Inference of Distribution Grids Based on Crowdsourced Grid Data and Drone Imagery

Hans-Arno Jacobsen, *IEEE Fellow*, Pezhman Nasirifard,
Jose Rivera, *Member, IEEE*, and Prerona Ray Baruah

Abstract—Distribution System Operators (DSOs) face several challenges in managing comprehensive and up-to-date models of distribution grids. To address these problems, we propose a crowdsourcing framework for collecting grid devices. We also provide an inference approach for generating topological models of the distribution grids. Since distribution cables are often underground, we use spatial data analytics on the collected data in combination with other open data sources to infer the topology of the distribution grid. Additionally, to increase the quality of crowdsourced data, we propose a cost-effective approach for collecting and detecting grid elements in urban areas using commercial drones with an RGB camera. To evaluate our approach, we organized a crowdsourcing campaign to map and infer a district in Munich, Germany. The results are compared with the ground truth of the distribution system operator. Our results report a precision of up to 82% and a recall of up to 65% for the correctly crowdsourced grid devices. We also observe that the inferred models achieve a power length accuracy of 88% compared to the ground truth. We evaluated the detection of solar panels from aerial imagery by conducting field experiments, showing precision and recall levels of 68% and 69%, respectively.

Index Terms—Power grids, Power distribution, Distribution grid inference, Geographic information systems, Crowdsourcing, Solar panels, Aerial imagery, Drone, Image processing, Nonmaximum Suppression

1 Introduction

Over the past few years, the majority of the international community has grown increasingly committed to the reduction in greenhouse gases, especially CO2 emissions [1]. Since the electric power industry is responsible for producing a significant portion of the CO2 emissions [2], researchers and practitioners have proposed several approaches to address these issues by facilitating further integration of renewable resources such as solar energy and introducing new electrical devices such as electric vehicles and local storage units [3]. However, before implementing the proposed solutions, the practicality and stability of the approaches should be comprehensively evaluated based on the actual distribution grid models. Nevertheless, the majority of studies are based on standardized test feeders, such as the IEEE test feeders [4], PNNL feeders [5] and CIGRE test feeders [6], which are considerably simplified models and fail to reflect the complexity, geographic features, and limitations of real individual power grids. Although some distribution system operators (DSOs) maintain digitized models of their grids, the operators do not publicly publish the grid models due to security and legal reasons. Additionally, in several cases, the grid data, especially for distribution grids, are either incomplete or outdated, and information regarding solar power capacity and locations of solar panel installations at a granular geographical scale is scarce. Additionally, periodically collecting and updating the grid data are time consuming, intrusive, and a significant financial burden for operators [7].

In this paper, we introduce a nonintrusive crowdsourcing framework for collecting and inferring distribution grid models. The crowdsourcing approach considerably reduces the cost, effort, and time required for gathering grid data by distributing the data collection tasks among the crowd. Furthermore, to improve the quality of the collected grid data, we merge the crowdsourced grid data with the extracted distribution grid elements from free and publicly available OpenStreetMap (OSM) data [8]. Moreover, we propose an approach for inferring a distribution grid topological model for a particular region based on the position of grid devices and the spatial features of the area.

To capture, analyze, and model the medium- and low-voltage grid models, the power industry has been utilizing geographic information systems (GIS) for a long time. However, no utility manages to maintain the most up-to-date and exhaustive model of their grids [9]. The novelty of our work lies in implementing a complete crowdsourcing framework for gradually and consistently collecting valid and verified grid data, which we use for inferring accurate grid models. In this work, we also explain how to conduct a crowdsourcing campaign based on our framework. We use the result of the crowdsourcing campaign to evaluate the performance of the participants and confirm the practicality of crowdsourced grid data in comparison to official grid data provided by the distribution grid operator.

Accurate and complete data regarding solar panels in urban areas, especially the panels mounted on rooftops, have to be collected manually. Although crowdsourcing is a potentially practical approach for collecting such devices, it is challenging for the crowd on the ground to identify and capture the solar panels that are installed on the higher elevation of buildings rooftops. Therefore, we propose an approach for detecting solar panel installations in urban areas from aerial images obtained with an affordable commercial drone equipped with an RGB camera. Most

Hans-Arno Jacobsen, Pezhman Nasirifard, Jose Rivera, and Prerona Ray Baruah are with the Department of Computer Science, Technical University of Munich, D-85748, Garching, Germany. (`p.nasirifard@tum.de`)

previous works focused on the detection of solar panels installed in solar farms from drone and satellite imagery [10]. However, urban areas have more obstructions than solar farms, such as buildings and vegetation, which makes the recognition of solar panels more challenging.

We recognize the following contributions in this paper:

1) We introduce a crowdsourcing framework for collecting, verifying, storing, and publishing power grid elements.
2) We design and conduct a crowdsourcing campaign with several participants, where evaluation of the acquired data verifies crowdsourcing as a practical approach for collecting and maintaining the grid data.
3) We propose an inference approach for generating distribution grid models of an area based on the geographical position of power grid elements, the consumer endpoints, and the pathway's structure of the area.
4) We implement an approach for automatic detection of solar panels in urban areas from images collected by drones equipped with an RGB camera.

We structure the rest of the paper as follows: In Section 2, we review the previous works on crowdsourcing and inference approaches for the distribution grids that we build on. In Section 3, we introduce the platform we built for executing the crowdsourcing event. Then, we discuss our crowdsourcing framework and describe the insights and results derived from the conducted crowdsourcing campaign in Section 4, where we also explain the surveyed feedback of the participants and evaluate the quality of the crowdsourced grid data. In Section 5, we describe and evaluate the inference approach for generating distribution grid models, followed by Section 6, where we explain our approach for detecting solar panels from drone imagery. In Section 7, we discuss the results and limitations of our work. Finally, in Section 8, we provide concluding remarks and describe future works.

2 Related Work

Crowdsourcing is a time-efficient and cost-effective method for collecting detailed and highly accurate geographical data, including electrical grid elements [11], [12]. However, there has not been a coherent crowdsourcing approach for collecting electrical grid devices. A potential resource of crowdsourced power-related data is OSM, which uses a community approach to locate and map the physical structures in an area. As of December 2018, OSM contains approximately 18 million components marked with power-related tags all around the globe [13]. However, the majority of the power-related OSM data are transmission-level elements. Medium- and low-voltage grid devices are scarce in OSM. Furthermore, the OSM community often tags the power-related components with wrong values due to the lack of expert knowledge or just errors; e.g., some transformers are marked as cable cabinets. An approach for crowdsourcing grid data and integrating the collected data with OSM data is presented in [14]. We extend the previous works by improving the crowdsourcing approach by designing a robust crowdsourcing framework as a supplement to OSM grid data to take advantage of all available datasets and to record grid elements for inferring distribution models. Moreover, we provide an evaluation of the approach against a ground truth, which was missing in the literature.

The studies on inferring topological models of distribution grids are somewhat limited [15]. Although several studies propose approaches for inferring topological models of transmission grids based on complex network theories and some from OSM data [16]–[21], these approaches do not apply to distribution grids. The main reasons are that the distribution grid components and structures are inherently different, the number of devices in distribution grids is more extensive than transmission grids, and the structures tend to be more complicated than in transmission grids. These challenges make distribution networks much more difficult to map and to build accurate models.

