REPRESENTING TOPOLOGY UNCERTAINTY FOR DISTRIBUTION GRID EXPANSION PLANNING

Domenico TOMASELLI
Siemens AG – Germany
domenico.tomaselli@siemens.de

Paul STURSBERG
Siemens AG – Germany
paul.stursberg@siemens.com

Michael METZGER
Siemens AG – Germany
michael.metzger@siemens.de

Florian STEINKE
Technical University of Darmstadt – Germany
Florian.Steinke@eins.tu-darmstadt.de

ABSTRACT

The rising penetration of distributed renewable energy sources (RES), electric vehicle (EV) home-chargers, and heat pumps in power distribution systems can lead to violations of the grid operating conditions. To design suitable grid expansion measures for this challenge, grid planners need a good understanding of the existing infrastructure. Trustworthy, readily usable grid models are, however, often not available. This holds especially for distribution grids at the low-voltage level. In this work, a framework is proposed to generate a probability distribution over an ensemble of different, possible grid topologies for a given area of interest. This probabilistic approach allows to explicitly account for the uncertainty implied by the scarcity of the available information. In a case study with EV home-chargers, it is demonstrated how the proposed ensemble-based framework leads to a robust, uncertainty-aware interpretation regarding the assessment of the existing distribution grid.

INTRODUCTION

Power distribution systems are currently undergoing a substantial transformation due to the rising penetration of distributed RES, EV home-chargers, and heat pumps. The integration of these solutions can often affect the stability of the grid, since the increased and now potentially bidirectional power flow can lead to feeder congestions, i.e., excessive voltage drops and rises, or the overload of grid equipment. To design mitigation measures for these challenges, grid planners need a digital representation of the existing infrastructure.

Trustworthy and readily usable grid models are, however, often not available. Several distribution grids were designed and built decades ago, so grid models sometimes no longer exist. Sometimes grid models are available on paper, but not readily usable in a digitized fashion. Sometimes grid models exist even digitally but are not reliable since they were not continuously updated after the numerous grid modifications that took place over the years.

Grid planners thus must often reconstruct appropriate models of the existing grids, relying on scarce information. This holds especially for the distribution grids at the low-voltage level. Knowledge about the distribution substations and the end-consumers can be assumed from asset management and billing systems, but the grid topology is often less certain. Moreover, distribution lines are typically located underground, making available prior knowledge difficult and expensive to verify [1].

Existing approaches mostly focus on deriving a single best estimate of the connecting topology based on the available information and the typical characteristics of distribution grids. In [2], the authors propose an approach for the automated creation of a distribution grid model based solely on publicly available GIS and statistical data. Similarly, in [3] algorithms are proposed to build a single, best estimate of the topology of large-scale distribution systems at the medium- and low-voltage level using the available information (e.g., consumers, information of the street maps, etc.) and assumed characteristics of a distribution grid. Both approaches, however, do not consider the scarcity of available information nor the resulting uncertainty about the assumed characteristics. The best estimate of the grid topology might well be wrong, and with it all design decisions based on it. Current efforts to overcome these challenges often only represent a partial solution to the underlying problem. For example, in [4] a crowdsourcing approach is proposed to collect the required information to infer a distribution grid, but information about the power distribution lines is often unobtainable considering that they are mainly located underground. In [1,5,6], the authors propose to rely on voltage fingerprints derived from advanced meter infrastructures (AMIs) to overcome the scarcity of information. While AMI represents a reliable source of information, the infrastructure is not installed everywhere yet. Moreover, AMI measurements are often subject to privacy restrictions. In [7,8,9] the authors propose probabilistic models to produce networks with scalable size and random topologies. The quality of the generated power grids is estimated according to how well their topological and electrical features reflect those of standardized test feeders. The main objective of these works is, however, to create numerous test feeders for further research and analysis, rather than utilizing the samples to produce an uncertainty-aware estimate of a specific true grid topology. Furthermore, they do not take specific location information into account to produce the random topologies.

