Energy Informatics

Quantifying geospatial interdependencies of ICT and power system based on open data

Anna Volkova[*], Jonas Terzer and Hermann de Meer

From The 10th DACH+ Conference on Energy Informatics
Virtual. 13–17 September 2021

*Correspondence:
anna.volkova@uni-passau.de
University of Passau, Innstrasse 41,
94032 Passau, Germany

Abstract

In order to evaluate the effectiveness of innovative services and technologies spanning over the ICT-enabled power system, realistic models are required. Although nowadays there is a wide range of power system data models, these models do not include a lifelike representation of the ICT system architecture and do not consider the level of interconnectedness of the two systems. In this paper, we propose a methodology for quantification of the geospatial relations between ICT and power system based on openly available data. We describe a graph-theoretic approach, formulate a problem of assessing geospatial relations and discuss the methods required to process publicly available data, retrieve scenarios for different regions and quantify the level of interdependence between ICT and power system.

Keywords: Data analysis, Interconnected systems, Interdependency, Cyber-physical energy systems, Smart grid

Introduction

With the growing penetration of distributed generators and renewable energy sources, the introduction of virtual power plants and the subsequent need for coordination of distributed smart power system components, the role of the Information and Communication Technology (ICT) becomes indispensable. One of the most visible trends of grid digitization is a continuously growing number of distributed intelligent communication-based power grid services. In order to evaluate the reliability and effectiveness of these services, analytical and simulation-based techniques are usually applied to representative evaluation scenarios, which describe the composition and dependencies between components of the system.

Such evaluation scenarios have to take the spatial properties and interdependencies of the real infrastructures into account. However, while geospatial interdependencies have been identified as one of the key dependency types (Rinaldi et al. 2001; Hokstad et al.

Volkova *et al. Energy Informatics* 2021, **4**(Suppl 3):25

Page 2 of 12

2012; Tøndel et al. 2018), they are usually not considered. Over the last decade, multiple simulation approaches for the interconnected ICT and power systems have been developed for wide-area smart grid applications (Müller et al. 2012; Lin et al. 2012; Georg et al. 2013). Pietsch et al. (2020) have proposed an approach to design scenarios for the communication network between distributed generators. Oest et al. (2021) have described an evaluation model for the optimized scheduling of virtual power plants. These publications have different objectives, but they all use various assumptions regarding the composition of the interconnected ICT and power system in order to construct evaluation scenarios. The validity of these assumptions is even more crucial, when the reliability of the interconnected system should be evaluated, for example, in case of the distribution grid restoration (Stark et al. 2021).

Our key insight is that realistic assumptions regarding the structure of the interconnected ICT and power system can be obtained from openly available data. In contrast to the transmission grid, the distribution grid significantly relies on public communication infrastructures, therefore, both power grid and Internet service provider data is required. While official data is reasonably unavailable, open data can still highlight the regional penetration of ICT access points. In the domain of the power systems, multiple projects (CIGRE Study Committee C6 2014; Meinecke et al. 2020) have achieved visible results in developing representative data models of actual power grids. Thus, the Sim-Bench project (Meinecke et al. 2020), has already used an open data approach in order to extract geospatial properties of distribution grids. Thus, the methodology described by Kays et al. (2016) defines several vital methods applicable to extract geospatial properties of the power grid from the open data, such as application of Voronoi tessellation. Furthermore, a valid approach to determine supply area and type of a distribution grid scenario is discussed in Kittl et al. (2018). However, for the interconnected ICT and power system, there is a lack of insights into how public communication systems are collocated with the power grid.

In this work we focus on cellular ICT systems due to the following aspects. Being already present in all types of regions (rural, semi-urban, urban), the role of cellular networks to enable future wide-area power gird communications will only grow (Borenius et al. 2021) with the introduction of 450MHz (Wissner et al. 2020) and 5G (Dragičević et al. 2019) technologies, especially in rural regions. Next, cellular network geodata for the area of Germany is significantly more complete due to the fact that cellular nodes are more visible at the surface compared to usually hidden underground Digital Subscriber Line (DSL) or optical networks and easier to capture by using radio-based crowd-sourcing applications. However, retrieving assumptions regarding the placement of ICT nodes in relation to the distribution power grid from open data is a challenging task. First, interdependencies between power and ICT system should be formalized in a way quantitative methods are applicable. Second, open and crowd-sourced data are neither complete nor reliable, thus requiring extensive correction, enrichment and merging of different data sources for the various region types.

In this paper we consider described challenges and make the following contributions. We describe a graph-theoretic approach, which provides a model of an interconnected ICT and power system within a geographic scenario and formulate a problem of finding geospatial relations. We apply the methodology on openly available datasets and perform multiple steps for extraction, correction and analysis of the crowd-sourced data as well as

Volkova *et al. Energy Informatics* 2021, **4**(Suppl 3):25

Page 3 of 12

for scenario selection. We quantify geospatial properties of power and ICT system by performing a statistical analysis as well as discuss the dissimilarities in the system structure for different types of regions.