Furthermore, in contrast to transmission grid elements, in many countries, a significant portion of distribution-level grid elements are underground; e.g., in Germany, 73 percent of the medium-voltage and 87 percent of low-voltage cables are buried [22]. As a result, locating the accurate geographical position of grid components and their characteristics can be very challenging and require several assumptions and background knowledge.

Nevertheless, some studies propose intrusive methods for inferring distribution grid topologies based on the interaction among grid devices. In [23], the authors propose an approach for decentralized inference of distribution grids based on communication among a set of autonomous intelligent agents on an overlay network. In [24], they estimate the grid topology by applying correlation analysis on the voltage amplitude measurements of grid endpoints. However, none of these approaches consider geographical characteristics, and they require detailed information about the grid and interaction with grid elements, which we cannot apply to crowdsourced grid data. One similar area to our inference challenge is planning optimal and cost-efficient distribution grid systems based on the location of expected consumers, where the applications of a genetic algorithm and graph theory are accepted methods [25], [26]. Therefore, in this work, we use spatial analysis and graph theory for inferring distribution models when the location of the grid devices, consumers, and structure of the region is known. We differentiate ourselves from previous works in the combination of crowdsourcing and inference approaches to infer real power distribution grids. Moreover, our method is nonintrusive; i.e., we only require device locations and no grid measurements. Finally, in contrast to other works, we conduct field experiments and compare our results to the ground truth of the distribution grid operator.

Detection of solar panels from aerial imagery is often limited to satellite images [27]–[30] and analysis of solar farms with the aid of thermal cameras [31]. In [10], the authors proposed an approach for detecting solar panels using drones equipped with an infrared camera by applying image processing techniques such as thermal-based thresholding, contour detection, and morphological closing. However, this method does not apply to urban areas due to obstructions. Additionally, thermal cameras are expensive, making this approach not cost-effective. The authors of [27], [28] proposed an approach for detecting solar photovoltaic arrays from satellite images using random forest classifiers. The authors used the dataset of an American city in which the resolution of images was limited to 30 cm. The same dataset was used by the authors of [29] to detect solar arrays using a deep convolutional neural network (CNN), which yielded better results than the random forest approach. The authors of [30] used a dataset of annotated satellite images from Google Earth to estimate the electrical capacity of small-scale solar panels. In [32], the authors introduced a method for detecting regions with solar panels from satellite images, using an SVM classifier to classify the regions based on their unique features. However, their approach does not count the number of solar panels; hence, it is not able to estimate the solar power production.

The detection of individual solar panels from satellite images is a challenging task due to the low resolution of images. The highest ground resolution available currently is 31 cm from the *Worldview-4 Satellite* [33], which refers to the pitch or length of the side of a pixel. Thus, a constant error of 31 cm is always present. Given that a solar panel is made up of 5 pixels in an image, the calculated length of the panel may have a least count error of 31 cm (20%), which would significantly negatively affect results. In contrast, the pitch of a pixel of the images captured by the used drone at an altitude of 50 meters is approximately 2.5 cm [34]. Hence, the area covered by 1 pixel is much smaller. A smaller pitch leads to higher accuracy in the detection of small features, such as solar panels [35]. In our approach, we use widely available and cost-effective RGB cameras, which capture 4K images, enabling us to detect solar panels in urban areas.

3 OpenGridMap Crowdsourcing Platform

Our crowdsourcing campaign heavily makes use of the OpenGridMap (OGM) project [36]. The OGM project offers a platform for collecting, organizing, and openly publishing a broad range of transmission and distribution grid data and models. Additionally, OGM provides researchers and practitioners with a crowdsourcing platform for collecting high-, medium-, and low-voltage-level grid devices. OGM extracts and combines the power-related grid data of OSM with the verified submissions of volunteers. For example, electrical utility crews can use the OGM platform to record the continuously changing electrical grid to maintain the most recent information of the grid.

The OGM crowdsourcing platform consists of two primary components, including a smartphone application and a web application. The OGM Android application is available free of charge on the Google PlayStore [37]. The participants of the crowdsourcing activity are required to download and install this application on their phone. Afterward, the participants follow a simple procedure for submitting the grid element upon identifying the element in their surroundings. First, the participants should select the type of discovered grid device. Then, the participants take a picture of the grid device with the application and review the location of the device on the map obtained from the location service of the smartphone. Since the location service may not be accurate, the participants can edit the location of the grid device manually. Finally, participants submit the recorded grid element to the OGM servers either immediately or when the smartphone has access to a WiFi connection. On the OGM web application, the expert in the loop reviews and accepts the submitted grid elements. The expert has the option of examining the grid element; correcting the assigned metadata, such as the type of device; and merging the device with existing devices to avoid duplicate entries.

4 Crowdsourcing Distribution Grid Data

In the following section, we describe the designed and developed crowdsourcing framework for collecting distribution grid devices. We also performed a crowdsourcing campaign in the Freimann district of Munich, Germany. A comparison with the ground truth reveals that crowdsourcing is a practical method for mapping distribution grid.

4.1 Crowdsourcing Framework

Crowdsourcing provides an appealing option for collecting distribution grid data because the grid devices are widely distributed and their positions are previously unknown [38]. Nevertheless, we require a consistent framework to clarify and divide up the data collection activity into smaller, precise, and realizable tasks that several participants can accomplish independently and in parallel. We base our crowdsourcing framework on the collective intelligence framework developed by Malone et al. at MIT's Center for Collective Intelligence [38]. Malone's framework consists of four elements, also known as "genes", that are required for recognizing the building blocks of collective intelligence. The four elements are described as the four fundamental questions of "Who, Why, What and How", which we utilize to structure our crowdsourcing method.

The detailed description of Malone's framework is out of the scope of this paper. According to Malone's framework, we identify the following requirements and attributes of an organized crowdsourcing campaign for collecting distribution grid devices:

- We require the crowdsourcing movement to be relevant anywhere on the planet, independent of the geographical location and participants' previous knowledge and training. Therefore, we utilize a custom-designed smartphone application with a simplified data collection procedure as a medium for collecting the grid data. Due to the high prevalence of smartphones, the application can be used by any participant without requiring any previous training.
- Since any practical crowdsourcing activity depends on the number of participants, providing an appropriate incentive for the participants is crucial. Therefore, we present the participants with a detailed description of the project and its objectives, emphasizing the project's potential for reducing greenhouse gases and integrating renewable energy resources, presumably increasing the intrinsic joy of the participants in engaging in such a community.
- The distribution grid devices have a wide range of designs and types. However, we require collecting only specific devices such as transformers and cable

cabinets. Therefore, we provide the participants with a protocol, explicitly defining the type of necessary grid elements and the tasks that they need to fulfill. Because we presume that the participants do not have any previous training and they may detect incorrect grid devices, we verify the validity of the collected grid elements after the crowdsourcing campaign with the help of an expert in the loop.
- To break down the crowdsourcing activity into recognizable smaller tasks, we divide the data collection region into smaller subareas, and we assign each subarea to a group of participants. However, to prevent any potential duplicate recordings, the expert in the loop monitors the elements based on their position and removes the duplicated entries.

Accordingly, any crowdsourcing event that we manage follows a precise procedure. First, before beginning the crowdsourcing event, we hold a preliminary meeting with all participants to describe the objectives of the event and build a group of two persons. To increase the community spirit of participants and incentivize them better, we explain the objectives and benefits of their contributions clearly and discuss how their contribution could potentially help researchers with finding new solutions for challenges facing the electrical grids, namely, integrating renewable energy resources that could lead to addressing climate change.

