) announced that the 450connect GmbH obtained the usage right to establish a nationwide, highly available and fail-safe mobile LTE network for smart grid applications [3]. This 450 MHz LTE network has been under construction since the end of 2021. The network is 3GPP standard-compliant and supports LTE and LTE-M. It is the first LTE network using band 72 in Germany.

Therefore, in this work, we introduce a newly developed framework to evaluate common smart grid communication protocols over mobile networks. In particular, we used our framework to evaluate the efficiency of different protocols using an LTE-M cell operating at 450 MHz. In our framework, different protocols can be emulated and their packet and data volumes as well their effects on the utilization of the LTE resources are recorded. The framework is modular and can be extended by the user with additional protocols or operated with a different mobile network generation (i.e. 5G). To stimulate further research, we made our framework open source[1] including a detailed guide how to use and extend it.

The rest of this work is structured as follows. In Section II, we briefly summarize related work. In Section III we introduce our framework and give an overview of the architecture. To illustrate the practical application of our framework, we showcase an example usage in section IV. Finally, in Section V, we summarize the paper, emphasizing the key contributions and highlights of our framework.

II. Related Work

To the best of our knowledge, there is no modular open-source framework which provides the possibility to evaluate the combination of mobile networks and smart grid protocols. The analysis of the LTE load is particularly relevant when using security mechanisms such as Virtual Private Network (VPN) or Transport Layer Security (TLS), as these can generate a large overhead. For VPN protected IEC 60870-5-104 (IEC 104) connections, the VPN overhead is approximately five times that of an IEC 104 message, although this improves with a larger number of information objects per Application Service Data Unit (ASDU) [4]. This is consistent with the work in [5], which estimates the security overhead of an IEC 104 connection, that is secured according to the IEC 62351 security standard, to be two to three times the bandwidth, if both a VPN and TLS are used.

Smart grid simulation is an important approach for investigating and evaluating the behaviour and resilience of power grids. In this context, various works have contributed to the development of simulation frameworks [6]–[8].

In the field of power grid cybersecurity, several efforts have been made to address the increasing vulnerability of digitally connected grids to cyberattacks. In [8], *Wattson* was introduced, a research tool that combines a power flow solver with network emulators to model real-world power grids including the communication layer. Our framework has the potential to be used for co-simulation with Wattson to perform cybersecurity studies in smart grids using a mobile network.