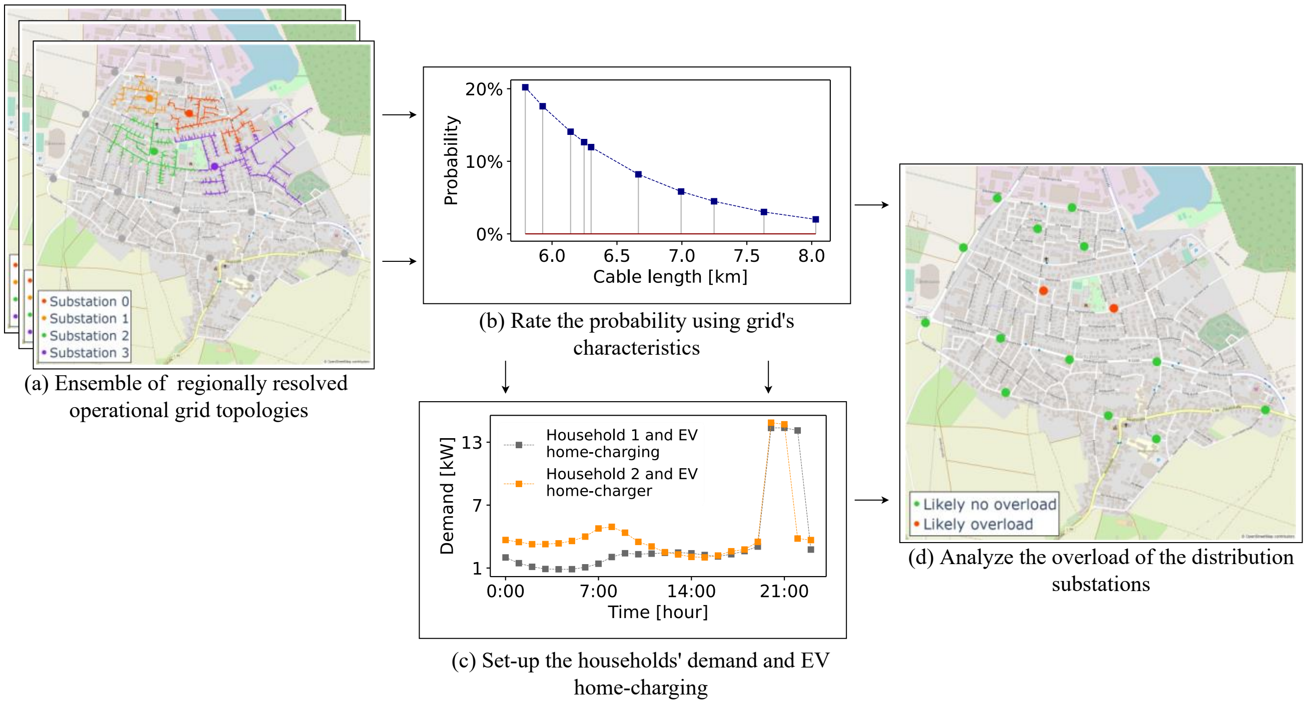

Fig. 1: Schematic of the proposed framework and its application. (a) and (b) depict the two main modules of the approach, i.e., generation of the operational grid topologies for a region of interest and rating the probability of the produced samples. (c) and (d) show the additional information required for the application of the ensemble to analyze the overload of the distribution substations along with the outcome of the analysis, i.e., the substations that likely require reinforcement.

This work proposes to use the available information about a specific distribution grid to generate an ensemble of different, possible grid topologies that appropriately represents the uncertainty in the estimation procedure. To this end, a randomized neighbourhood-growth model is used to assign end-consumers to distribution substations and a minimum spanning tree to derive the connecting topology from the resulting assignment. The probability of each sample is rated according to the estimated grid investment costs as well as domain knowledge about key characteristics of the grid. The presented framework can be applied to conduct uncertainty-aware follow-up grid analyses. For a distribution grid with an increased roll-out of EV home-chargers, this work demonstrates that relying on a single best estimate of the grid topology can mislead grid planners regarding its feasibility, whereas the proposed framework leads to a more robust interpretation. The approach can thus help designing robust mitigation measures for the rising penetration of EV home-chargers when only imperfect grid knowledge is available. The framework along with its application are graphically described in Fig. 1.