Each group receives a package containing the group protocol and a stamped letter describing the intentions of the crowdsourcing campaign that they present to the police or other security personnel in case of any inquiry. We also provide an agenda that uses visual examples to explain the conventional design, signs, and characteristics of the desired grid devices specific to the area.

In each group, one person has the role of navigator, and the other serves as a data collector. The group navigator has the responsibility of filling the group's protocol and navigating the group through the area during the crowdsourcing event. The collector follows the navigator and carefully monitors the area for the requested power devices, and upon identifying a new grid device, the collector snapshots and submits the element by using the OGM smartphone application. The group protocol, completed by the navigator, contains a printed map of the area assigned to the group. The navigator marks the streets and paths that are covered by the group. Furthermore, in the case of failing to inspect some parts of the area, due to time limitations or lack of accessibility, the navigator marks the missing parts on the protocol's map and documents the reason. To prevent losing any collected grid data due to any unexpected OGM platform failures, the navigator keeps a list of discovered transformers on the protocol's map. Additionally, the navigator keeps a record of the number of identified transformers and cable cabinets in the protocol.

To increase the safety of the participants, we introduce a few obligatory rules that all participants are required to follow. We prohibit participants from trespassing in any military, industrial, or private properties. Only power devices that are observable from the streets and public areas should be recorded. Furthermore, we provide the participants with phone numbers of the event's organizers, whom they can reach in case of an emergency.

4.2 Crowdsourcing Campaign in Munich Freimann

To measure the quality of our framework, we organized and conducted a crowdsourcing campaign in the German city of Munich's Freimann district on the 9th of May 2017 [39]. We selected the Freimann district because we received the official distribution grid data from *Stadtwerke München* (SWM) [40], the Munich city utilities, and DSO, which we use as the basis for our evaluation.

Initially, we divided up the Freimann district into several subareas with approximately equal areas. However, since we recruited only 22 participants for the crowdsourcing event, we covered eleven areas with two persons assigned to each area, and we gave each group 90 minutes to perform the mapping in the designated area. Although we did not record the exact distance each group traversed, we intended each group to cover 4 kilometres of routes on average. Furthermore, since participants installed the application on their phones, their participation did not enforce any initial cost on us. Although the OGM crowdsourcing platform is capable of storing any distribution of grid devices, for the sake of simplicity, we asked the participants to collect only transformers and cable cabinets in their area. In the agenda, we provided detailed information about the transformers and cables cabinets used by SWM, Munich's DSO. We should mention that after the end of the crowdsourcing event, we offered an incentive (in the form of gift cards) to the best performing group.

After the event, we used a questionnaire to survey the overall crowdsourcing experience of all participants. We include the complete result of the survey in the appendix[1] and briefly explain the most interesting responses. In general, the results are in the affirmative upper third in all indicators, confirming that the participants are satisfied with the organization and execution of the OGM crowdsourcing platform and our framework. Furthermore, for the significant majority of the participants, the community spirit is a more valuable incentive than monetary prizes when deciding to join the event, indicating that our community-spirit-oriented incentives were attractive to the participants.

Although 75% of the participants expressed willingness to participate in such a crowdsourcing event again, we must remark that recruiting a large number of participants is challenging. We recruited 70% percent of the participants from the course instructed by one of the event organizers, and their friends informed the rest of the participants; no participant discovered the event from the public Facebook event we had created two months before the event. Therefore, we suggest investing enough time and publicity for gathering participants before any crowdsourcing event.

Finally, the survey reveals that identifying distribution grid devices is a challenging task because the grid devices are well hidden. However, since we only conducted the crowdsourcing event in Munich, we cannot argue that this difficulty extends to other urban or rural areas on the planet. The effort and time required for mapping an area depends on the complexity and density of the area and located grid devices more than it does on the size of the area.

1. The appendix is included as supplemental material.

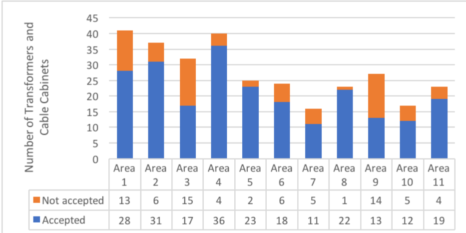


Fig. 1: The number of accepted and rejected grid devices in each area.

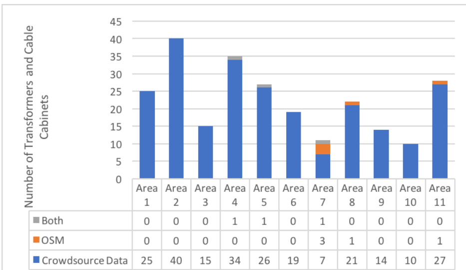


Fig. 2: OGM grid data separated based on their origin.

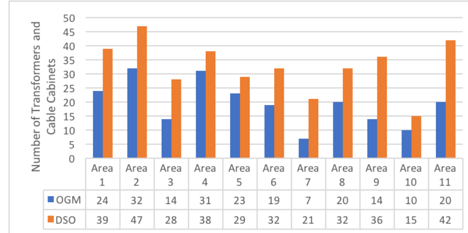


Fig. 3: The number of transformers and cable cabinets from the OGM and DSO datasets.

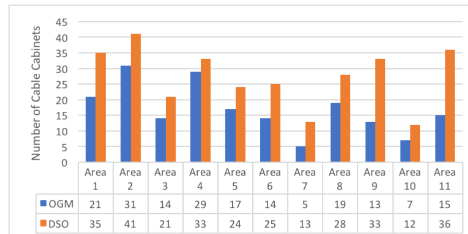


Fig. 4: The number of cable cabinets from the OGM and DSO datasets.

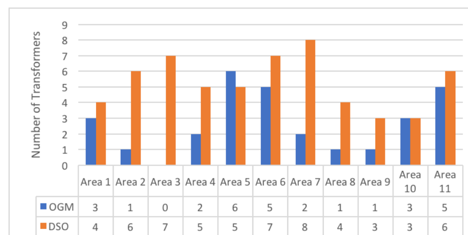


Fig. 5: The number of transformers from the OGM and DSO datasets.

4.3 Evaluation of the Crowdsourced Grid Data

After the crowdsourcing event, our experts in the loop use the OGM web application to verify the submitted devices. The grid elements that are not classified correctly, such as telecommunication cabinets that are incorrectly marked as cable cabinets, are removed from the dataset. The results show that, on average, 75% of the collected devices are correctly identified by the participants and that Area 8 has the best result with 96% accuracy because Group 8 strictly adhered to the agenda and captured only devices with detectable SWM signs. Figure 1 summarizes the number of correctly verified devices, including the cable cabinets and transformers combined, and the number of rejected devices from all eleven areas.

After the verification, we use the OGM platform to merge the grid data collected during the crowdsourcing event with the existing OGM grid data. As mentioned in the previous section, the OGM grid data consists of extracted power-related OSM data combined with a few potential submissions by other volunteers since the beginning of the project. During merging, we removed the duplicate submissions of the same grid devices that have identical type and position. We carry out the rest of the evaluation by using the merged grid data, which we refer to as OGM grid data. The reasons for using the merged data are that one of the contributions of the crowdsourcing approach is its function as a supplement to the extracted OSM grid data and that it is more crucial to evaluate the quality of aggregated publicly available grid data. However, at the time of conducting the crowdsourcing event, hardly any of the OGM transformers or cable cabinets originated from OSM or any volunteers, so the crowdsourcing event created a significant majority of the grid elements shown in Figure 2.