[1]https://github.com/mclab-hbrs/Smart-Grid-Mobile-Network-Framework

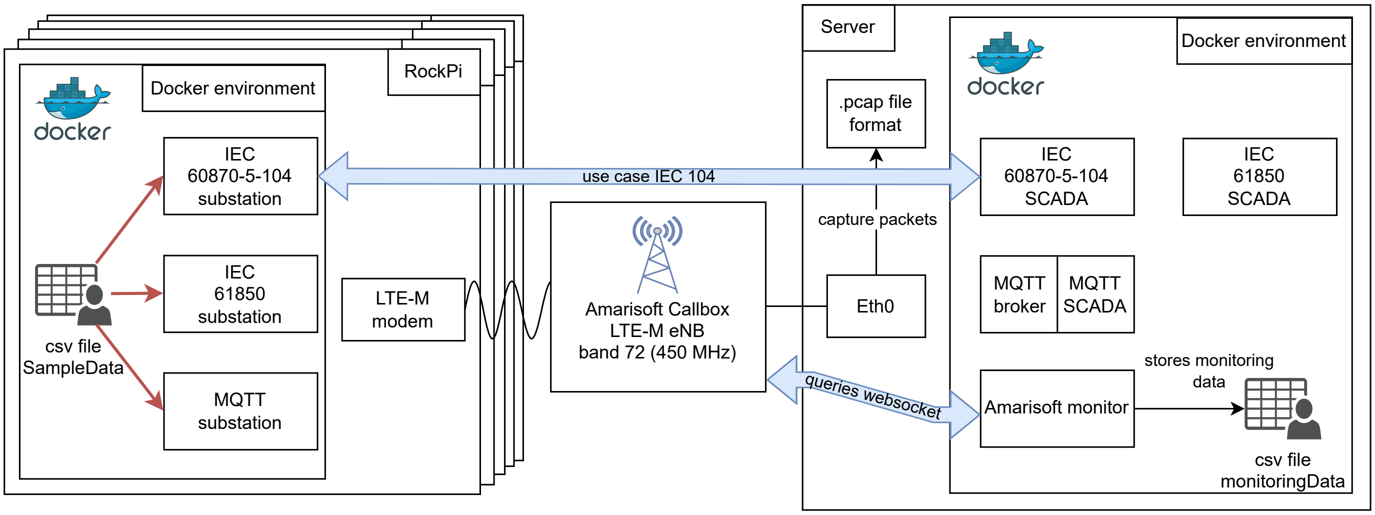

Fig. 1: The main components of our modular framework. The left side corresponds to client devices transmitting data (i.e. in a substation). The right side corresponds to a server application requesting and monitoring data (i.e. a Supervisory Control and Data Acquisition (SCADA) system).

III. Framework

In this section, we provide an overview of the architecture and the main components of our modular framework which is visualized in Figure 1. The framework consist of three main blocks. A User Equipment (UE) side (i.e. a substation), which is the Single-Board Computer (SBC) RockPi 4b+ in our case, a mobile communication technology (i.e. LTE-M) and a server side (i.e. a SCADA system).

To be modular and easily extendable our framework is build upon the idea of virtualization and microservices. We use Docker for this, a software for encapsulating and virtualizing applications using containers. In our scenario, Docker and the virtualization approach provides several advantages. First, during the development phase, containers can be run locally without setting up a complex network infrastructure or test setup. Second, the containers behave in the same way on any of the hardware and operating systems that are supported by Docker. This means the experiments can be repeated later with different hardware and technologies, and the development is not dependent on the UE and server platform used. Third, because all interfaces are well-defined in our framework, adding a new protocol is as simple as implementing two new containers, one representing the substation and one representing the SCADA system.

In our framework, containers are named according to their function in a smart grid. This means that the container which emulates a specific protocol in a substation is e.g., referred to as “IEC 61850 substation”. Similar, the container that provides the protocol implementation which queries the substations is referred to as SCADA container. To emulate a real substation, a SBC or other computer with a mobile network modem is needed. The server can run on a PC or a virtual machine.

As visualized in Figure 1, all protocol containers access the same Comma-Separated Values (CSV) file. This file contains generated data to represent different use cases in a smart grid (cf. Section IV-A). There are no strict requirements for that file, so the data can either be artificially generated or extracted from real world applications. The file is parsed line by line by a protocol container and the data is mapped to the structure and requirements of the individual protocol. Afterwards, the data is transmitted via the modem and the mobile network to the corresponding SCADA container.

The generated packets are captured in the well known *libpcap* file format on the server as shown in Figure 1. Therefore, various tools like Wireshark can be used to conduct an in-depth analysis of the communication. Furthermore, network monitoring data is retrieved via a web socket of the mobile network and stored in a CSV file.

While the overall framework is designed in a generic architecture, we implemented concrete examples for the smart grid protocols and the network components.

A. Protocol Implementations

Three different protocols have been implemented in the aforementioned container representing a substation and a SCADA application.

1) IEC 61850: To include the IEC 61850 standard, the open-source library libiec61850 from MZ-Automation is used [9] in our framework. The library is written in the C programming language and has been translated to Python resulting in the Pylibiec61850. This Python translation is used in other research groups as well [10].

The IEC 61850 standard includes multiple protocols. For the following experiment, only the Manufacturing Message Specification (MMS) is examined, as this is typically used for communication from a substation to the SCADA application and is thus transmitted via the mobile network.

2) IEC 104: The program for the experiments, which implements the IEC 60870-5-104 standard, was created using a fork of the lib60870 from the MZ-Automation library [11].

