# GENERATING LOW VOLTAGE GRIDS ON THE BASIS OF PUBLIC AVAILABLE MAP DATA

André SEACK
TU Dortmund University – Germany
andre.seack@tu-dortmund.de

Jan KAYS
TU Dortmund University – Germany
jan.kays@tu-dortmund.de

Christian REHTANZ
TU Dortmund University – Germany
christian.rehtanz@tu-dortmund.de

## ABSTRACT

*Being subject of historic growth and focusing on the load supply in the past, only few low voltage grid models are available for research purposes. Due to energy system changes with large amounts of renewable feed-in being redistributed within these grids, it becomes more important to create a reliable data basis to evaluate innovative control approaches.*
*Facing this lack of data, in this paper, low voltage grids are constructed on the basis of public available geographical data from OpenStreetMaps (OSM) for selected residential areas.*
*The presented approach includes the definition of the supply task in the study region, the positioning of low voltage substations and a fitness evaluation of the generated network to optimize the positioning process.*
*In comparison to real network data it is demonstrated, that the proposed approach provides reasonable approximation of the real network for research purposes.*

## INTRODUCTION

In the past, low voltage grids were just constructed to meet the required load demand of the supply region. Based on adapted planning and operation guidelines, the initial state of the grids were planned to allow load growth for several years. However, being subject to historic expansion of the grid's structure because of adjustments and additions in the corresponding area, today's structure of the grids might not be the most efficient topology to serve the given supply task. With their focus being only on the load supply role, the low voltage networks did not get enough attention in the past. Therefore, in contrast to the medium and high voltage grids, the low voltage grids are usually not maintained in network calculation programs and just available as grid maps. Additionally, only few low voltage model grids are available in common literature. Due to the increasing amount of renewable in-feed it is necessary to improve the data basis for these networks to evaluate real loading situations in the planning process and innovative concepts and approaches in research [1]
Facing these discrepancies, in this paper low voltage grid models are constructed on the basis of public available geographical data from OpenStreetMaps (OSM) for selected residential areas. In common literature this approach is not well tackled to date. While [2] describes an approach to derive benchmark grids based on clustering of a high number of classified network information, [3] finds a solution to a topology derivation problem when only substation positions are available. There, the authors describe an algorithm to create different topologies for the same supply task. This contribution still lacks the positioning of new substations if the supply task is changing and the evaluation of the designed topology. To face this challenge, [4] describes an approach to create a loss optimized low voltage network with an ideal substation positioning. Extended in [5], this work is applied on a larger scale to optimize the deployment of a high number of substations in a selected planning zone.
With the main idea of [5] in mind, the definition of an ideal exemplary grid, which is only based on publicly available geographic data of the area, is presented in this work. This enables the evaluation of the actual situation in the grid. Additionally, the presented approach allows for validation of scientific findings in operation and planning of grids without being dependent on data of the local utilities. If the utilities' construction principles are taken into account, these grids can also be used for target grid planning purposes.
The paper starts with the developed approach and the definition of the supply task in a selected area. The following section focuses on the description of the load clustering and the assignment of grid connection points to obtained substations in the network. Afterwards the fitness evaluation and variation of substation position within the target networks is described. The proceeding of the grid generation is presented on the basis of a real discrete residential area. The paper concludes with an evaluation of the approach and an outlook on extension of the system

## PROPOSED APPROACH

The general process of the grid creation is depicted in Fig. 1. Public available geographic data is imported for the corresponding study region. The imported dataset contains residential housing information that is extracted. The buildings are then assigned to the nearest infrastructure axis (streets). In coherence with [2] it is assumed, that in residential areas only existing infrastructure axis are valid paths for electric lines.
With a clustering method low voltage substations locations are determined to assign the households evenly distributed to the foreseen substations in the study region. During this process, a typical load is

assumed for each house, based on the buildings base area. Every load is connected to the nearest substations via the existing axes (streets). Common planning and operation guidelines are taken into account. This includes the reallocation of nodes to another substation, if the allocation to the nearest is not compatible with the planning guidelines.