The remainder of the paper is structured as follows. In Section II, the proposed approach for generating grid topologies and representing their uncertainty is introduced. In Section III, a case study for its application is presented. Section IV concludes the work with a summary and outlook on avenues for future work.

FRAMEWORK FOR GRID TOPOLOGY ENSEMBLE GENERATION

The ensemble-based framework consists of three parts. First, a randomized neighbourhood-growth model assigns the end-consumers to the available distribution substations. Then, a grid topology using a minimum spanning tree is derived. Finally, the grid topology generation procedure is repeated, and the probability of the resulting samples is rated. A flowchart describing the workflow of the proposed framework is depicted in Fig. 2.

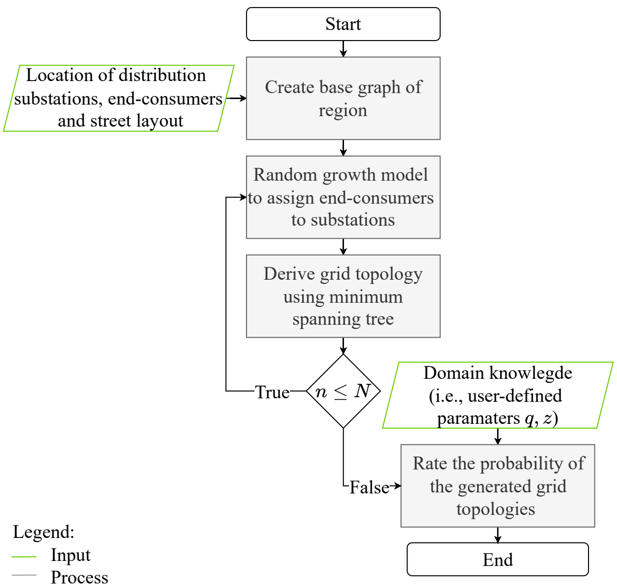

Fig. 2: Flowchart describing the workflow of the proposed approach.

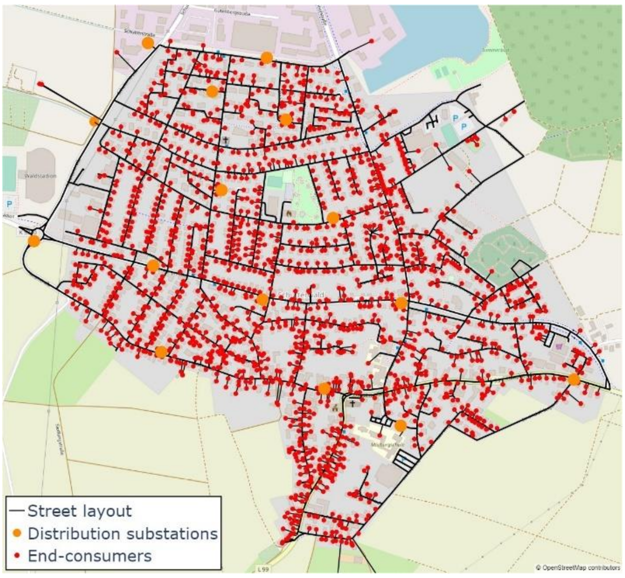

Fig. 3: Base graph $\mathcal{G}_{base}$ for a region in Schutterwald, Germany, including the available distribution substations (orange circles), the end-consumers (red circles) and the street layout (black).

Problem settings

This work considers the problem of modelling regionally resolved operational grid topologies [10]. The operational grid topology (i.e., topology of the grid's bus-branch model) describes how the grid equipment is connected considering the current switch status and the existence of temporary elements [11]. In the following, operational grid topology is referred simply as grid topology. Conceptually, the grid topology is represented as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ is the set of nodes (i.e., distribution substations, end-consumers, and nodes interconnecting the electrical lines) and $\mathcal{E}$ is the set of edges (i.e., electrical lines). Let $\mathcal{G}_{\text{base}} = (\mathcal{V}_{\text{base}}, \mathcal{E}_{\text{base}})$ be the street layout of the considered region, see, e.g., Fig. 3. $\mathcal{E}$ is assumed to follow the street layout, as is often the case in residential, urban areas, i.e., $\mathcal{V} \subseteq \mathcal{V}_{base}$ and $\mathcal{E} \subseteq \mathcal{E}_{\text{base}}$ [3]. Given K distribution substations $\mathcal{G}$ is the union of K disjoint sub-grids, where each sub-grid $\mathcal{T}_k = \left(\mathcal{V}_{\mathcal{T}_k}, \mathcal{E}_{\mathcal{T}_k}\right)$ is a tree (i.e., an acyclic, connected graph) with a distribution substation $s_k \in \mathcal{V}, k \in \{1, \dots, K\}$ at the root.