The measured values were modelled with a float 32 value (Short measured value) data type with the IEC 104 designation M_ME_NC_1(13). A data type without a timestamp was selected because the time stamps are not retrieved in the other test cases and are, therefore, not transmitted. Furthermore, it was decided to use float values, as these are used in the test set-up of the IEC 61850 standard. Scaled values would not be comparable, as the values delivered across the different protocols would differ in accuracy.

The SCADA application can receive readings periodically over the mobile network on request or via interrogation. Interrogation is a method in which the SCADA application requests the substation to transmit all current values of all data points [12]. For this purpose, 16 groups and general interrogation are available [12]. The experiment implements only the general interrogation, as all measured values must be retrieved for comparability. The use of groups would allow the retrieval of a defined set of measured values, for example, only the values of a low voltage feeder.

3) MQTT: The experimental setup for MQ Telemetry Transport (MQTT) uses the open-source application Mosquitto [13] as a broker. The substation and SCADA program use the Python library Paho [14]. The broker is executed on the same hardware as the SCADA container.

The substation publishes the measured values in a JavaScript Object Notation (JSON) format to the broker. The JSON format was chosen because it is widely used and supported by most programming languages. One of the advantages of JSON is the included description of the measured values. This is important because no fixed structure for measured values is defined or agreed by the energy industry for MQTT. This limits the interoperability of systems from different manufacturers. Furthermore, the values of substation 1 are published in the topic 'Substation1/Values'. Publishing the measurements in separate topics would generate a high volume of packets. However, one disadvantage of this approach would be that not all measured values are transmitted simultaneously. For example, the medium-voltage measurements could be sent every minute, and the low-voltage measurements every five minutes. Furthermore, an application that only needs specific measured values would always have to parse the entire JSON data set. It should be noted that the length of the description for the values, the values themselves and the topic name impact the size of the packet. A name based on IEC 61850 was used to describe the measured values, such as *MMXU_MV_Hz_f*. These are already unique and descriptive of the value they represent. The description also directly indicates the datatype.

B. Mobile Network Components

In addition to the smart grid protocol implementations, a mobile network is the key component for our framework. In the following, we shortly describe the modem (the UE) and the commercial Radio Access Network (RAN) and core network.

1) Quectel Modem BG95M4: In our concrete experiments, we use the Quectel modem BG95M4. The modem is a Cat M1 device running the firmware *BG95M4LAR02A02_01.009.01.009*. The modem is connected via USB to the SBC. The connection to the mobile network is established using scripts supplied by Quectel via the Point-to-Point Protocol (PPP). These scripts contain sequences of AT commands sent to the modem via the serial interface. After the successful connection, the script creates a ppp0 interface. This can be used like a standard network interface.

2) Amarisoft Callbox: For the RAN and the core network we used a commercial product from the company Amarisoft [15] called Callbox Classic. The Callbox Classic is a network in a box solution implementing all mayor components of the RAN and core network for a 3GPP mobile infrastructure. For the configuration of the eNodeB, we slightly adapted the default configuration supplied by Amarisoft. The number of available resource blocks was reduced from 25 to 15, corresponding to a Frequency Division Duplex (FDD) bandwidth of 3 MHz — 3 MHz for the uplink and 3 MHz for the downlink. Furthermore, LTE band 72 is selected.

One part of our framework is a python tool to monitor the Amarisoft radio network. This tool queries the Amarisoft application programming interface (API) via a web socket. Among other data, the API returns status information in a JSON format. Selected information, such as the number of Radio Resource Control (RRC)_Connects, the uplink and downlink bit rate, retransmission rates and the latency, are stored in a CSV file (cf. Fig 1). Our tool queries the Amarisoft API every second during the experiments. Different values are returned in an already aggregated way. One of these values is the average percentage allocation of the resource blocks. The value transmitted is the average since the last query of the API, i.e. over a one-second interval. Furthermore, the number of connected UEs is recorded with further information concerning the end device.