To evaluate the accuracy and validity of the collected grid data, we compare the OGM grid data from Freimann subareas one to eleven with the official DSO grid data in the corresponding subareas that we acquired from SWM. First, we examine the DSO grid elements and discover that for a few transformers, there are multiple identical entries that are overlapping on the map. Therefore, we merge the overlapping DSO transformers into one before evaluating the OGM grid elements. Then, we compare the number of OGM grid elements with the number of DSO grid elements in each area without any constraints on the distance, meaning that we do not enforce any maximum distance threshold between the exact geographical position of the OGM grid element and its corresponding DSO grid device. On average, the OGM grid data contain 60% of the DSO grid elements, and Area 4 reports the best accuracy of 82%, as shown by Figure 3, which summarizes and compares the number of extracted OGM and DSO transformers and cable cabinets. Furthermore, in more detail, our comparison reports a 61% coverage of DSO cable cabinets, with Area 4 offering the best precision of 88%. The comparison of OGM transformers to DSO transformers shows a 50% coverage, with Area 5 and 10 having 100% accuracy. Figure 4 and Figure 5 display the comparison between OGM and DSO grid data in the numbers of cable cabinets and transformers, respectively.

For the second phase of the evaluation, we define an eight-meter maximum distance threshold between the OGM grid element and the corresponding DSO element. As an

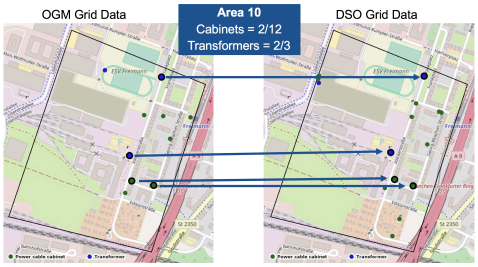


Fig. 6: OGM and DSO grid devices in Freimann subarea 10.

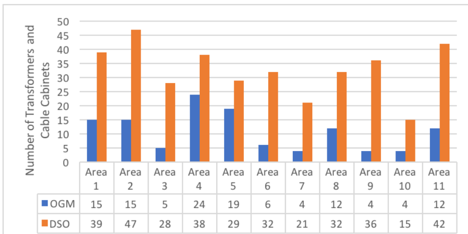


Fig. 7: The number of transformers and cable cabinets from the OGM and DSO datasets based on an eight-meter distance threshold.

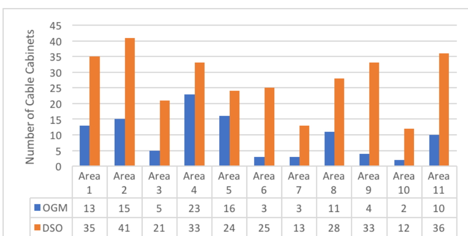


Fig. 8: The number of cable cabinets from the OGM and DSO datasets based on an eight-meter distance threshold.

example, Figure 6 displays the OGM and DSO grid elements on the map, where we only count OGM elements as valid if there exists an identical DSO element with the same type within an eight-meter proximity of the OGM element. With this constraint, we observe that on average, 33% of DSO elements are covered by the OGM dataset, where Area 5 has the highest coverage of 66%, as shown by Figure 7, which compares the numbers of OGM and DSO transformers and cable cabinets in each area. In more detail, the results report a 35% coverage of DSO cable cabinets, with Area 4 showing the highest coverage of 70%. The results also show a 26% coverage of DSO transformers, where Area 10 reports the best accuracy of 67%. Figure 8 and Figure 9 illustrate the number of recognized cable cabinets and transformers.

Given the fact that we mapped a large area of Freimann district in 90 minutes with the participation of 22 people and collected 230 valid transformers and cable cabinets, we argue that crowdsourcing is a practical approach for fast and cost-efficient data collection. Although none of the participants had expert knowledge, our experts in the loop verified the correctness of 75% of the collected devices, and the high precision (the number of true positives over the number of true positives and false positives) and recall (the number of true positives over the number of true positives

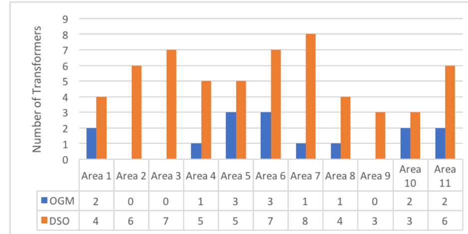


Fig. 9: The number of transformers from the OGM and DSO datasets based on an eight-meter distance threshold.

TABLE 1: Precision and Recall of Area 5

Area 5	Precision	Recall
Transformers + cable cabinets	82.61%	65.52%
Transformers	94.12%	66.67%
Cable cabinets	50%	60%

and false negatives) rates of Area 5 (the group with the best result) [41], as Table 1 displays, confirm the practicality of crowd intelligence. However, the participant's training and motivation are necessary for consistent performance. Nevertheless, the limited precision of the location services of the smartphones and the difficulty of detecting and locating well-hidden grid devices are the primary sources of inaccuracy, as shown by the varying accuracy of different areas and groups.

5 Inference of the Distribution Grid

In the following section, we introduce our approach for inferring distribution grid models based on the geographical position of grid devices and consumers' endpoints. Furthermore, we discuss the ground truth, which we derive from the official DSO grid data that we received from SWM, and we finally evaluate the accuracy of two inferred models compared to the ground-truth model.

5.1 Distribution Network Inference based on Grid Data

We propose an approach for inferring distribution grid models of an area based on the grid data of the area. In other words, we infer a model of the low-voltage underground power cables located within a specific area based on the position of transformers, cable cabinets, and consumer endpoints such as buildings and the structure of pathways in the area. We base our inference method on the primary assumption that the majority of underground cables are installed along roads and pathways. Therefore, we take advantage of the complete and freely available road information of the area from OSM. Inference algorithm 1 denotes the method we develop for heuristically inferring a minimum spanning tree (MST) as the distribution grid model. We consider the provided position of grid devices and consumers as target nodes taking into account the structure of the roads. Each road consists of several nodes representing a line, which we simplify to enforce the inference of MST along the roads. Therefore, we select the start and endpoints of the roads and the intersection point of each road pair. To avoid including duplicate points in our dataset, we use a set ensuring the existence of only one copy of any point.

Fig. 10: Projection of the target node on the nearest roads with an edge between them.

Afterward, we determine the projection of the target nodes on the nearest road. As an example, Figure 10 displays target nodes in blue dots, their projection as green nodes, and the edge between target and projection nodes as red lines. We merge the projected nodes and filtered road nodes to create a base graph containing edges between every pair of nodes in the union of the two sets, where we consider the distance between nodes in meters as the weight of the edge. Then, we use the base graph to infer an MST, and we finally add the edges between target nodes and projected nodes to the inferred MST and return the tree as the distribution grid model. We use projected nodes instead of target nodes for creating the base graph because otherwise, we could not create a clean MST along the roads.