Randomized neighbourhood-growth model

A randomized neighbourhood-growth model is used to assign end-consumers to the available distribution substations.

Let $\mathcal{V}_{\text{base}}$ be subdivided into a set of already assigned nodes $\mathcal{L}$ and a set of so far unassigned nodes $\mathcal{U}$. Initially, $\mathcal{L}$ consists of the substations s_k, $k \in \{1, \dots, K\}$, only. In each of the following iterations, a node from $\mathcal{U}$, neighbouring the current $\mathcal{L}$, is assigned to one substation, until no nodes are left in $\mathcal{U}$.

Let the tuple $(a_v)_{v \in V}$ encode the current assignments, i.e., $a_v \in \{1, \dots, K\}$ if $v \in \mathcal{L}$ and $a_v = 0$ if $v \in \mathcal{U}$. For a node $v \in \mathcal{U}$ the weight w_v^k, $k \in \{1, \dots, K\}$ is defined as

$$w_v^k = \sum_{\{u \in \mathcal{N}(v) | a_u = k\}} {}^{1}\!/_{d(v,u)}$$

where $\mathcal{N}(v)$ are the neighbouring nodes of v in $\mathcal{G}_{base}$ and $d(v, u)$ the physical distance between node v and u along edges in $\mathcal{G}_{\text{base}}$. v is assigned to the substation s_k with the largest weight w_v^k.

Utilizing this mechanism, that is adapted from [12], often results in only one non-zero weight, i.e., when v has only one labelled neighbour in $\mathcal{L}$. However, if several neighbours are already labelled, the closer label is preferred.

To generate diverse possible node assignments, a stochastic element is added into the growth model. Specifically, the next node $v \in \mathcal{U}$ to assign to a substation is selected by randomly sampling from a pool of candidates. The candidates comprise the current unlabelled nodes with at least one neighbour already assigned to a substation.

Deriving the grid topology

Given a node assignment $(a_v)_{v \in V}$, the final grid topology $\mathcal{G}$ using minimum spanning trees is derived.

Let $\mathcal{G}_{\text{base},k} = \left(\mathcal{V}_{base,k}, \mathcal{E}_{base,k}\right)$ be a disjoint sub-graph of $\mathcal{G}_{\text{base}}$ defined by $\mathcal{V}_{base,k} = \{v \in \mathcal{V}_{\text{base}} | a_v = k\}$ and $\mathcal{E}_{base,k} = \{(v, u) \in \mathcal{E}_{base} | v \in \mathcal{V}_{base,k} \land u \in \mathcal{V}_{base,k}\}$. To derive the desired sub-grid $\mathcal{T}_k$, a minimum spanning tree on $\mathcal{G}_{\text{base},k}$ is calculated and all edges not leading to a terminal, i.e., a substation or an end-consumer node, are removed.

Rating the probability of the grid topologies

The grid topology generation procedure described above is repeated N times to generate an ensemble of N possible grid topologies.

The probability p_i of a generated grid topology $\mathcal{G}_i$, $i \in \{1, \dots, N\}$, is assumed to depend on the grid's investment costs, that in turn are assumed proportional to the aggregated length of the grid's electrical lines $l_{\mathcal{G}_i}$. Specifically, p_i is calculated as

$$p_i = \frac{e^{-\lambda l_{\mathcal{G}_i}}}{\sum_{j=1}^{N} e^{-\lambda l_{\mathcal{G}_j}}}$$

where λ represents the sensitivity of p_i on $l_{\mathcal{G}_i}$ and encodes domain knowledge. The latter is defined via user-defined parameters z and q as

$$\lambda = -\frac{1}{z} \ln\left(\frac{1}{1-q}\right),$$

where z represents an increase in the length of the lines (e.g., 10%) and q the corresponding reduction in probability (e.g., 50%).