IV. Usage Example

In this section, we describe a step-by-step process to use our developed framework. In particular, our goal is to evaluate the efficiency of smart grid protocols using a LTE-M cell operating at 450 MHz and to assess an upper bound for the maximum number of devices per cell. The process consists of the following steps:

(A.) Definition of use cases
(B.) Specification of protocol configuration
(C.) Configuration of the framework, running the experiment
(D.) Evaluation of the generated data
(E.) Interpretation of the results

A. Definition of use cases

We defined two uses case in a typical smart grid to be tested. These use cases, in particular the traffic profiles associated with them, were provided by energy industry companies.

As shown in Figure 2, we consider our example substation containing three low-voltage feeders and a medium-voltage

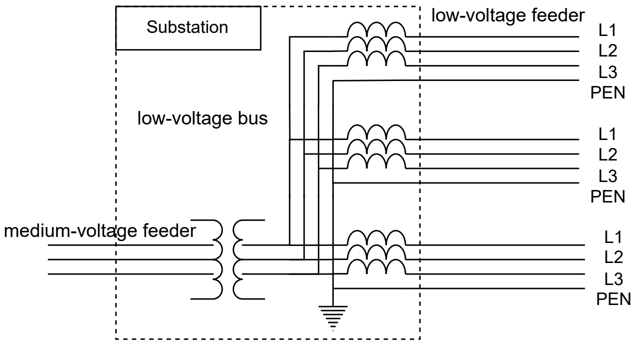

Fig. 2: Sketch of the substation used as a use case.

feeder. We interviewed energy industry companies about the values and collection methods they typically use. The survey revealed that for each voltage level and feeder, power and reactive power are measured for each of the three phases. For the medium-voltage level, the voltage of all phases to the protective earth neutral conductor is also measured. This results in a total of 27 values per measuring interval. Two models for the interrogation methods are used:

1) The measured values of the medium-voltage and low-voltage levels are queried every minute.
2) The measured values are transmitted every 15 minutes as the 15-minute average of the individual measured values.

We will refer to these two models as two different use cases.

B. Specification of protocol configuration

IEC 61850, IEC 104 and MQTT provide various protocol options which affect the results in our experiment. Due to space limitations, we can not list all options directly in this work. However, to ensure reproducibility, they are published alongside our code (cf. Section I).

In general, we configured the protocols with two goals in mind. First, using the same data types (i.e. floats) and update intervals. Second, taking into account best-practices and previous research to reduce the overhead. Some protocols do not have a fixed implementation, such as IEC 61850. Here, the SCADA operator needs to work with the device manufacturer to ensure interoperability [16].

In the following, we describe two example configurations for IEC 61850. First, we configured a continuously open Transmission Control Protocol (TCP) connection which is more efficient than conducting a handshake for a new connection for each data transmission. Second, the overhead of IEC 61850 protocol strongly dependent on the implementation of different functions, such as the query interval of the complete structure of the data model. Since the query of the complete structure of the data model in IEC 61850 has a significant influence on our results, we decide to conduct the experiments with and without model retrieval.

C. Configuration of the framework, running the experiment

Each protocol has been evaluated using an independent experiment with an increasing number of UEs, i.e. substations. While from an architectural perspective the number of emulated UEs is not an issue, but we used up to five due to cost constraints. In future research, additional UEs can be integrated into the system. The limiting factor is the maximum number of UEs the Amarisoft can handle. As SBCs we chose RockPi 4b+ to run the containers, but SBCs with less processing power would have been suitable as well. Figure 3 shows the experimental setup.

Fig. 3: Picture of the experimental setup in our laboratory with one eNodeB (left) and five substations (right).

To assess the current radio cell utilization, several indicators can be considered and accessed via the Amarisoft web socket. For example, the average latency, the average bit-rate or the utilization of available uplink resource blocks. We choose the latter since this value is best suitable for the evaluation of the protocol impact on the radio cell.

We focus the evaluation on the average resource block utilization of the uplink. This focus emerges from the observation that more data volume and packet quantity originates from the substations which has been described by other researchers [4]. When analyzing the values from the Amarisoft API, it is noticeable that in idle mode, when no UE is connected, the average uplink resource block utilization is already at 44%. We critically evaluated this high idle load and found the reason in the LTE-M signalling overhead.

