TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

AUSTRIAN INSTITUTE OF TECHNOLOGY

Master Thesis

Systematical Analysis of Low Voltage-Networks for Smart Grid Studies

under the supervision of

ao.Univ.-Prof. Dipl.-Ing. Dr.techn. Gerhard Theil

Vienna University of Technology

Dipl.-Ing. Benoît Bletterie

Austrian Institute of Technology

Dipl.-Ing. Franz Zeilinger

Vienna University of Technology

by

Serdar Kadam, BSc

Matriculation number: 0425089

Donaufelderstraße 103/2/4 A-1210 Vienna

Executive Summary

One of the prerequisite for implementing smart grid solutions into LV networks is to have a better understanding of these networks. The systematical analysis of LV-networks can be done on two levels. On the feeder level or on network level. A LV-network can supply different types of feeders: short or long feeders, feeders supplying a rural or urban area, etc. To group feeders with similar characteristics, indicators are needed. In a first step, new indicators and indicators already introduced in previous studies were implemented in the network simulation program DIgSILENT PowerFactory using DPL (DIgSILENT Programming Language). The main results from the network computations were written to excel and further processed using macros. In a second step, the information content of all indicators was analysed with a linear regression model. With this analysis, indicators that could be calculated by a combination of other indicators were identified and removed. The aim of this approach was to clarify if the indicators that can only be computed by simulations in PowerFactroy are significant. If not, the network models in PowerFactory would not contribute to the characterization of feeders and networks. A principal component analysis (PCA) was done in order to try to reduce the number of dimensions and the three most important principal components to visualize data. In a fourth step, feeders and networks have been classified in similar groups on the basis of different clustering algorithms. In order to complete the cluster analysis, a criterion has been introduced to determine the number of clusters suitable to classify the whole set of feeders or networks. For each cluster the hypothetical median LV-network or feeder was calculated by the cluster members. Then the most similar LV-network or feeder was identified as the most descriptive element of the clusters. The most descriptive global element was also found by electrical and non-electrical indicators on feeder and network level. The most descriptive global element was used to find outliers. Finally, two different Snap-shots of two feeders of the LV-network network 01 were available until the end of this work. The balancing gain for these two Snap-shots was calculated by analysing the voltage range for the Snap-shot and the optimally symmetrised case in PowerFactory.

Acknowledgement

This thesis is the final step to finish my studies in electrical engineering, therefore I would like to start to express my thanks. I would like to thank Prof. Gerhard Theil, DI Benoît Bletterie and DI Franz Zeilinger for their input and discussions during this work.

Special thanks go to my parents, Gülay and Mithat, for their continuous support and encouragement for my siblings and me. I would like to dedicate this work to them.

Serdar Kadam

Kurzfassung

Eine der Voraussetzungen für die Implementierung von Smart Grid Konzepten in Niederspannungsnetzen, ist ein ausreichendes Verständnis dieser Netze. Niederspannungsnetze können auf zwei Ebenen systematisch untersucht werden: Auf Strang- oder Netzebene. Ein Niederspannungsnetz kann aus verschiedenen Arten von Strängen bestehen. Stränge können kurz oder lang sein, ländliche oder städtische Gebiete, etc. versorgen. Um Stränge in Gruppen mit ähnlichen elektrischen Eigenschaften zu gruppieren, werden Indikatoren benötigt. Im ersten Schritt wurde die Berechnung neuer und bereits eingeführter Indikatoren in der Skriptsprache DPL der Netzberechnungssoftware DigSILENT PowerFactory implementiert. Indikatoren, die in PowerFactory berechnet wurden, können als elektrische Indikatoren beschrieben werden. Die Ergebnisse wurden nach Excel exportiert und mit Makros vervollständigt. Danach wurden die Ergebnisse mit einem multidimensionalem Regressionsmodell analysiert um vorhandene Redundanz in den Indikatoren zu erkennen und damit die Anzahl der Indikatoren zu reduzieren. Danach wurde eine Hauptkomponentenanalyse durchgeführt um die Relevanz der Indikatoren zu untersuchen und um die 3 wichtigsten Hauptkomponenten für die Darstellung der Ergebnisse in einem Koordinatensystem zu verwenden. Als nächstes wurden die Stränge und Netze mit einem Hierarchischen Clustering Algorithmus gruppiert. Für jede Gruppe wurde der Charakteristischste Strang (Clusterzentrum) bestimmt. Danach wurden die Clusterzentren, die mittels elektrischen Indikatoren gefunden wurden beschrieben. Am Ende der Arbeit wurden erste Snapshots die innerhalb des Projektes ISOLVES:PSSA-M aufgenommen wurden analysiert.