CASE STUDY: OVERLOAD ANALYSIS OF DISTRIBUTION SUBSTATIONS

The following case study demonstrates that relying on a single, best estimate of the grid topology (e.g., the most likely grid topology in the ensemble) can mislead grid planners regarding the feasibility of a distribution grid, whereas the proposed ensemble-based framework leads to a more robust interpretation. To this end, a low-voltage distribution grid subject to 5% penetration of level-2 EV home-chargers is examined.

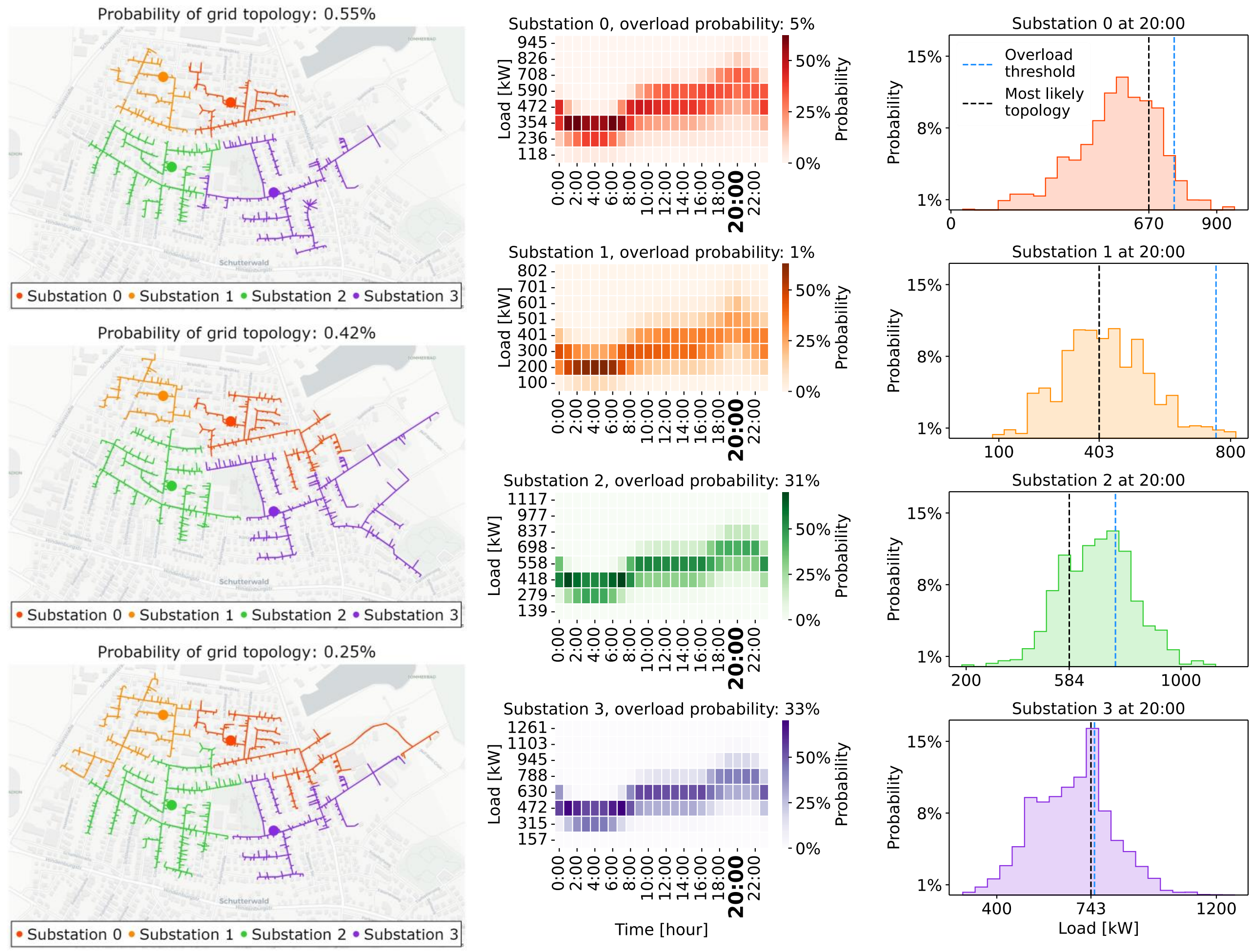

Figure 4: (Left column) Three exemplary grid topologies and their probability rating for four substations in the region considered in the case study. (Middle column) Hourly analysis of the load probability over a one-day period for the four substations. (Right column) Comparison of the load of the most likely grid topology with the load probability distribution over the ensemble for the four substations at an hour of peak load (i.e., 20:00).

Settings

A region located in Schutterwald, Germany, is considered with 15 distribution substations of capacity 630 kVA and in total 1751 households. The location of the distribution substations is obtained from [13], the households are extracted from the open-source geographic database OpenStreetMap [14].

Load profiles with hourly resolution $P_h^d(t), t \in \{1, \dots, 24\}$ over a one-day period are derived for all households $h \in \{1, \dots, H\}$ as follows. Existing load profiles obtained from [15] are randomly selected, averaged for a certain hour of the day t, and normalized w.r.t. to the total energy consumption. The resulting time-series for household h is then scaled with a daily energy consumption factor, that is proportional to the area of the household's shape extracted from OpenStreetMap.

Level-2 EV home-chargers are added to 5% randomly selected households. For a charger in household h a year-long charging profile from [16] is selected at random and the charging probability for each hour of the day t is derived from it. Then, 11 kW are added to $P_h^d(t)$ according to the determined charging probability.

Application

The ensemble of grid topologies generated is used to investigate overloads of the distribution substations in the modelled region. The overload of a distribution substation s_k is assumed to occur when s_k operates at or above 120% of its nameplate capacity $P_{s_k,\mathrm{np}}$ for at least one hour. For a substation s_k, the overload probability $p_{s_k,\mathrm{overload}}$ is calculated over the ensemble of N grid topologies as