Methodology

This section describes the modelling approach, identifies required parameters for both systems and proposes a methodology to retrieve these parameters from openly available data.
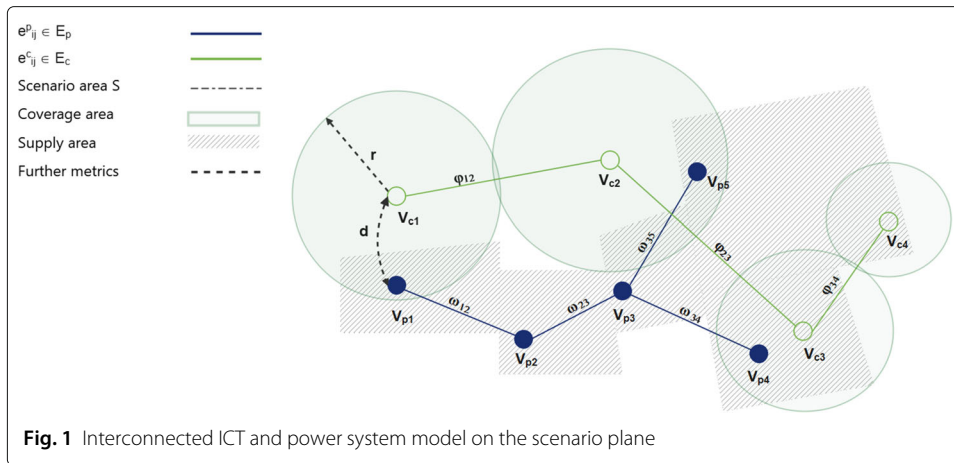
Problem formulation

We define power grid as a weighted graph $G_p = (V_p, E_p)$, where V_p is a set of n power system nodes and E_p is a set of power lines. Each power line $e_{p_{ij}}$ has a weight $\omega_{ij} \in \Omega$, which represents the length of the power line between two power system nodes V_{p_i} and V_{p_j}. The supply area A_i is defined as the geographic area around the power system node V_{p_i}. In this work, we focus on the MV and MV/LV grid level and assume that V_p is a set of MV and MV/LV voltage substations. We also assume that two power system nodes are always connected by a power line through the shortest path, therefore, we assume that ω_{ij} defines the distance between two power system nodes.

The ICT system is defined as a weighted graph $G_c = (V_c, E_c)$, where V_c is a set of k ICT nodes and E_c is a set of edges. Each edge $e_{c_{lm}}$ has a weight $\phi_{lm} \in \Phi$, which represents the distance between two ICT nodes V_{c_l} and V_{c_m}. This approach allows to model multiple wireless and wired ICT technologies. In this work, we focus on cellular networks and define V_c as cellular base stations. We assume a cell to be a circle. This assumption is not always true for modern cellular networks transmitting in an arc or wedge, however, is required to merge our base station model with the models used by the most popular ICT data sources. Thus, we define our base station model with two parameters: r_l as the range of a cellular base station V_{c_l} and $Cov_l = \pi r^2$ as coverage of the base station.

Scenario $S = \{area, type\}$ is a plane in a two-dimensional space, which represents a particular geographical region and is defined by the following two parameters: the area size S_{area} in km^2 and type $S_{type} = \{rural, semi\text{-}urban, urban\}$. Scenario types have been selected according to known grid data models, such as Kays et al. (2016). The problem of quantifying the geospatial relations of the interconnected system is defined as *how to quantify the co-location of a power system graph G_p and an ICT system graph G_c in scenario S, so that both graphs keep realistic geospatial relation to each other and the scenario S*. Figure 1 depicts an exemplary model of interconnected ICT and power system and indicates different types of the geospatial relations between both systems.

We quantify the geospatial relations between ICT and power system according to the following distance metrics. First, ω_{ij} and ϕ_{lm} define the distances between nodes of power and ICT system respectively. Second, the geospatial distance between power system node V_{p_i} and ICT system node V_{c_l} can be defined as a distance d_{il}. Third, the amount of nodes within the supply area and coverage quantifies in which intensity one system is dependent on the other. We say V_{p_i} is dependent on ICT node V_{c_l}, if V_{p_i} belongs to Cov_l. This implies that a power system node uses communication services of an ICT node. However, a power system node can use services of multiple ICT nodes. ICT node V_{c_l} is dependent on V_{p_i}, if V_{c_l} belongs to A_i, which implies the power system node supplies the ICT node.