Algorithm 1: Distribution Grid Inference Approach

```
InferDistributionGrid (TargetNodes, Roads)
    input : TargetNodes, the set of transformers, cabinets and
            buildings.
    input : Roads, the line geometry of roads.
    output: GridModel, A minimum spanning tree representing the
            distribution grid.
    filteredRoadNodes = set()
    foreach road_i ∈ Roads do
        foreach road_j ∈ Roads do
            ;                                  // Roads start points
            filteredRoadNodes.add(road_i[0], road_j[0])
            ;                                  // Roads end points
            filteredRoadNodes.add(road_i[−1], road_j[−1])
            if roadsAreIntersecting(road_i, road_j) then
                intersectionPoint =
                  findIntersectionPoint(road_i, road_j)
                filteredRoadNodes.add(intersectionPoint)
    projectedNodeSet =
      projectTargetNodesOnRoad(TargetNodes, Roads)
    mergedNodeSet = filteredRoadNodes ∪ projectedNodeSet
    baseGraph = Graph()
    baseGraph.addNodes(mergedNodeSet)
    foreach node_i ∈ mergedNodeSet do
        foreach node_j ∈ mergedNodeSet do
            if edge(node_i, node_j) not in baseGraph AND i! = j
              then
                edgeWeight =
                  getDistanceInMeters(node_i, node_j)
                baseGraph.addEdge(node_i, node_j, edgeWeight)
    gridModel = generateMinimumSpanningTree(baseGraph)
    projectedEdges =
      makeEdgeProjectedTargetNode(TargetNodes)
    gridModel = gridModel ∪ projectedEdges
    return gridModel
```

The inference engine is implemented in Python with the help of several packages, including but not limited to NumPy [42], Shapely [43], and NetworkX [44], for creating and manipulating complex networks. We store our grid data on a PostgreSQL [45] database with the PostGIS [46] extension enabled. To visualize and inspect the data on the area's map, which we import from OSM, we also use the open-source geographical information system QGIS [47]. Furthermore, our source code is open sourced [48].

5.2 Ground-Truth Model of Freimann District

To evaluate the quality of the inference algorithm, first, we generate the ground truth of the distribution model of Freimann based on the official DSO grid data. We need to construct compatible ground-truth models because the acquired DSO grid information is in shapefile format (shape format, shape index format, and attribute format) [49], which is not compatible with our inferred grid models. Furthermore, the DSO data require cleaning due to some data inconsistencies, data duplication, and errors. After importing the DSO grid data of the Freimann district into our database, including the location of transformers, cable cabinets, consumer connections, and location of underground cables, we perform a few steps of data cleaning.

For some transformers, cable cabinets, and consumer nodes, there exist duplicate copies of nodes that are either overlapping or located within a few meters of each other. We merge these nodes by defining a maximum distance of one to five meters between them. The reason for using a varying range as a distance threshold is that we inspect the data visually to find the best threshold based on the type of the node. Afterward, we inspect the transformers, cable cabinets, and consumer endpoints to detect the ones that are disconnected from the grid due to the absence of a connection to any neighboring cables. We connect these isolated nodes by connecting the nodes to the closest cable located within a node's five-meter proximity, and if such a cable does not exist, the node is removed from the dataset. In the end, we review any remaining separated cable that is not connected to any node or another cable at any endpoints and remove the isolated cables from the dataset.

After cleaning and preprocessing, we use the cleaned data to create a graph representing the distribution grid of the area. However, we only use the cables from the DSO dataset since all cables are presumably connected to either endpoints or other cables. We build the base ground-truth graph by iterating through each DSO cable line and retrieving the geographical representation of the line from the database. Then, we convert the line's data into a set of nodes with an edge connecting the points that are in the row behind each other. Afterward, we select the largest connected subgraph as the model representing the distribution model of the area. We follow this approach for creating the ground-truth model because the cable line data are stored according to EPSG:31468 Projection [50], which requires conversion into the fundamental longitude and latitude that we use. As an example, Figure 11 displays the ground-truth graph of the Freimann subarea, in which the largest connected subgraph is identified with green edges and the smaller marked disconnected subgraphs are discarded. Figure 12 shows the created ground truth model of Freimann subareas one to five, with a total cable length of 46484 meters.

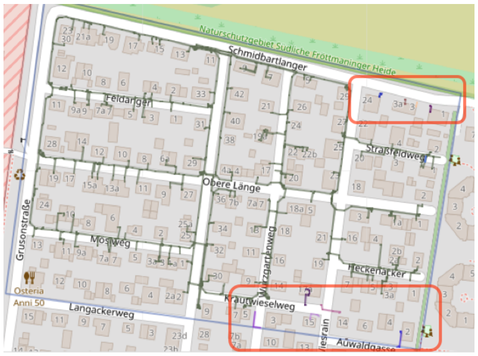


Fig. 11: Freimann subarea ground-truth subgraphs.

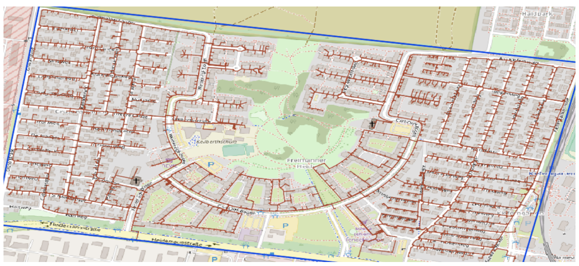


Fig. 12: Ground-truth model of the Freimann subarea.

5.3 Evaluation of the Inference Approach

To evaluate the accuracy of the inference algorithm, we infer two separate models based on the DSO and OGM transformers and cable cabinets data of Freimann and compare the models with the generated ground-truth model. Since our inference algorithm also requires the information of roads and consumer endpoints of the area, we integrate the grid data with the extracted related OSM data, including the structure of the roads and the position of residential and commercial buildings. The OGM grid data that we use for generating models are the verified grid devices, which we evaluated in Subsection 4.3. Figure 13 and Figure 14 illustrate the inferred models of the DSO and OGM grid data, respectively. We generate the ground-truth model and the DSO-inferred model, both of which are based on the acquired DSO data of the area. However, the difference is that the ground truth is created based on the structure of previously known underground cables, whereas the inferred DSO grid model generates the structure of the underground cables based on the position of transformers and cable cabinets.

The visual comparison of the models reveals high coverage of the ground-truth model with the inferred models. Figure 15 displays the two ground-truth and DSO-inferred models overlapping on the map. Furthermore, we compare the models based on the length of the models' inferred cables. Table 2 summarizes the calculated length of cable for each inferred model, reporting the 88% coverage of the ground-truth model by the DSO-inferred model and 75% coverage by the crowdsourced OGM grid model. The results indicate that the inference approach is capable of inferring accurate grid models. However, the lower availability of grid data in the OGM dataset than the DSO dataset indicates that the quality of the model heavily depends on the availability of distribution grid elements in the area.

Fig. 13: The DSO-inferred model of the Freimann subarea.

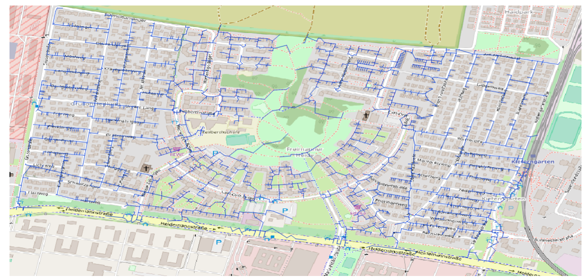


Fig. 14: The OGM-inferred model of the Freimann subarea.

Fig. 15: Overlapping ground-truth and DSO inferred-models. (Orange lines are ground truth, and the blue lines are DSO.)