$$p_{s_k,\mathrm{overload}} = \sum_{i=1}^{N} p_i * \max_t \mathbb{1}_{P_{s_k}^d(t) \geq 1.2 P_{s_k,\mathrm{np}}}$$

where $P_{s_k}^d(t)$ is the aggregated load profiles for all households assigned to the substation s_k and $\mathbb{1}$ the usual indicator function. To calculate $p_{s_k,\mathrm{overload}}$ for the substations in the modelled region, an ensemble comprising $N = 500$ grid topologies is generated.

Results

Three exemplary generated grid topologies along with their probability rating are shown in the left column of Fig. 4 for four of the 15 available substations.

To provide a comprehensive overview regarding the overload probability of the substations, the middle column of Fig. 4 analyses the hourly load probability over a one-day period. Substations with higher $p_{s_k,\mathrm{overload}}$ experience, particularly over the hours of peak load (e.g., $t = 20$), an increased load probability around and above the assumed overload threshold of 756 kW.

The right column of Fig. 4 demonstrates that relying on a single best estimate of the grid topology can mislead the grid planner regarding the feasibility of the modelled distribution grid. To this end, the load of the most likely grid topology for a designated hour, i.e., the hour of the peak load, $t = 20$, is compared with the load probability distribution over the ensemble of generated grid topologies. Specifically, the resulting peak load does not exceed the overload threshold of the four examined substations considering the most likely grid topology only. However, there are other possible grid topologies in the ensemble where the resulting peak load exceeds the overload threshold of the substations. This should be considered in a robust planning procedure when only imperfect grid knowledge is available.

CONCLUSION

In this work, a novel approach is presented to generate an ensemble of different, plausible grid topologies using the information about the distribution substations and the street layout as the only inherent attributes. The advantage of an ensemble pays off when evaluating the feasibility of the distribution grid. It is demonstrated how relying on the most likely estimate of the grid topology can mislead grid planners when imperfect grid knowledge is available, whereas the ensemble-based framework provides a more robust interpretation of the grid feasibility.