Volkova *et al. Energy Informatics* 2021, **4**(Suppl 3):25

Page 4 of 12

Fig. 1 Interconnected ICT and power system model on the scenario plane

The amount of ICT nodes of V_c within the supply area A_i of the power system node V_{p_i} is an instance of a point-in-polygon problem (Haines 1994) and is defined as follows. Let V_{c_l} be a point and A_i a polygon generated by V_{p_i} on a plane S, then the amount of the nodes within the supply area A_i is defined as:

$$\Theta(A_i) = \sum_{l=1}^{k} \theta_l, \ \theta_l = \begin{cases} 1, \text{ if } V_{c_l} \in A_i \\ 0, \text{ otherwise} \end{cases} \tag{1}$$

The number of power system nodes $V_{p_i} \in V_p$ in coverage Cov_l of the ICT node V_{c_l} is defined as follows:

$$\Gamma(Cov_l) = \sum_{i=1}^{n} \gamma_i, \ \gamma_i = \begin{cases} 1, \text{ if } V_{p_i} \in Cov_l \\ 0, \text{ otherwise} \end{cases} \tag{2}$$

The geospatial relation of both systems to scenario S is represented by the number of nodes of both systems $|V_p|$ and $|V_c|$ with respect to the scenario area as a density of nodes per km^2: $density_p = \frac{|V_p|}{S_{area}}$, $density_c = \frac{|V_c|}{S_{area}}$. The presented model also allows to determine the different cases of dependency between nodes of both systems: from the complete absence of dependency to mutual dependency, when the ICT node belongs to the supply area of the power system node, which in turn uses communication services of this ICT node.

We also consider that the ICT system structure can vary depending on the scenario type. Thus, shifting the view from an urban to a rural scenario will change not only the density of the ICT nodes but also the typical cell size r_{avg}. In order to create a realistic model of interconnected ICT and power system for a particular scenario S, multiple metrics are required. While the assumptions regarding ω_{avg} and r_{avg} can be estimated based on the different data sources such as Prettico et al. (2016), real values of Φ, Θ, Γ and d are neither available in the related work nor published by power system or communication infrastructure operators. The resulting research objective is formulated as follows: *Given a geospatial dataset of power system nodes with locations and voltages, a geospatial dataset of ICT nodes with locations and ranges, a set of the observation scenarios S_o, compute the values Φ_{avg}, Θ_{avg}, Γ_{avg}, d_{avg}. Observation scenario S_o is an instance of scenario S on a real geographical area. Location features of datapoints in power and ICT system datasets define V_p and V_c for the given scenario respectively. The research objective therefore can be tackled as a set of problems of computational geometry.*

Volkova *et al. Energy Informatics* 2021, **4**(Suppl 3):25

Page 5 of 12

Quantification of the geospatial relations

The estimation of the spatial distances Φ, Ω for power and ICT system respectively and d for the interconnected system is done based on the geographic location of power and ICT nodes in terms of coordinates. We assume that crowd-sourced datapoints follow the pattern of a realistic system, and for both power and ICT system the nearest neighbours are most likely connected. Distance matrix algorithm (Dokmanic et al. 2015) can be employed to estimate the distances between the nodes within a scenario S. This algorithm aims to find datapoints with the closest spatial features to the features of a given datapoint.

In order to estimate the values of Θ, the supply area A_i of a power system node V_{p_i} is identified. We assume that loads are not evenly distributed over the scenario area S and the supply areas must therefore have different sizes. Since to the best of our knowledge, no openly available data contains a realistic description of supply areas of power system substations, we split scenario S by assigning a supply area to the nearest substation, while considering the geographic location. Thus, in order to identify A_i for each power system node in V_p, Voronoi tessellation can be applied. According to Okabe et al. (2008), Voronoi tessellation can be performed for bounded scenarios $S \in \mathbb{R}^2$ and a set of n generator points $P = (p_1, \ldots, p_n)$. In general form, Voronoi region Vor_i, $i = (1, \ldots, n)$ is defined via Euclidean distance $d(p, p_i)$ between an arbitrary point p and the generator point p_i as:

$$Vor_i = \{p \mid d(p, p_i) \leq d(p, p_j)\}, \text{ for } j = (1, \ldots, n) \text{ and } j \neq i \tag{3}$$

A supply area A_i of power system node V_{p_i} corresponds to the Voronoi region of this node. The amount of ICT nodes of V_c within the supply area A_i of the power system node V_{p_i} is therefore calculated according to Equation (1).

Different to the supply area estimation problem, the coverage Cov_j of ICT node V_{c_j} is estimated based on the location and range r_j values known from the data. In general form, for a given geospatial dataset $X = \{x_i \mid i = 1, \ldots, N\}$, where N is the total amount of objects in X, the algorithm generates the buffer area with a given radius *rad*. The buffer area B_i is calculated as Zhou et al. (2021):

$$B_i = (p : d(p, X_i) \leq rad_i), \tag{4}$$

where $d(p, X_i)$ is the Euclidean distance between an arbitrary point p and object X_i, rad_i is the buffer radius. By taking the set of ICT nodes V_c as objects and cell range value r_j of ICT node V_{c_l} as radius, Cov_l can be estimated. The observed amount of power system nodes from V_p in coverage Cov_l of V_{c_l} is calculated according to Equation (2).