6 Solar Panel Detection from Drone Imagery

To improve the quality of the inferred distribution models, we need to improve the quality and availability of grid data by integrating other resources, such as automatic detection of grid devices from aerial imagery. In this section, we propose a novel approach for detecting solar panels installed in urban areas from aerial images through an automated and efficient approach. We divide the procedure for detecting the solar panel into two separate phases. The first phase is flight planning and capturing images of the areas of interest with a drone. The second part involves preprocessing the images and detecting the solar panels.

TABLE 2: Comparison of Ground-Truth Model with DSO and OGM Inferred Models

Grid Model	Cable Length (m)	Coverage
DSO-Inferred Model	40840	87.86%
OGM-Inferred Model	34866	75.01%

6.1 Data Collection Strategy and Image Acquisition

We selected an area in Garching, Germany, with high availability and density of solar panels. We used a *Phantom 3 Professional* quadcopter from *DJI* [34] with a 4K RGB camera attached. We flew the camera 50 meters above the ground for 40 minutes in total. We used the *Pix4D* Capture app to execute the flight plans with a double-grid flight pattern [51]. Although this pattern is more time consuming, it provides maximum overlap and coverage, ensuring good quality orthophotos. The acquired images are high-quality RGB images.

6.2 Image Processing and Solar Panel Detection

The collected images are preprocessed using two different parameter sets to remove noise and improve contrast, resulting in two outputs for each input image. Then, the procedure is followed by edge and contour detection of the solar panels and the application of feature constraints to minimize false positives. The results are combined, and redundancies are removed using a nonmaximum suppression method to provide a higher number of correctly detected solar panels. In the following, we explain each step in more detail.

6.2.1 Preprocessing

We perform preprocessing to enhance the detectability of the solar panels. Two parameter sets are used for preprocessing the images: one for highlighting panels in different lighting conditions and orientations and the other for providing two outputs for each input at every step. Finally, the nonmaximum suppression method removes redundancies from the combined results. We apply the following preprocessing steps:

- **Thresholding**: For increasing the contrast of solar panels against the background, the blue channel of images is thresholded at 130 for the parameter set one. No thresholding is done for the parameter set two in this step.
- **Grayscaling**: This step converts each pixel to an intensity value in the range of 0 to 1 for reducing the amount of data used during image processing. The grayscaled images are thresholded again at a value of 140 for both parameter sets to further improve the contrast.
- **Gaussian Blurring**: We apply this step to smooth the grayscaled image and to remove the noise. For making the boundaries of the solar panel more significant and detectable, we apply two filter sizes of $\sigma = 3$ for parameter set one and $\sigma = 5$ for parameter set two.

6.2.2 Edge Detection

After removing the noise from the images, we use the Canny edge detection algorithm to detect the edges in the image for identifying the boundaries of objects by selecting pixels whose intensity values change suddenly.

6.2.3 Morphological Closing

Several detected edges are interrupted by small gaps between them, which hinder the detection of contours or closed polygons in the next step. For this reason, we perform a closing operation to close these small gaps.

Fig. 16: The result of contours detected (in red) using the first parameter set.

Fig. 17: The result of contours detected (in blue) using the second parameter set.

6.2.4 Contour Detection

Contours are derived from edges. To be able to define a contour, the detected edges must be a set of closed curves forming a closed polygon. For detecting solar panels, we identify the quadrilateral since the distortions in aerial images cause rectangles to appear as quadrilaterals.

6.2.5 Constraints

To eliminate the false positives from a large number of detected polygons during contouring, we apply the following four constraints:

- We limit the number of vertices for each polygon to four.
- Contours that have an approximate area outside the range of typical solar panels are discarded.
- We set the valid range of the interior angles of the polygons from 25 to 140 degrees.
- We define the valid range for the length of each of the segments of the polygons to (34, 55).

Figure 16 and Figure 17 show the result of this step for the parameter set one and two on an example figure.

6.2.6 Nonmaximum Suppression

In the last step, we combine the outputs of the two parameter sets and remove the redundancies by using the nonmaximum suppression method [52]. We modified the

Fig. 18: The result of nonmaximum suppression (in green) combining the detected contours.

algorithm to enable us to detect the polygons and tilted rectangles. This algorithm calculates the value of the local maxima in the image. All of the pixels that do not fall within the local maxima are set zero. Figure 18 shows the result of this step on the sample image.

6.3 Evaluation of the Solar Panel Detection

For evaluation, a set of ten input images and their outputs were taken into consideration, as Table 3 displays. For each image, we manually counted the number of available solar panels to quantify the efficiency of our detection. True positives refer to the number of correctly detected solar panels, false positives are the polygons falsely detected as solar panels, and false negatives are the solar panels that are not detected in the final output images. To quantify the quality of our approach, we use precision and recall. A high recall level is desirable, as it indicates a higher number of correctly detected solar panels.

Figure 19 shows the results for the total numbers of solar panels, true positives, false negatives, and false positives for parameter sets one and two and for the final output after combining the results and removing the redundancies with the nonmaximum suppression method. We see that the number of true positives increased significantly to 266 in the final output compared to 211 and 224 for parameter sets one and two, respectively. The number of false negatives decreased in the final output to 118. This result is desirable since we aim to correctly detect as many solar panels as possible. The results show that the recall increased significantly up to 0.69 in the final output compared to 0.57 and 0.60 for parameter sets one and two, respectively. This is a desirable result, as the aim is to maximize the correctly detected solar panels. The different parameters allow the detection of solar panels in different lighting conditions and orientations, which leads to an increase in recall. The precision decreased slightly to 0.68 since the combination of results led to a marginal increase in the number of outliers. This outlier increase can be rectified at a later stage during postprocessing.

TABLE 3: Calculation of Precision and Recall for a Sample of Images

Image	TP	Found Polygons	FP	FN	Precision	Recall
1	71	80	14	1	0.835	0.986
2	14	28	14	3	0.5	0.823
3	14	28	14	17	0.5	0.451
4	84	86	9	5	0.903	0.943
5	6	13	14	11	0.3	0.352
6	6	25	19	9	0.24	0.4
7	36	48	12	33	0.75	0.521
8	7	9	2	10	0.777	0.411
9	14	28	14	3	0.5	0.823
Total	252	345	112	92	0.68	0.69

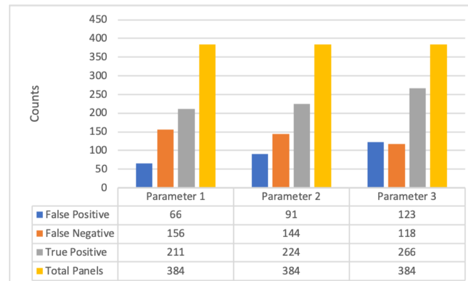


Fig. 19: Solar panel counts for the parameter set one, the parameter set two and for the final output after nonmaximum suppression.

7 Discussion and Limitations

The results of the campaign and inference approach demonstrate that crowdsourcing is a practical, fast, and cost-efficient approach for collecting grid data of an area that can be used for generating accurate distribution grid models. Although the inferred models can be used for simple academic simulation studies, they lack sufficient accuracy to be used by TSOs and DSOs for power infrastructure control purposes. To improve the quality of inferred distribution models, we need to improve the quality and the availability of crowdsourced grid data as well as the inference approach. To improve the quality of grid data, we need to increase the accuracy of collection devices, invest in training of participants and integrate grid data and models from various official resources. Although these approaches can be beneficial, they introduce new financial and legislative burdens. Furthermore, although we did not use solar panels for the inference of the grid in this work , the discussed panel detection approach can be adapted for use in detecting transformers since we could benefit from the low flying altitude of drones and the various techniques we employed for identifying objects and angles.