Nomenclature and abbreviations

Nomenclature

ε	[1]	equivalent load location
Z_Σ	[Ω]	equivalent sum-impedance
u,v	[p.u.]	phase to neutral voltage
S_k"	[MVA]	initial short circuit power
Z_k	[Ω]	short circuit impedance
d	[m]	distance
dvdP	[%/kW]	real power sensitivity
dvdQ	[%/kVAr]	reactive power sensitivity
P_{Load}	[kW]	total active power that is consumed at a node (by loads)
P_{Flow}	[kW]	total active power that is transmitted to a node
adtn	[m]	average distance to neighbour nodes
NON	[1]	number of neighbour nodes
$\overline{NON}$	[1]	average number of neighbour nodes
NVR	[p.u.]	node voltage range
FVR	[p.u.]	feeder voltage range
GVR	[p.u.]	grid voltage range
PR	[1]	power ratio
TR	[kVA]	tranformer rating
BG	[%]	balancing gain

Abbrevations

LV	Low voltage
LVG	Low voltage grid
PSS	Power Snap-Shot
PSSA-M	Power Snap-Shot Analysis by Meters
ISOLVES	Innovative Solutions to Optimise Low Voltage Electricity Systems
PV	photovoltaic
DER	distributed energy resource
LDF	Load-flow
DNO	Distributed network operator
PCA	Principal component analysis

Contents

1 Introduction

One of the prerequisite for implementing smart grid solutions into LV-networks is to have a better understanding of these networks. When investigating or developing innovative smart grids concepts to enable an optimal integration of DER (distributed energy resources), one of the first questions that arise is how to model the system. As mentioned in [2], LV-network modelling remains a challenging task due to the lack of data (i.e. load profiles, phase information, neutral earthing). In the absence of detailed models, the validity of network studies can be questioned since they are based on unrealistic assumptions. In order to address the mentioned problems, some research work is on-going within the project ISOLVES:PSSA-M (Innovative Solutions to Optimise Low Voltage Electricity Systems: Power Snap-Shot Analysis by Meters) [1].

After the liberation of the electricity market in Austria ([7]), investments were reduced until the year 2005 started to increase since then [4]. According to the outlook for the next 10 year of Energy Control Austria, Smart Meters will play an important role in the energy market. Customers will be informed in shorter intervals about their electricity use which will raise their awareness on costs and potential savings and new price models will be offered. Nevertheless, the overall consumption is predicted to rise. With the introduction of Smart Meters, network operators will have more precise information about consumption. On 24th of April 2012 a new Smart Meter regulation came into effect in Austria [5]. This regulation which is the national implementation of the European directive forces that 10% of all metering points have to be equipped with Smart Meters until 2016 and 95% until 2019. From a technical point of view, such meters can be also used as measurement instruments. These measurements can be used to collect data about the loading of the network, identify load situations that are corresponding to the highest stress conditions (voltage or loading). In the project ISOLVES:PSSA-M smart meters are used to take 'Snap-Shots' of the network. A Snap-Shot consists of data synchronously measured by all meters at a certain timestamp. The meters transmit the measured data of active and reactive power and the line-neutral voltages to a data concentrator in the transformer station. Later, the data is transmitted to the PSSA-Host and can be accessed over a database. In the frame of the project, Snap-Shots will be taken in 34 networks of EAG in Upper Austria. The models of the 34 LV-networks selected for the study have been built in the simulation software DIgSILENT PowerFactory.