Data analysis

In this section, we describe the steps required to apply the proposed methodology for extraction of the geospatial relations from publicly available data for the region of Germany at MV and MV/LV grid level.

Data source selection

In this work we have identified several potential publicly available data sources for the power grid and cellular communication network and evaluated those based on four qualitative suitability criteria. First, relevance of data has been estimated based on the presence of the required features, such as location and voltage tags for the power system data and location and range for ICT system data. The relevance is evaluated as high if both of the

Volkova *et al. Energy Informatics* 2021, **4**(Suppl 3):25

Page 6 of 12

features are available, and limited, if only one feature is available. However, the relevance of data can be increased by merging several data sources. Second, availability of the data set for download in a machine-readable data format under an open, permissive license has been examined. Next, completeness and accuracy have been estimated as amount of data per scenario and frequency of the updates respectively. We have set up the following categories: *high* if the data is released by an official source, *medium* if the data is released by a crowd-sourced project, dense data regions and recent updates have been identified during a preliminary analysis and *low* otherwise. According to our observations, relevance and availability of the data appeared to be superficial requirements.

For power system data, the OpenStreetMap (OSM) (OpenStreetMap Contributors) dataset has been selected as data source. Alternatively observed projects either use OSM data or only observe higher voltage levels. The positional accuracy of the OSM geodata has been estimated as accurate within a region of maximum 18 meters (Fan et al. 2014). However, the quality of tagging at the distribution grid level in OSM is rather low, what can be explained by a limited motivation of the community to track power system infrastructures, high amount and concealed placement of power system objects, especially in the bigger municipalities. We continue to use OSM data in this work in order to evaluate the developed methodology and add several additional steps to enhance the quality of results, such as selection of accurate and complete scenarios.

Regarding ICT data, several potential sources, including Federal Network Agency of Germany (BNA) (Bundesnetzagentur 2021), Mozilla Location Service (MLS) (Mozilla 2021), OpenCelliD (Unwired Labs), CellMapper (CellMapper 2021), Funkloch-Karte (Bundesnetza- gentur) and OSM have been observed and summarized in Table 1. BNA and MLS datasets have been selected. The BNA dataset includes official information regarding the location of cellular base stations and has a license preventing publishing and visualization of derivatives. The information regarding the range of base station is not available and has to be extracted from MLS data. The MLS uses radio environment information collected by users of the Ichnaea (Mozilla 2020a) geolocation tracking service to determine the location and range of cell towers including radio technology, operator and country codes. If a cell tower includes several cellular equipment units, e.g. antennas with different Cell Global Identification (CGI), each device will be observed as a distinct ICT node and will therefore have its own dataset entry. This behaviour will be further corrected by merging the data with BNA data, which contains precise base station locations. Further data sources have been studied in detail, such as OpenCellID, however, significant data anomalies, such as the same default range value shared by 58.2% of nodes have been explored. The position accuracy of the BNA dataset is expected to be high, while the location and range values for MLS are estimations based on crowd-sourced data.

Table 1 Review of the observed ICT data sources

Name	Features	Accuracy	Completeness	Availability
BNA	location	high	high	CC BY-ND 4.0
MLS	location, range	limited	limited	CC 0
OpenCelliD	location, range	low	limited	CC BY-SA 4.0
CellMapper	location, range	limited	limited	not available
Funkloch-Karte	coverage	high	high	DL-DE BY-2.0
OSM	location	limited	limited	ODbL

Data extraction and correction

In order to estimate the density and distance metrics for the distribution power system, only substations of certain voltage levels (in this work, we look at MV and MV/LV) have been retrieved. We selected the classification approach based on IEC 60038 (International Electrotechnical Commission 2009), which defines $V \leq 1\ kV$ as low voltage and $1\ kV < V \leq 35\ kV$ as medium voltage. According to this classification, 8224 of 19558 entries of the total dataset were considered to operate in MV and MV/LV voltage levels and selected.

The MLS and BNA data has been carefully extracted with respect to the licensing. For MLS data reverse geocoding has been applied to each data entry by using Nominatim (OpenStreetMap Contributors) and entries with low sample rates have been detected. A low amount of samples implies limited accuracy of the location and range of the datapoint. As discussed in Velasco et al. (2008), triangulation is a valid technique to locate an object based on three or more radio signals, it is thus assumed that the location of datapoints with at least three samples is more precise. Therefore, 8.48% of the datapoints (51124 entries) with less than three samples have been removed from the dataset to increase confidence in the location data.