The proposed approach explicitly represents the limits implied by the scarcity of the available information by modelling the probability of the generated samples using domain knowledge. The domain knowledge is encoded via two user-defined parameters, and investment costs are used as indicator of the plausibility of a grid topology. Moving forward, it would be interesting to investigate additional design criteria to rate the probability of the generated samples.

There are several, other avenues of future work. For example, the grid topology generation mechanism can be expanded to the medium-voltage level. In general, there is significant potential in the ensemble-based grid topology generation, especially if one aims to robustly quantify the impact that future penetration levels of distributed RES, EV home-chargers and heat pumps can have on current power distribution systems.

REFERENCES

[1] Y. Liao, Y. Weng, G. Liu, Z. Zaho, C.W. Tau, R. Rajagopal, 2019, "Unbalanced multi-phase distribution grid topology estimation and bus phase identification", *IET Smart grid* vol. 2, 557-570.

[2] H.K. Çakmak, L.Janecke, V. Hegenmeyer, 2022, "Automated generation of large-scale distribution grid models based on open data and open source software using an optimization approach", *arXiv preprint arXiv:2202*.

[3] C. Mateo, F. Postigo, F. de Cudra, T.G. San Roman, T. Elgindy, P. Dueñas, B.M. Hodge, V. Krishnan, B. Palmentier, 2020, "Building Large-Scale U.S. Synthetic Electric Distribution System Models", *IEEE Transactions on Power Systems* vol. 26, 190-197.

[4] P. Nasirifard, J. Rivera, Q. Zhou, K.B. Schreiber, H. Jacobsen, 2018, "A crowdsourcing approach for the inference of distribution grids", *Proceedings of the Ninth International Conference on Future Energy Systems*, 187-199.

[5] Y. Liao, Y. Weng, M. Wu, R. Rajagopal, 2015, "Distribution Grid Topology Reconstruction: An Information Theoretic Approach", *2015 North American Power Symposium (NAPS)*, IEEE, 1-6.

[6] Y. Liao, Y. Weng, G. Liu, R. Rajagopal, 2019, "Urban MV and LV Distribution Grid Topology Estimation via Group Lasso", *IEEE Transactions on Power Systems* vol. 34, 12-27.

[7] Z. Wang, R.J. Thomas, A. Scaglione., 2008, "Generating Random Topology Power Grids", *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*, IEEE, 183-183.

[8] Z. Wang, A. Scaglione, R.J. Thomas, 2010, "Generating Statistically Correct Random Topologies for Testing Smart Grid Communication and Control Networks", *IEEE Transactions on Power Systems* vol. 1, 28-39.

[9] P. Schultz, J. Heitzig, J. Kurths, 2019, "A Random Growth Model for Power Grids and Other Spatially Embedded Infrastructure Networks", *The European Physical Journal Special Topics* vol. 12, 2593-2610.

[10] D. Deka, M. Cherrkov, S. Backhaus, 2020, "Topology Estimation Using Graphical Models in Multi-Phase Power Distribution Grids", *IEEE Transactions on Power Systems* vol. 35, 1663-1673.

[11] T. Kovač, 2019, "Translating CIM model to bus-branch model", *2019 27th Telecommunications Forum (TELFOR)*, IEEE, 1-4.

[12] S.A. Macskassy, F. Provost, 2003, "A Simple Relational Classifier", *New York Univ Ny Stern School of Business*.

[13] J.M. Kisse, M. Braun, S. Letzgus, T.M. Kneiske, 2020, "A GIS-Based Planning Approach for Urban Power and Natural Gas Distribution Grids with Different Heat Pump Scenarios", *Energies* vol. 13, 4052.

[14] OpenStreetMap, 2022, Open-source geographic database, https://www.openstreetmap.org/.

[15] NREL, 2022, "End-Use Load Profiles for the U.S. Building Stock dataset for electric consumption of households", https://www.nrel.gov/buildings/end-use-load-profiles.html.

[16] M. Muratori, 2017, "Impact of uncoordinated plug-in electric vehicle charging on residential power demand - supplementary data", National Renewable Energy Laboratory, Golden, USA.