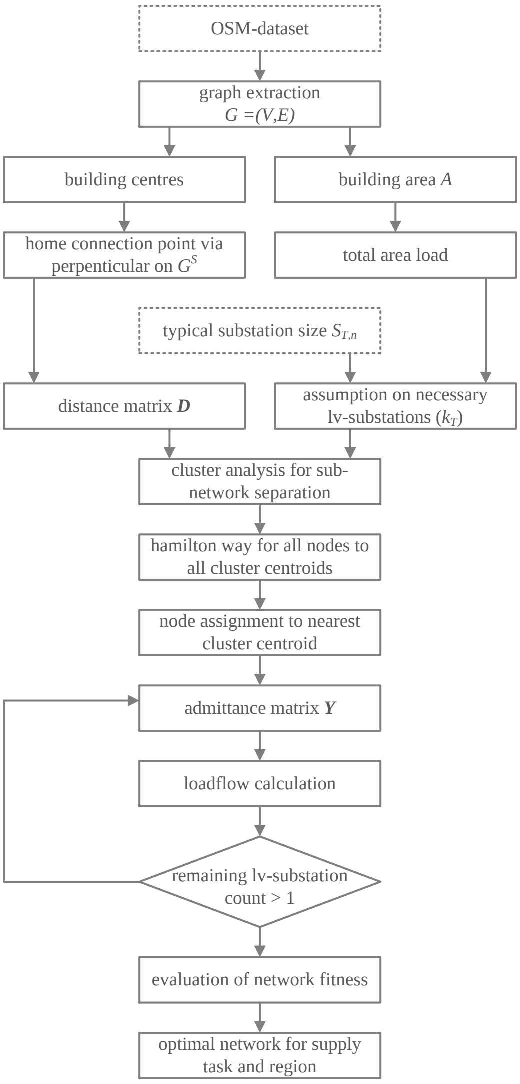


Fig. 1. Flow chart of the network generation process

By use of typical low voltage grid assets, the geographic information of the low voltage networks is transferred to static distribution grid models to enable the evaluation from a technical point of view. Besides checking the compliance with given technical boundaries like asset loading and voltage profile, the quality of the generated network is rated via a fitness function. This process is carried out for all sub-network nodes as possible location for the substation. The result is an optimal network for the regarded area and supply-task, considering boundary conditions of planning criterions.

## SUPPLY TASK DEFINITION

For demonstration purposes of the proposed approach a study region in the south of Germany is selected. The residential area consists of around 150 buildings. From the public available dataset of OSM the information on the location of residential buildings and street-sections is extracted. The universal graph definition is derived from [6]. The obtained graph

$$G = (V, E)$$

is defined by the set of nodes $V$ and the set of edges $E$, a two-dimensional set of $V$:

$$E \subseteq [V, V]$$

The edges E represent straight connections between two nodes of

$$V = \{v_1, \dots, v_n\}$$

with the total number of nodes $n$. $G$ is the input data for the target network generation algorithm.
All extracted building edges and street sections are part of $G$. The classification matrix $\boldsymbol{\Psi}$ includes the assignment of the nodes V to the geographical longitude φ and latitude λ as well as the node category $\psi$ (e.g part of a building or street).

$$\boldsymbol{\Psi} = \{V, \varphi, \lambda, \psi\}$$

Depending on the node category $\psi$, $G$ is separated into the sub-graphs for buildings $G^B$ and for streets $G^S$:

$$G^B = (V^B, E^B)$$

$$G^S = (V^S, E^S)$$

They include the graph information of all buildings and streets in the examined area. The extracted geographic information of the residential area is illustrated in Fig. 2.