Next, range values both too high and too low have been discovered and filtered out by determining lower and upper bounds. According to Kabalci et al. (2016), the maximum cell range for Long Term Evolution (LTE) can reach up to 100 km with reduced performance. For Universal Mobile Telecommunications System (UMTS) the range limit is estimated as 70 km (Mozilla 2020b), for Global System for Mobile Communications (GSM) according to extended timing advance feature a range of 120 km is possible (3GPP 2020). 7138 datapoints with a range lower than 10 meters, smallest radio access nodes deployed by mobile operators (European Telecommunications Standards Institute 2018), or a range exceeding the limit of their respective radio technology have been removed from the dataset. Furthermore, the ICT data has been enriched by combining the MLS coverage data with official base station locations from the BNA dataset as follows. For every BNA datapoint the spatially closest MLS datapoints were determined. The coverage areas of the closest MLS datapoints were then merged and assigned to the BNA datapoint as a single, combined coverage area of this particular datapoint.

Scenario selection

Having the power and ICT system datasets, we are now able to capture the differences of the interconnected power and ICT system for different municipalities as scenarios. In contrast to the clustering-based classification methodology used by Kittl et al. (2018); Kays et al. (2016), we focus only on the population density and do not consider land use type. In order to retrieve the settlement data and classification of Germany, we use a dataset of 2019 from the Federal Statistical Office of Germany (Statistisches Bundesamt (Destatis) 2020). This dataset uses the Degree of Urbanisation Classification (DEGURBA) methodology and classifies the local administrative units by considering the geographical contiguity and population density (Eurostat 2018). Thus, 11007 scenarios were retrieved, classified into three types (rural, semi-urban and urban).

However, the limited quantity and quality of power system data hinders the processing of all scenarios. Thus, an approach to evaluate the data completeness was developed, including completeness requirements and scenario evaluation. These steps can be omitted if reliable power system data, e.g. Distribution System Operator (DSO) geodata, is

available. The superficial requirement of scenario completeness is the presence of data points. Thus, the scenarios with zero ICT datapoints and zero both MV and MV/LV datapoints have been filtered out. From 11007 initially available scenarios, 9698 have been removed as zero data scenarios. In order to evaluate the completeness of the scenario regarding the power system nodes, we have calculated a metric of the *completeness of supply* as ratio of number of the nodes per 1 km^2 to population per 1 km^2. We assume that the scenarios with greater completeness of supply have a higher density of power system nodes. The selected scenarios were later validated by applying a demand estimation. The demand per 1 km^2 was estimated as a product of the number of households and an average demand per household, assuming the average household size as 2 people (Bundesinstitut für Bevölkerungsforschung 2021) and demand as 3 kW per household. We estimated the required number of transformer substations to meet this demand based on the average transformer capacity for each region type (Prettico et al. 2016) and compared it to the number of the substation nodes in the datasets. Thus, the scenarios which fit this estimation were considered as conditionally complete. From the 1309 remaining municipalities only 53 satisfied the completeness requirement.

Results

In the previous section we performed multiple steps to increase the quality of datasets. However, due to the limited completeness of the power system data especially for urban zones, we obtained a limited amount of conditionally complete scenarios and merged therefore MV and MV/LV nodes into a single set. At the same time, for the semi-urban and rural scenarios the power system data showed a higher level of completeness. The availability of BNA dataset with reliable location data has significantly increased the quality of the ICT dataset and therefore the precision of the communication-related values estimation.

The geospatial relation values have been retrieved based on the prepared datasets and 53 observation scenarios S_o. Algorithms described in Methodology section were applied in a Python Geographic Information System (GIS) application using the QGIS (QGIS Development Team 2021) Python API PyQGIS, which provides implementations of the distance matrix, buffer area, Voronoi tessellation and count points in polygon algorithms. Overall, we investigated areas of 261.08 km^2 for 3 urban, 71.20 km^2 for 5 semi-urban and 828.04 km^2 for 45 rural scenarios. The smaller area size for the semi-urban scenarios can be explained by the nature of these municipalities, which are usually parts of larger agglomerations.

First, we estimated the values of node density for each scenario. Figure 2a presents the observed average $density_c$ of ICT nodes that was observed over all scenarios. Figure 2b presents the average density $density_p$ of power system nodes. These values follow the expected correlation between the population density and the amount of required infrastructure. The node density of the power system in semi-urban areas with a median of 10.05 nodes per 10 km^2 lies closer to rural values. For the ICT system the density of nodes in semi-urban scenarios is rather closer to rural with a median of 1.63 nodes per 10 km^2, 1.03 nodes per 10 km^2 for rural and 13.22 per 10 km^2 urban.