For each protocol, we ran five experiments with an increasing amount of connected substations (UEs). Each experiment lasts one hour.

D. Evaluation of the generated data

The evaluation of the results from the conducted tests are divided into the analysis of the generated packets and the uplink usage reported by the radio cell.

In table I and II, the average uplink load during the trial period of the two use cases is stated. The reported basic utilization (44%) is subtracted to show only the utilization of the UEs.

TABLE I: Scalability of use case 1 based on the average utilization of the uplink resource blocks without base utilisation.

Count	uplink average usage			
	IEC 104	IEC 61850 with model	IEC 61850 no model	MQTT
1	0.00411%	0.00544%	0.00462%	0.00614%
2	0.00893%	0.01198%	0.01004%	0.0128%
3	0.01493%	0.01783%	0.01548%	0.01969%
4	0.01888%	0.02302%	0.02067%	0.0263%
5	0.02363%	0.0285%	0.02505%	0.0313%

TABLE II: Scalability of use case 2 on the average utilization of the uplink resource blocks without base utilisation.

Count	uplink average usage			
	IEC 104	IEC 61850 with model	IEC 61850 no model	MQTT
1	0.00236%	0.00331%	0.00302%	0.00248%
2	0.00443%	0.007%	0.00601%	0.00493%
3	0.00664%	0.01044%	0.00915%	0.00874%
4	0.00877%	0.01404%	0.01207%	0.01107%
5	0.01028%	0.01702%	0.01493%	0.01339%

E. Interpretation of the results

To assess the scalability of a 450 MHz LTE-M cell, the method of a simple linear regression is used in this work. We choose this simple method since our data shows a strong tendency to grow linearly. Based on the generated data for one to five UEs for the different protocols, the maximum number of UEs is estimated. It should be noted that this method has a certain statistical uncertainty since the projection is conducted from five UE to several thousands. Furthermore, no other factors such as the maximum number of simultaneous RRC sessions were taken into account. In addition, the high base load of 44%, which was neglected in the previous tables, must be considered now. This means the radio cell has only 56% of uplink capacity left even if no device is connected.

1) Use Case 1: In Figure 4 on the left, the linear regression result in the interval from 1 to 10 is shown together with the measured values from the first use case. For example, this plot reveals, that six sensors using the MQTT protocol generate similar load as eight sensors using the IEC 104 standard.

With the help of the equations resulting from linear regression method, the maximum number of devices per radio cell can be determined, assuming that the connections are optimal distributed and scale perfectly linear. Furthermore, no limitations, such as the maximum simultaneous RRC sessions, are taken into account. Based on this set of assumption a maximum number of terminals of 11,431 for the IEC 104 protocol is calculated. The IEC 61850 standard shows a maximum number of 9,797 if the model is retrieved. The number of devices can be as high as 10,875 if this is omitted. The fewest number of devices, can be connected when using MQTT, with a value of 8,774 devices.

2) Use Case 2: The right side of Figure 4 depicts the values for the second use case using up to five UEs. This shows the protocols are further apart than in the first use case. Again, the IEC 104 protocol performs best. Five sensors using the IEC 104 standard put about as much load on the system as three sensors with the IEC 61850 standard with model retrieval. It is also interesting to note that MQTT scales better than both versions of the IEC 61850 test setup.

If the maximum number of devices is calculated, up to 27,750 end devices could be connected using the IEC 104 protocol. With MQTT as the protocol with the second most terminals, the radio cell can accommodate 20,028 devices, 38.5% fewer than the IEC 104 protocol. With the IEC 61850 standard, the fewest amount of terminals can be connected to the radio cell. The calculated numbers are 18,741 terminals without and 16,250 terminals with model polling.

Finally, it should be noted that all projections are purely theoretical values. The number of possible UEs would change under using different configurations significantly. In addition, it cannot be ruled out that other limitations may occur. One of these is the number of maximum simultaneous RRC sessions. In the paper [4], the maximum number is given as 400 to 500, with a release timer of five seconds [4]. Assuming that each RRC session is only disconnected after five seconds, one would estimate at 4,800 to 6,000 devices. Since a connection must be established at least every minute in each use case examined. This is also the maximum number for the second use case, as the UEs have to transmit a keep-alive message every 60 seconds. If this could be avoided, more connected devices would be possible.