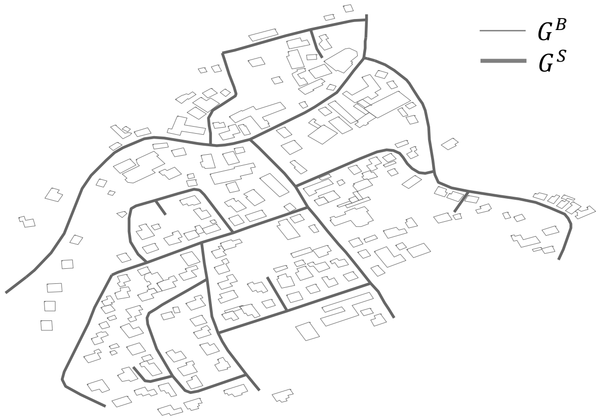


Fig. 2. Extracted building $G^B$ and street graps $G^S$ associated with $\boldsymbol{\Psi}$

Subsequently, $G^B$ is used to calculate the centre points of all buildings $\boldsymbol{M}$ in the area. For a building $m \in \boldsymbol{M}$ with the graph $G_m^B = (V_m^B, E_m^B)$, its centre coordinates $(\varphi_{m,mean}, \lambda_{m,mean})$ are determined by

$$\varphi_{m,mean} = \frac{1}{2}\left(max\left(\varphi(V_m^B)\right) + min\left(\varphi(V_m^B)\right)\right)$$

$$\lambda_{m,mean} = \frac{1}{2}\left(max\left(\lambda(V_m^B)\right) + min\left(\lambda(V_m^B)\right)\right)$$

The set of all building centres

$$\boldsymbol{C}(\varphi_{m,mean}, \lambda_{m,mean}) \forall m \in \boldsymbol{M}$$

provides an origin for the building grid connection points determination. With their set being $\boldsymbol{P}$, the final grid connection for one building $m$ is defined by

$$p_m = min[(c_m \perp E^S) \cap dist(V^S)]$$

It is either the dropped perpendicular to the nearest edge $e \in E^S$ or the nearest street node out of $V^S$. The resulting building grid connections are depicted in Fig. 3.

Fig. 3. Representation of all building center $\boldsymbol{C}$ and grid connection points $\boldsymbol{P}$ in the study area

The next step is the distance calculation between neighbouring grid connection nodes, considering the infrastructure topology $G^S$. For $n$ nodes in the area, the distance matrix $\boldsymbol{D}$ contains geographic distances $d_{ij}$ between two neighbouring nodes $n_i$ and $n_j$ by taking into account $\boldsymbol{\Psi}$ and $G^S$. Due to the structure of the dataset, $\boldsymbol{D}$ is symmetric and the diagonal elements for $i = j$ are $d_{ij} = 0$.
Another major input for the subsequent clustering process is the approximation of the total load $S_{L,total}$ in the regarded area. Since the base area of a building $A_m$ is the only available input parameter for the estimation of the electric load $S_{L,m}$, proportionality is assumed between these parameters.

$$S_{L,m} \propto A_m$$

With an assumed average living area $A_{mean}$ per household, the connection rating power $S_{L,m}$ is an integer multiple of $S_{L,mean} = 2\ kW$.

$$S_{L,m} = S_{L,mean} \frac{A_m}{A_{mean}}$$

The sum of all connection rating powers in the area is the basis for the quantity $k_T$ of local transformers, assuming a nominal transformer power $S_{T,n}$:

$$S_{L,sum} = \sum_m S_{L,m}$$

$$k_T = \frac{S_{L,sum}}{S_{T,n}}$$

The buildings, their aggregated load and the quantity of transformers define the supply task for the analysed area.

## LOAD CLUSTERING

The determined grid connection points have to be assigned to one of the possible substations, which need to be placed in the supply area. The best geographic position of the transformers is identified with a clustering of all grid connection points $\boldsymbol{P}$. The number of cluster centroids is equivalent to the minimal number of required transformers $k_T$.
The grid connection points of every building in the assessed region are assigned to the nearest cluster centre. However, the infrastructure's topology is not yet considered.
With $\boldsymbol{D}, \boldsymbol{\Psi}$ and $G^S$ all valid paths $X$ from each node $n$ to all cluster centroids $K$ are evaluated and combined to the way matrix $\boldsymbol{W}$. It represents the minimal geographic distance of every node to every cluster center $k$.

$$\boldsymbol{W} = min[\,dist(n, K)] \forall X$$

For the node assignment matrix $\boldsymbol{A}$ these distances are evaluated for every node and all nodes are assigned to the nearest cluster centroid. This assures that every grid connection point is linked to the closest substation, considering the given topology. The resulting low voltage substations with the associated subnetworks are depicted in Fig. 4. All grid connection points within the red ellipse are reallocated to the green sub-network in the final step of the grid generation process, on the basis of the way matrix.