Next, we have observed the average distance between nodes in power and ICT systems. Figure 2c presents the average distance Φ_{avg} between the ICT nodes, while Fig. 2d presents the average distance Ω_{avg} between power system nodes. Regarding the values of

Volkova *et al. Energy Informatics* 2021, **4**(Suppl 3):25

Page 9 of 12

Fig. 2 Results: Quantification of the geospatial relations between ICT and power system
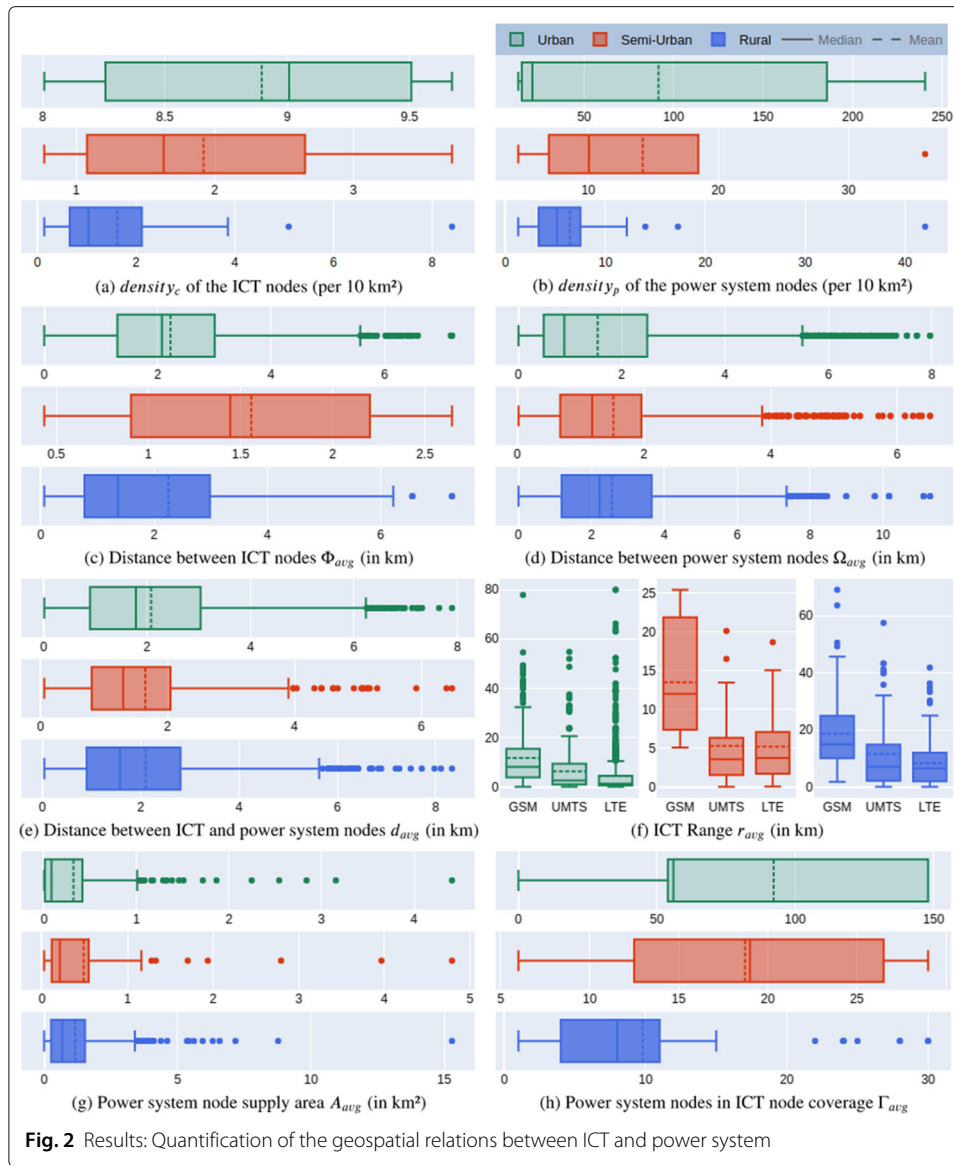
Ω_{avg}, the median values of the distance between power system nodes grow thrice as long from 0.89 *km* for urban zones to 2.23 *km* for rural. These values fall in the same range as observed by Meinecke et al. (2020), which additionally to other methods surveyed several grid operators. For the ICT system, the average distances between the base stations are a mean of 2.25 *km* for rural, 1.56 *km* for semi-urban and 2.23 *km* for urban. Figure 2e shows the average distance between power and ICT nodes d_{avg}. Interestingly, the proximity values of the power system nodes and communication nodes follow the same pattern with a median value of 1.78 *km* for urban, 1.30 *km* for semi-urban and 1.56 *km* for rural areas. The observed minimum and maximum distances correlate in a similar manner.

The high amount of MLS data points with confirmed observations from 18 April 2020 to 18 April 2021 allows studying the average cell range r_{avg}. Figure 2f shows r_{avg} for different region types and technology tags, available in the MLS dataset. Thus, 45.24% of the rural base stations use LTE standard, while 26.49% and 28.27% provide access

according to UMTS and GSM standards respectively. The following pattern repeats for the urban zones with clear domination of LTE with 56.62%, 23.44% UMTS and 19.94% GSM respectively. For the semi-urban scenarios, the domination of LTE cells increases to 69.67%.