It should be noted here that a radio cell normally consists of three sectors, each of which would support this number of UEs. The radio network of 450connect GmbH will not only be used to connect end devices for smart grid network but also devices such as smart meter gateways or end devices for mission-critical push-to-talk communication. Concerning the protocols and our configurations investigated, the IEC 104 standard seems to be the most efficient.

V. Conclusions and Future Research

By using our framework we have demonstrated that it is possible to perform scalability studies of energy protocols such as IEC 104, IEC 61850 and MQTT, over the radio interface via an LTE-M mobile network. For the protocol configuration used, and use cases we have outlined, we have found IEC 104 to be the most efficient standard. For other configurations and use cases, this may not be the case. In the absence of our framework, conducting these investigations would have been much more challenging since all the experiments would have needed to be carried out manually. While we only used five sensors for our experiments, it is possible to use many more sensors simultaneously. With our framework, a real eNodeB is used to monitor the traffic of sensors in substations over

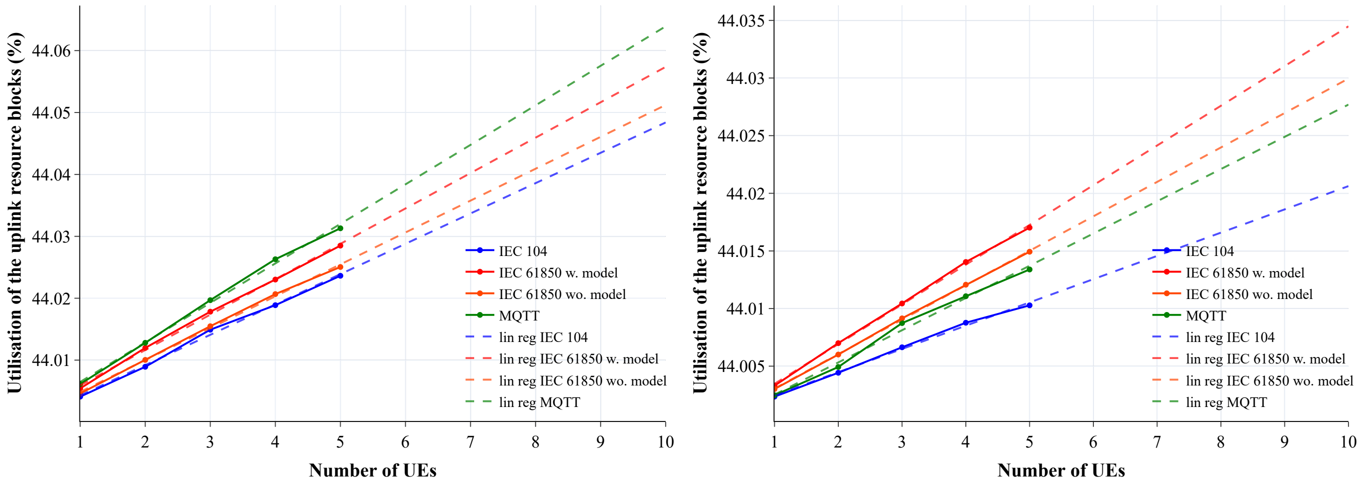

Fig. 4: Linear regression of the average uplink resource block utilisations for both use cases. Use case 1 is shown on the left, while use case 2 is shown in the right. Data has been obtained for 1 to 5 UEs and the regression has been for up to 10 UEs.

the LTE-M network. This allows us to take into account real behaviour of real sensors and networking hardware. In addition, by using real implementations of common smart grid protocols, it is possible to conduct experiments which are close to the reality.