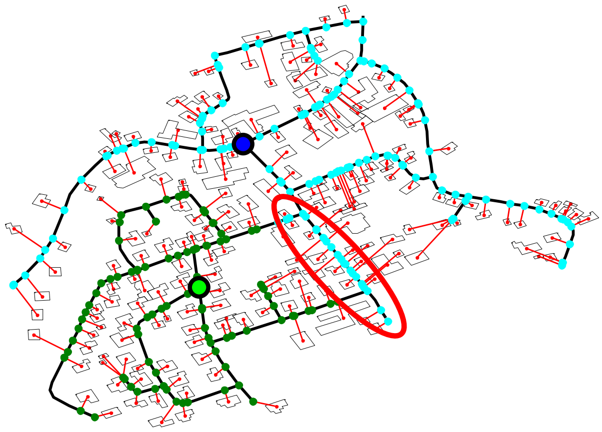

Fig. 4. Resulting low voltage sub-networks after load clustering

## FITNESS EVALUATION

The final step of the network generation process is the fitness evaluation of the generated low-voltage networks. The network admittance matrix $\boldsymbol{Y}_K$ is generated based on standard assets für low voltage network and the distance between individual nodes $\boldsymbol{D}$. In conjunction with the substation position and the individual loads $S_L$ the defined topology is sufficient for the static network analysis calculation for each sub-network.

Since the substation position was determined without considering the path topology, the position may not be optimal for the resulting grid. For the identification of the optimal position, the substation's position in the clustered sub-network is varied, examing every node of the sub-network as potential location.

With the substation position the reference node for the analysis calculation is changed. The nodal admittance matrix $\boldsymbol{Y}_n$ is set up from the combination of the distance matrix $\boldsymbol{D}$ and electric parameters for standard assets in the low voltage grid. $\boldsymbol{Y}_n$ links the nodal voltage vector $\boldsymbol{U}_n$ with the nodal current vector $\boldsymbol{I}_n$ for all nodes $n$:

$$\boldsymbol{I}_n = \boldsymbol{Y}_n \boldsymbol{U}_n$$

$$\boldsymbol{I}_n = [\bar{I}_1, \bar{I}_2, \dots, \bar{I}_i, \dots, \bar{I}_n]^T$$

$$\boldsymbol{U}_n = [\bar{U}_1, \bar{U}_2, \dots, \bar{U}_i, \dots, \bar{U}_n]^T$$

Detailed definition and the iterative calculation method are derived from [7]. The result of the iterative static network analysis is the complex nodal voltage vector $\bar{U}$.

$$\bar{U} = \bar{e} + j\bar{f}$$

In the following it is assumed, that a minimal voltage band throughout the generated network leads to minimal losses in the sub-network. Therefore the main criterion for an optimal substation position $p_{\tau,opt}(\varphi_{\tau,opt}, \lambda_{\tau,opt})$ of the substation $\tau$ is the minimal difference between the maximal $U_{p,max}$ and minimal $U_{p,min}$ voltage occuring for all nodes $n$. The criterion is checked for every position $p$ as element of all possible substation positions $\boldsymbol{P}$:

$$\Delta U_p = \max_n U_n - \min_n U_n$$

$$p_{\tau,opt}(\varphi_{\tau,opt}, \lambda_{\tau,opt}) = p(x | x = min(\Delta U_p) \forall p \in \boldsymbol{P})$$

The position with the minimal occurring nodal voltage interval $\Delta U_p$ defines the optimal postion for the substation $p_{\tau,opt}$. The evaluation for the green network ist depicted in Fig. 5, where the optimal position is marked in red.