One of the core observations of this work is the level of interconnectedness of both systems. In order to estimate it, we have observed the value of Θ_{avg} of ICT nodes within the supply area of a power system node according to Equation (1). First, we have investigated the power system node supply area size, which is presented in Fig. 2g. The mean supply area size grows from 0.32 km^2 and 0.49 km^2 for the urban and semi-urban areas respectively to 1.15 km^2 for the rural areas. The rural areas include multiple outliers with an area size up to 15.29 km^2, which can be explained by the Voronoi tessellation being applied to the whole region and not only the used fraction of land. The observed amount of ICT nodes within the supply area is 0.11 nodes for rural, 0.10 nodes for semi-urban and 0.45 nodes for urban scenarios. The maximum observed values are 3 ICT nodes in one supply area for the rural scenario and 7 for an urban scenario. Regarding the observation of Θ_{avg}, the share of the power system nodes supplying ICT nodes, 18.11% and 11.99% of the power system nodes for rural and semi-urban areas, respectively, supply at least one ICT node, whereas for the urban areas that value grows to 35.75%. Thus, in rural areas every sixth, in semi-urban every ninth and in urban areas every third power system node supplies power to the ICT system.

Analogously, the average value of power system nodes within the coverage of ICT nodes Γ_{avg} was calculated according to Equation (2). The dependency of the power system on the cellular network grows from the rural region (median 8 nodes per one ICT node) to urban (median 56 nodes per one ICT node). The observed values are presented in Fig. 2h. Despite the growing dependency in urban zones, the importance of the cellular network is higher in rural areas, where this type of communication is most likely the only one available, next to power line communication, while the urban environment usually provides a variety of available technologies.

Conclusion

In this paper, we presented a methodology for quantifying the geospatial relations between nodes of the interconnected ICT and power system. We applied our methodology on publicly available geospatial datasets for ICT and power systems. In the absence of official data from the grid and ICT operators, we have used multiple data selection, extraction, correction and enrichment steps to create datasets with all required features. In our future work, we will conduct further validation of the methodology with official power and ICT system datasets and real-world scenarios, once these are made available. Next, a supply area of the power system node will be modeled in detail by looking at lower voltage grid structures and land use types. Finally, we will study geospatial relations between ICT and power system, taking into account different grid voltages, penetration of renewable energy sources, and create an approach to generate realistic models for a given municipality or region.

Volkova *et al. Energy Informatics* 2021, **4**(Suppl 3):25

Page 11 of 12

Declarations

References

3GPP (2020) 3GPP Technical Report 43.030. https://www.etsi.org/deliver/etsi_tr/143000_143099/143030/16.00.00_60/tr_143030v160000p.pdf. Accessed 15 Apr 2021

Borenius S, Hämmäinen H, Lehtonen M, Ahokangas P (2021) Smart grid evolution and mobile communications—scenarios on the finnish power grid. Electr Power Syst Res 199:107367

Bundesinstitut für Bevölkerungsforschung (2021) Zahl der Privathaushalte und Durchschnittliche Haushaltsgröße (1871-2018). https://www.bib.bund.de/DE/Fakten/Fakt/L49-Privathaushalte-Haushaltsgroesse-ab-1871.html. Accessed 10 Apr 2021

Bundesnetzagentur (2021) Elektromagnetische Felder - EMF. https://www.bundesnetzagentur.de/DE/Vportal/TK/Funktechnik/EMF/start.html. Accessed 5 May 2021

Bundesnetzagentur Funkloch-Karte. https://www.breitband-monitor.de/funkloch/karte. Accessed 5 May 2021

CellMapper (2021) CellMapper Project. https://www.cellmapper.net. Accessed 5 May 2021

CIGRE Study Committee C6 (2014) Benchmark Systems for Network Integration of Renewable and Distributed Energy Resources. https://e-cigre.org/publication/575-benchmark-systems-for-network-integration-of-renewable-and-distributed-energy-resources. Accessed 20 Jan 2021

Dokmanic I, Parhizkar R, Ranieri J, Vetterli M (2015) Euclidean distance matrices: essential theory, algorithms, and applications. IEEE Signal Proc Mag 32(6):12–30

Dragičević T, Siano P, Prabaharan S, et al. (2019) Future generation 5g wireless networks for smart grid: A comprehensive review. Energies 12(11):2140

European Telecommunications Standards Institute (2018) ETSI Technical Report 103 230: Fixed Radio Systems; Small Cells Microwave Backhauling. https://www.etsi.org/deliver/etsi_tr/103200_103299/103230/01.01.01_60/tr_103230v010101p.pdf. Accessed 20 Apr 2021

Eurostat (2018) Methodological manual on territorial typologies-2018 edition. Technical report

Fan H, Zipf A, Fu Q, Neis P (2014) Quality assessment for building footprints data on openstreetmap. Int J Geogr Inf Sci 28(4):700–719