To improve the accuracy of the inference approach, we need to consider the exceptional inference cases. For example, the discussed inference algorithm does not determine a difference between the various types of roads and pathways and assumes the existence of underground cables along any paths, such as pedestrian ways and dirt roads, but in reality, the cables are often not installed along dirt roads. Furthermore, MSTs are a suboptimal solution when inferring distribution grid models because DSOs often implement loops in the distribution systems for increasing the resilience and reliability of the power grid. Therefore, we suggest investigating more complex network theory approaches or power flow-based network design approaches.

Furthermore, we should mention that although several OSM power-related elements contain useful information such as voltage level, several other essential grid characteristics, such as the line's thermal parameters, are missing, as acquiring such information requires expert knowledge of the local distribution grid. We limit ourselves in this work to inferring the topological model of the distribution grids.

Although we limited this work to infer distribution grid models of specific regions in Munich, the explained approach for collecting grid data and generating models is not limited to any geographical region. However, researchers may encounter a few challenges when applying our approach to other regions, especially areas outside Germany. One of the challenges is the different appearance of grid elements in various countries. For example, various regions may have different regulations for marking transformers or cable cabinets. Therefore, the contributors should first be instructed on the local characteristics of grid devices. Furthermore, researchers must also pay attention to the assumptions we considered for inferring models, such as whether the placement of underground cables along roads holds for the corresponding region.

8 Conclusions

In this work, we introduced a crowdsourcing framework developed with the help of the OGM platform, which we used for conducting a crowdsourcing campaign. We also proposed an inference approach for generating distribution grid models based on the position of grid devices, consumer endpoints, and roads. We also introduced an approach for detecting solar panels from drone imagery. The results of the crowdsourcing event achieve a precision of up to 82% and a recall of up to 65% depending on the performance of the participants. Furthermore, the evaluation of the inferred distribution grid model based on the official complete DSO grid dataset shows a power length accuracy of 88% compared to the ground truth. Additionally, the evaluation of our field experiments for solar panel detection revealed 68% precision and 69% recall. Our results confirm crowdsourcing as an efficient and beneficial data collection approach for distribution grid device mapping, which, in combination with an inference algorithm, can provide a practical method to obtain realistic distribution grid models.

As future work, we will work on improving the quality of the data collection methods and the inference approach by integrating additional data sources and methods, such as automatic detection of distribution grid elements and automatic improvement and classification of grid data by utilizing deep learning approaches. Furthermore, we intend to investigate the applicability of our crowdsourcing approach and inference algorithm to the North American distribution grid, where the grid elements are more spread out over larger geographical areas and the grid models are also less accurate than the German grid. To improve the precision of the inference algorithm on the crowdsourced data, we plan to consider the use of more complex graph structures, genetic algorithm-based methods, and spatial clustering approaches of smart meter data. Finally, we focus on adapting the techniques used for detecting solar panels to identify transformers, and we improve the accuracy of the solar panel detection mechanism by improving the constraint definitions to remove the false positives.

Acknowledgment

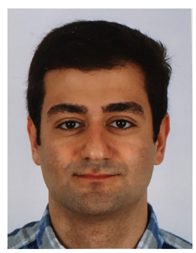
The authors would like to thank Dr. Carsten Sperl from Stadtwerke München for his contribution to the project. Most importantly, we would like to thank all the contributors that have helped crowdsource geographical data with the help of the OpenGridMap project. This research was supported by the German Federal Ministry of Education and Research Grant (BMBF 01IS12057) and the Alexander von Humboldt Foundation.

References

[1] U. N. T. Collection, “Paris agreement,” 2015.
[2] T. W. Bank, “Co2 emissions from electricity and heat production,” 2014, accessed: 2018-01-03.
[3] J. Rivera, C. Goebel, and H. A. Jacobsen, “Distributed convex optimization for electric vehicle aggregators,” *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1852–1863, 2017.
[4] W. Kersting, “Radial distribution test feeders,” *IEEE Transactions on Power Systems*, vol. 6, no. 3, pp. 975–985, 1991.
[5] P. N. N. Laborattories, “Modern grid initiative distribution taxonomy final report,” Tech. Rep., 2008.
[6] K. Strunz, R. Fletcher, R. Campbell, and F. Gao, “Developing benchmark models for low-voltage distribution feeders,” in *Power Energy Society General Meeting, 2009. PES ’09. IEEE*, 2009, pp. 1–3.
[7] Esri, “Enterprise gis and the smart electric grid,” Tech. Rep., 2012.
[8] OpenStreetMap, “Openstreetmap wiki,” 2017, accessed: 2018-01-03. [Online]. Available: https://wiki.openstreetmap.org/wiki/Main_Page
[9] B. Meehan, *Modeling Electric Distribution with GIS*. Esri Press Redlands, 2013.
[10] S. Dotenco, M. Dalsass, L. Winkler, T. Wrzner, C. Brabec, A. Maier, and F. Gallwitz, “Automatic detection and analysis of photovoltaic modules in aerial infrared imagery,” in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2016, pp. 1–9.
[11] C. Heipke, “Crowdsourcing geospatial data,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 6, pp. 550–557, 2010.
[12] G. Paolacci, J. Chandler, and P. G. Ipeirotis, “Running experiments on amazon mechanical turk,” 2010.
[13] OpenStreetMap, “Openstreetmap power tag info,” 2018, accessed: 2018-12-12. [Online]. Available: https://taginfo.openstreetmap.org/keys/power
[14] J. Rivera, C. Goebel, D. Sardari, and H. A. Jacobsen, *OpenGridMap: An Open Platform for Inferring Power Grids with Crowdsourced Data*. Springer International Publishing, 2015, pp. 179–191.
[15] P. Nasirifard, J. Rivera, Q. Zhou, K. B. Schreiber, and H. A. Jacobsen, “A crowdsourcing approach for the inference of distribution grids,” in *Proceedings of the Ninth International Conference on Future Energy Systems*. ACM, 2018, pp. 187–199.
[16] R. Albert, I. Albert, and G. L. Nakarado, “Structural vulnerability of the north american power grid,” *Phys. Rev. E*, vol. 69, 2004.
[17] B. Cloteaux, “Limits in modeling power grid topology,” *2013 IEEE 2nd Network Science Workshop (NSW)*, pp. 16–22, 2013.
[18] J. Rivera, J. Leimhofer, and H. A. Jacobsen, “Opengridmap: Towards automatic power grid simulation model generation from crowdsourced data,” *Computer Science - Research and Development*, vol. 32, no. 1, pp. 13–23, 2017.
[19] P. Hines, S. Blumsack, E. C. Sanchez, and C. Barrows, “The topological and electrical structure of power grids,” in *2010 43rd Hawaii International Conference on System Sciences*, 2010, pp. 1–10.
[20] V. Hagenmeyer, H. Kemal Çakmak, C. Düpmeier, T. Faulwasser, J. Isele, H. B. Keller, P. Kohlhepp, U. Kühnapfel, U. Stucky, S. Waczowicz, and R. Mikut, “Information and communication technology in energy lab 2.0: Smart energies system simulation and control center with an open-street-map-based power flow simulation example,” *Energy Technology*, vol. 4, no. 1, pp. 145–162, 2016.
[21] J. Rivera, P. Nasirifard, J. Leimhofer, and H. Jacobsen, “Automatic generation of real power transmission grid models from crowdsourced data,” *IEEE Transactions on Smart Grid*, 2018.
[22] S. von Roon, M. Sutter, F. Samweber, and K. Wachinger, “Netzausbau in deutschlan,” 2014, accessed: 2018-01-03. [Online]. Available: http://www.kas.de/wf/doc/kas_38837-544-1-30.pdf