It can be stated, that in this particular case the best substation position is ideltical to the cluster centroid. However, the technical optimal position is dependent on the network topology and might differ from the geographic optimal positioning.

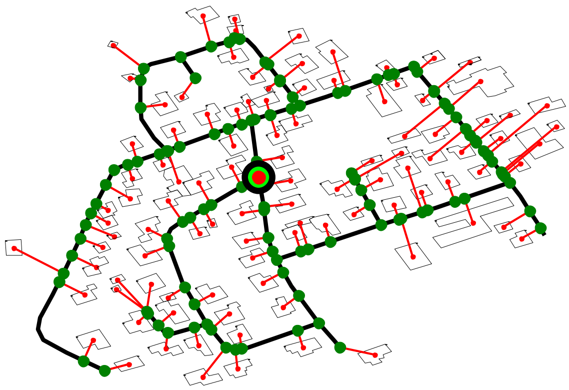

Fig. 5. Best substation position marked in red

## COMPARISON TO REAL NETWORK DATA

In conventional distribution grid planning process the utility's planning department has the same starting and boundary conditions as the presented network generation algorithm. When new settlement areas are planned, the geographical position of new buildings and infrastructure axis is obligatory and has to be respected by the network planning department when defining the grid's topology. Especially in low- and medium voltage networks, the planning personal will use the paths of the infrastructure if the network is designed as a cable-infrastructure.

Therefore it is reasonable to compare the results of the developed network generation approach to existing network data from the distribution grid operator for benchmarking the identified grid. The comparison of the supply area's grid with the real grid in Fig. 6 states, that the presented approach delivers a good approximation of the real network configuration. The determined substation position is very close the real location. The main differences in the feeder structure are justified in the usage of different technologies (overhead line vs. cable) and planning principles.

Fig. 6. Real grid extract of the studied area

## CONCLUSION AND OUTLOOK

The developed and presented proceeding allows for the generation of a large number of grids based on geographic data. Especially for research purposes, this network data is necessary to evaluate innovative approaches for the distribution grid planning and operation. The comparison with the real grid in the chosen test supply area demonstrates that the developed approach provides a good approximation for the real network infrastructure data.

In future work this approach will be extended to larger districts of low voltage network (e.g. complete villages) as well as medium and high voltage layer of distribution grids to establish a base for an integral distribution grid analysis for a selected area. Additionally, different asset combinations will be included to reproduce typical and historically grown network infrastructure.

Since the information on the positioning of distributed generation is publicly available in Germany, this database and the corresponding coordinates will be included as well to enhance the supply task definition in the regarded area. Another integral step to improve the example grid generation is the automatic detachment of meshed structures.

## REFERENCES

[1] German Energy Agency (dena), ed.: *dena* “Distribution Grid Study – German electricity distribution grids in need of significant expansion by 2030”, *Available at: www.dena.de (German)*

[2] J. Dickert, M. Domagk, P. Schegner, 2013, “Benchmark low voltage distribution networks based on cluster analysis of actual grid properties” *IEEE PowerTech, Grenoble*, 1–6.

[3] C.M. Domingo, T.G.S Román, A. Sánchez Miralles, J.P.P González, A.C.Martínez, 2011, “A Reference Network Model for Large-Scale Distribution Planning With Automatic Street Map Generation”, *IEEE Transactions on Power Systems 26*, 190–197

[4] A. Navarro, H. Rudnick, 2009, “Large-Scale Distribution Planning—Part I”, *IEEE Transactions on Power Systems 24*, 744–751

[5] A. Navarro, H. Rudnick, 2012, “Large-Scale Distribution Planning—Part II”, *IEEE Transactions on Power Systems 24,* 752–758

[6] R. Diestel, 2010, *Graphentheorie,* Springer, Heidelberg, Germany, 01-30.

[7] E. Handschin, 1987, *Elektrische Energie-übertragungssysteme,* Hüthig, Heidelberg, Germany, 23-56.