Georg H, Müller SC, Dorsch N, Rehtanz C, Wietfeld C (2013) Inspire: Integrated co-simulation of power and ict systems for real-time evaluation. In: 2013 IEEE International Conference on Smart Grid Communications (SmartGridComm). IEEE, Vancouver. pp 576–581

Haines E (1994) Point in polygon strategies. Graph Gems 4:24–46

Hokstad P, Utne IB, Vatn J (2012) Risk and Interdependencies in Critical Infrastructures. Springer, London

International Electrotechnical Commission (2009) IEC 60038:2009 Standard Voltages. https://webstore.iec.ch/publication/153. Accessed 7 Feb 2021

Kabalci Y (2016) A survey on smart metering and smart grid communication. Renew Sust Energ Rev 57:302–318

Kays J, Seack A, Smirek T, Westkamp F, Rehtanz C (2016) The generation of distribution grid models on the basis of public available data. IEEE Trans Power Syst 32(3):2346–2353

Kittl C, Sarajlić D, Rehtanz C (2018) *k*-means based identification of common supply tasks for low voltage grids. In: 2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe). IEEE, Sarajevo, Bosnia and Herzegovina. pp 1–5

Lin H, Veda SS, Shukla SS, Mili L, Thorp J (2012) GECO: Global Event-Driven Co-Simulation Framework for Interconnected Power System and Communication Network. IEEE Trans Smart Grid 3(3):1444–1456

Meinecke S, Sarajlić D, Drauz SR, Klettke A, Lauven L-P, Rehtanz C, Moser A, Braun M (2020) SimBench - A Benchmark Dataset of Electric Power Systems to Compare Innovative Solutions Based on Power Flow Analysis. Energies 13(12):3290

Mozilla (2021) Mozilla Location Service. https://location.services.mozilla.com/. Accessed 20 Apr 2021

Mozilla (2020) Ichnaea. https://ichnaea.readthedocs.io/en/latest/. Accessed 6 Feb 2021

Mozilla (2020) Ichnaea Accuracy. https://ichnaea.readthedocs.io/en/latest/algo/accuracy.html. Accessed 20 Apr 2021

Müller SC, Georg H, Rehtanz C, Wietfeld C (2012) Hybrid simulation of power systems and ICT for real-time applications. In: 2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe). IEEE, Berlin. pp 1–7

Oest F, Radtke M, Blank-Babazadeh M, Holly S, Lehnhoff S (2021) Evaluation of communication infrastructures for distributed optimization of virtual power plant schedules. Energies 14(5):1226

Okabe A, Satoh T, Furuta T, Suzuki A, Okano K (2008) Generalized network voronoi diagrams: Concepts, computational methods, and applications. Int J Geogr Inf Sci 22(9):965–994

OpenStreetMap Contributors OpenStreetMap. https://www.openstreetmap.org/. Accessed 5 May 2021

OpenStreetMap Contributors Nominatim: Open-source Geocoding with OpenStreetMap Data. https://nominatim.org/. Accessed 18 Apr 2021

Pietsch M, Klein A, Steinke F (2020) Merging microgrids for optimal distribution grid restoration under explicit communication constraints. In: 2020 Resilience Week (RWS). IEEE, Salt Lake City. pp 48–54

Prettico G, Gangale F, Mengolini A, Lucas A, Fulli G (2016) Distribution system operators observatory - from european electricity distribution systems to representative distribution networks. Technical report

QGIS Development Team (2021) QGIS - Free and Open Source Geographic Information System. https://www.qgis.org/en/site/. Accessed 16 Apr 2021

Rinaldi SM, Peerenboom JP, Kelly TK (2001) Identifying, understanding, and analyzing critical infrastructure interdependencies. IEEE Control Syst Mag 21(6):11–25

Stark S, Volkova A, Lehnhoff S, de Meer H (2021) Why your power system restoration does not work and what the ict system can do about it. In: Proceedings of the Twelfth ACM International Conference on Future Energy Systems (e-Energy '21). Association for Computing Machinery, New York

Statistisches Bundesamt (Destatis) (2020) Verzeichnis der regional- und gebietseinheiten definitionen und beschreibungen. Technical report

Tøndel IA, Foros J, Kilskar SS, Hokstad P, Jaatun MG (2018) Interdependencies and reliability in the combined ICT and power system: An overview of current research. Appl Comput Inform 14(1):17–27

Unwired Labs OpenCelliD - The World's Largest Open Database of Cell Towers. https://opencellid.org. Accessed 5 May 2021

Velasco E, Chen W, Ji P, Hsieh R (2008) Challenges of Location Tracking Techniques in Wireless Forensics. In: 2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing. IEEE, Harbin. pp 3–6

Wissner M, Sörries B, Zander W (2020) Die 450 mhz-frequenz als wegbereiter der energiewende. Z Energiewirtschaft 44(3):163–175

Zhou G, Zhang R, Huang S (2021) Generalized buffering algorithm. IEEE Access 9:27140–27157

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.