The modular design of our framework allows the interested users to implement their own protocol modules and make them available to the research community. New use cases and sample data are fed into our framework via a CSV file and can thus be customized for the desired analyses. Future research setups should be looking at the impact on protocol scalability considering security protocols such as TLS or securing the connection through VPNs. We are currently investigation possibilities to add TLS and VPN support to the framework in order to carry out this kind of research. Due to the architecture used, this should be possible with minimal effort. The experiments have shown that the packet and data volume of the IEC 61850 standard is very dependent on how the data has been modelled and the implementation of mechanisms. Here, the impact of larger models or the end device-implementations could be investigated. The applied data can be exchanged in the framework. Another research focus could be optimizing the radio cell configuration. Adjustments to this aspect could significantly improve the scalability.

Another research aspect could be tests and experiments with an adapted scheduler. Moreover, the concurrency of application events through an adapted scheduler for smart grid applications or the use of random offsets to transmit regular data in the uplink could be of interest. This scheduler would need to be integrated into our framework.

Acknowledgment

This work has been partially funded by the German Federal Office for Information Security (BSI) under project funding reference number 01MO23003B (PlusMoSmart). The responsibility for the content of this publication lies with the authors.

References

[1] International Electrotechnical Commission, "Telecontrol equipment and systems - Part 5-104: Transmission protocols - Network access for IEC 60870-5-101 using standardtransport profiles," International Electrotechnical Commission, Tech. Rep., June 2016, [Online]. Available: https://webstore.iec.ch/publication/25035 [Accessed: 06-08-2023].

[2] ——, "Communication networks and systems for power utility automation - ALL PARTS," International Electrotechnical Commission, International Standard, January 2023, [Online]. Available: https://webstore.iec.ch/publication/6028 [Accessed: 06-08-2023].

[3] Bundesnetzagentur. 450 mhz. [Accessed: 06-08-2023].

[4] P. Muszynski, "D6.1.3: Advanced smart grid communication concept," [Accessed: 06-08-2023].

[5] K. Ghanem, I. Abdulhadi, A. Kazerooni, and C. McGookin, "Communication requirements for future secondary substations to enable dso functions," in *CIRED 2020 Berlin Workshop (CIRED 2020)*, vol. 2020. IET, 2020, pp. 451–454.

[6] M. Vogt, F. Marten, and M. Braun, "A survey and statistical analysis of smart grid co-simulations," *Applied energy*, vol. 222, pp. 67–78, 2018.

[7] K. Mets, J. A. Ojea, and C. Develder, "Combining power and communication network simulation for cost-effective smart grid analysis," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1771–1796, 2014.

[8] L. Bader *et al.*, "Comprehensively Analyzing the Impact of Cyberattacks on Power Grids," in *2023 IEEE 8th European Symposium on Security and Privacy (EuroS&P)*. IEEE, 2023.

[9] Mz-Automation, "GitHub - mz-automation/libiec61850: the open-source library for the IEC 61850 protocols," [Online]. Available: https://github.com/mz-automation/libiec61850 [Accessed: 06-08-2023].

[10] S. Chen, "pylibiec61850," 2023, [Online]. Available: https://gitlab.com/thu_smartgrids/pylibiec61850/ [Accessed: 06-08-2023].

[11] Ellepdesk, "GitHub - ellepdesk/lib60870: implementation of the IEC 60870-5-104 protocol," [Online]. Available: https://github.com/ellepdesk/lib60870 [Accessed: 06-08-2023].

[12] P. Matoušek, "Description and analysis of iec 104 protocol," [Accessed: 06-08-2023].

[13] The Eclipse Foundation, "Eclipse Mosquitto," [Online]. Available: https://mosquitto.org/ [Accessed: 06-08-2023].

[14] T. E. Foundation, "Eclipse Paho," [Online]. Available: https://www.eclipse.org/paho/ [Accessed: 06-08-2023].

[15] Amarisoft, "AMARI Callbox Series," 2023, [Online]. Available: https://www.amarisoft.com/products/test-measurements/amari-lte-callbox/ [Accessed: 06-08-2023].

[16] P. Bishop and N.-K. C. Nair, Eds., *IEC 61850 Principles and Applications to Electric Power Systems*, 1st ed., ser. Springer eBook Collection. Springer International Publishing and Imprint Springer, 2022.