In this thesis the 34 low voltage-networks will be analysed using specific indicators introduced to characterize low voltage grids. Indicators will be introduced to characterize low voltage grids. Some are based on indicators introduced for simplified network topologies but which need to be enhanced for use on more complex network structures. In [10], [8] some indicators are introduced to estimate the acceptable amount of PV generation for specific networks. The characterisation and classification of the 34 networks shall help DNOs in assessing possible smart grids concepts for the integration of DER into LV-networks.

2 Objectives and methodology

In this chapter the objectives and the methodology will be discussed. The objective of this work is the analysis of feeders and networks in PowerFactory to characterize and classify them. To classify networks or feeders, indicators are needed that describe networks or feeders and provide a metric for a comparison. Some indicators are defined by the information of the network infrastructure ('non-electrical indicators') usually available in network information systems (NIS) or the GIS. Others have to be calculated in a network simulation software ('electrical indicators'). Therefore it is targeted, that the usage of generic data will also deliver information to a certain amount, for characterization and classification on network and feeder level. In a first step, generic load data consisting of uniform load values will be used. The approach which allows analysing the network topology can then be improved by considering real load data from Power Snap-Shots.

2.1 General objectives

The objectives of this work are:

- Introduction and analysis of various indicators for the characterization of 34 LV-networks and 247 feeders
- Characterization of LV-networks on feeder and network level
- Methodological classification of networks or feeders by indicators and clustering on feeder and network level
- Exemplarily Analysis of feeders with already available PSS data

Figure 1 shows the approach of this work. In a first step, the topology of the 34 networks was modelled in the network simulation software PowerFactory. To characterize the networks and feeders, indicators are needed. In [8] for example, the transformer rating, the length of the feeders and the equivalent load location were suggested as indicators to distinguish between LV-networks. An indicator to describe feeders was introduced in [13]. In principle, topological information e.g. the total cable length or electrical information e.g. the initial short circuit power could be used as indicators for feeders or networks. The indicators will be discussed in the next chapter. The programmed scripts in PowerFactory were executed on all 34 networks with generic loads using an external loop. The used generic loads are characterized by a loading of 1kW and 0.1kVA symmetrically distributed on the 3 phases. The results of this generic analysis can be seen in section 5.

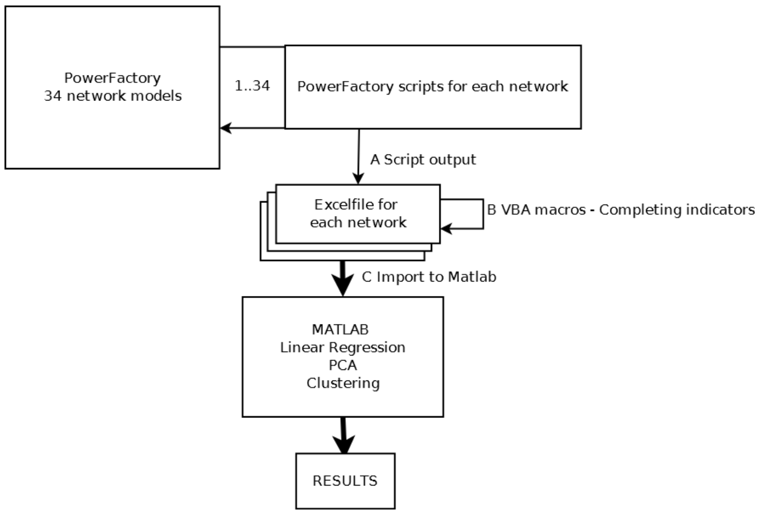


Figure 1: Methodology

After the definition and implementation of the indicators, grids or feeders could be characterized or classified. However, some implemented indicators could contain redundant information. Therefore it is targeted to reduce the number of describing indicators to a minimum with an appropriate model. This will be discussed in section 2.4. After the identification of the most essential indicators, a methodology to classify networks or feeders will be discussed.