[23] N. Honeth, A. Saleem, K. Zhu, L. Vanfretti, and L. Nordström, "Decentralized topology inference of electrical distribution networks," in *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*, 2012, pp. 1–8.
[24] S. Bolognani, N. Bof, D. Michelotti, R. Muraro, and L. Schenato, "Identification of power distribution network topology via voltage correlation analysis," in *52nd IEEE Conference on Decision and Control*, 2013, pp. 1659–1664.
[25] V. Kumar, R. Krishan, and Y. R. Sood, "Optimization of radial distribution networks using path search algorithm," *International Journal of Electronics and Electrical Engineering*, vol. 1, no. 3, pp. 182–187, 2013.
[26] C. G. Taroco, E. G. Carrano, R. H. C. Takahashi, and O. M. Neto, "A faster genetic algorithm for substation location and network design of power distribution systems," in *2012 IEEE Congress on Evolutionary Computation*, 2012.
[27] J. M. Malof, K. Bradbury, L. M. Collins, and R. G. Newell, "Automatic detection of solar photovoltaic arrays in high resolution aerial imagery," *Applied*, vol. 183, pp. 229 – 240, 2016.
[28] J. M. Malof, K. Bradbury, L. M. Collins, R. G. Newell, A. Serrano, H. Wu, and S. Keene, "Image features for pixel-wise detection of solar photovoltaic arrays in aerial imagery using a random forest classifier," in *2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA)*, 2016, pp. 799–803.
[29] J. M. Malof, L. M. Collins, and K. Bradbury, "A deep convolutional neural network, with pre-training, for solar photovoltaic array detection in aerial imagery," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 874–877.
[30] B. So, C. Nezin, V. Kaimal, S. Keene, L. Collins, K. Bradbury, and J. M. Malof, "Estimating the electricity generation capacity of solar photovoltaic arrays using only color aerial imagery," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 1603–1606.
[31] R. K. Fedorov, "Logical methods of object recognition on satellite images using spatial constraints," *arXiv preprint arXiv:1004.4793*, 2010.
[32] A. Devarajan, S. Manor, J. Perkins, R. Saboo, W. Zhang, K. Bradbury, L. Collins, T. Johnson, J. Malof, and R. Newell, "Solar power estimation through remote sensing," 2018, accessed: 2018-12-05. [Online]. Available: https://bigdata.duke.edu/sites/bigdata.duke.edu/files/images/SolarPoster_sm.pdf
[33] DigitalGlobe, "Satellite information," 2018, accessed: 2018-12-05. [Online]. Available: https://www.digitalglobe.com/resources/satellite-information
[34] DJI, "Phantom 3 4k specs," 2018, accessed: 2018-12-05. [Online]. Available: https://www.dji.com/phantom3-4k/info
[35] NanoLumens, "Pixel pitch," 2017, accessed: 2018-12-05. [Online]. Available: https://www.nanolumens.com/blog/what-is-pixel-pitch-and-why-should-i-care/
[36] OpenGridMap, "Opengridmap homepage," 2018, accessed: 2018-01-03. [Online]. Available: http://opengridmap.com/
[37] ——, "Opengridmap android application," 2018, accessed: 2018-01-03. [Online]. Available: https://play.google.com/store/apps/details?id=tanuj.opengridmap
[38] T. W. Malone, R. Laubacher, and C. Dellarocas, "Harnessing crowds: Mapping the genome of collective intelligence," 2009.
[39] C. of Munich, "Munich freimann district," 2018, accessed: 2018-01-05. [Online]. Available: http://www.muenchen.de/stadtteile/freimann.html
[40] S. München, "Munich city utilities," 2018, accessed: 2018-01-05. [Online]. Available: https://www.swm.de/
[41] K. M. Ting, *Precision and Recall*. Springer US, 2010, pp. 781–781.
[42] NumPy, "Numpy package," 2018, accessed: 2018-01-03. [Online]. Available: http://www.numpy.org/
[43] Shapely, "Shapely package," 2018, accessed: 2018-01-03. [Online]. Available: https://pypi.python.org/pypi/Shapely
[44] NetworkX, "Networkx package," 2018, accessed: 2018-01-03. [Online]. Available: https://networkx.github.io/
[45] P. G. D. Group, "Postgresql," 2018, accessed: 2018-01-10. [Online]. Available: https://www.postgresql.org/
[46] PostGIS, "Spatial and geographic objects for postgresql," 2018, accessed: 2018-01-05. [Online]. Available: https://postgis.net/
[47] Q. D. Team, "Qgis," 2018, accessed: 2018-01-04. [Online]. Available: https://www.qgis.org/en/site/
[48] OpenGridMap, "Opengridmap distribution inference engine," 2018, accessed: 2018-01-03. [Online]. Available: https://github.com/OpenGridMap/power-grid-detection
[49] E. ESRI, "Shapefile technical description," *An ESRI White Paper*, 1998.
[50] S. Reference, "Epsg:31468 projection," 2018, accessed: 2018-01-11. [Online]. Available: http://spatialreference.org/ref/epsg/31468/
[51] Pix4D, "Pix4dcapture," 2018, accessed: 2018-12-05. [Online]. Available: https://pix4d.com/product/pix4dcapture/
[52] A. Rosebrock, "Non-maximum suppression for object detection in python," 2014, accessed: 2018-12-05. [Online]. Available: https://www.pyimagesearch.com/2014/11/17/non-maximum-suppression-object-detection-python/

Hans-Arno Jacobsen's pioneering research lies at the interface between computer science, computer engineering, and information systems. After studying and completing his doctorate in Germany, France, and the USA, Prof. Jacobsen engaged in postdoctoral research at the INRIA near Paris before moving to the University of Toronto in 2001, as a professor in the Department of Electrical and Computer Engineering and the Department of Computer Science. After being awarded the prestigious Alexander von Humboldt Award, he started to engage in research activities at the Technical University of Munich.



Pezhman Nasirifard received his B.Sc. in Computer Software Engineering from the University of Tehran, Iran, in 2013 and his M.Sc. in Computer Science from the Technical University of Munich in 2017. He joined the Chair of Application and Middleware Systems at the Technical University of Munich in August 2017 as a Ph.D. candidate.

Jose Rivera received a Dipl.-Ing. degree in Electrical Engineering and Information Technology and a Ph.D. degree in Natural Sciences from the Technical University of Munich, Germany, in 2012 and 2017, respectively. Until his graduation, he interned at Siemens Corporate Technologies. He was also a visiting student at the Massachusetts Institute of Technology (MIT) in the Laboratory for Information and Decision Systems (LIDS).

Prerona Ray Baruah received a B.Tech. Degree in Computer Science and Engineering from Indraprastha University, India, in 2014 and an M.Sc. degree in Computer Science from the Technical University of Munich, Germany, in 2017. Her interests include image processing, database systems, and computer networks. She is currently working as a software developer at the SAP Headquarters, Germany.