2.2 Network modelling

The LV-networks were modelled in DIgSILENT PowerFactory. The identification of feeders was possible by assigning a zone to each feeder. Every object (line, node, etc.) of a feeder was assigned to the same zone. This allows to use feeder based scripts by sorting elements by zone. In some of the networks, ring switchers are connecting between feeders for reconfiguration purpose. They are however in normal conditions always opened. The cable and transformer data were stored as a common library for the 34 networks. All network models contain a medium voltage slack and a distribution transformer. The 230V node of the secondary side of the transformer will be named distribution node. 3 supply options are available:

- A medium voltage slack combined with a transformer
- 3 voltage sources at the distribution node
- A low voltage slack at the distribution node

For the generic analysis the first option will be used to include transformer loss effects that are depending on the loading of the grid. To include these effects in the generic analysis, a voltage setpoint of 1 p.u. was set at the slack on the medium voltage node. Another reason for this choice is that the size if the transformer impacts the short-circuit current and the network impedance. Since the transformers of all LV-networks are modelled realistically, it is advantageous to use a medium voltage slack. Therefore the voltage at the secondary side of the transformers depends on the overall loading in the network. Option two and three will be of interest for model validation, to get the closest results to the measured values. The difference between 3 single phase voltage sources and a low voltage slack at the distribution node is that 3 single phase voltage sources can simulate an unsymmetrical network, as the voltage set point for each source can be set individually. A slack on the other side, has just a symmetrical voltage set point for all the phases. The network models are based on the 4-wire model with 3 phases and neutral wiring. As the value of the grounding resistances is unknown and difficult to measure, a uniform grounding resistance of 2Ω was assumed. This assumption does not impact this work since loads were defined as symmetrical. Further investigations about the validity of this assumption will be done in the project. An important part of the analysis of LV-networks is the method of providing the network models in digital form. A manually input of detailed models is expensive and time-intensive. Therefore e.g. in [13] half automated approaches are used to import the network models from GIS systems. It is also mentioned that this approach has an error rate under 10% and therefore is still suitable for statistical analysis. The LV-networks in this study were manually entered in PowerFactory with all available details.

2.3 Used network simulation tools

In this section, the used network simulation tools will be discussed. The tools that will be used are the loadflow calculation, the short circuit calculation and the sensitivity analysis of PowerFactory. These tools are needed to calculate some indicators, that will be introduced in the next chapter. Indicators, that are calculated using PowerFactory will be labelled as 'electrical' indicators. The indicators based on short circuit calculations were calculated according to IEC 60909 which is available in PowerFactory. The IEC 60909/VDE 0102 method uses an equivalent voltage source at the faulted bus and is a simplification of the superposition method (Complete Method). The goal of this method is to accomplish a close-to-reality short-circuit calculation without the need for the preceding load-flow calculation and the associated definition of actual operating conditions [3].
The short circuit calculation was executed for nodes between the distribution node and the 'end node' of a feeder.

For the generic analysis, a symmetrical loadflow calculation was used. The loads were modelled as PQ-loads (1kW, 0.1kVA, per load). A load flow calculation gives the active and reactive power flows, the voltages for all nodes and the currents through elements. These values are needed to calculate some indicators presented in the next chapter.

With a load flow calculation the 'end node' of each feeder can be identified: it is defined as the node with the lowest voltage. If there was a relevant amount of DER, the identified node need not to be at the end of the feeder any more. Therefore the part from the identified node to the topological end node would not be part of the analysis and some indicators that are defined for this node would be calculated for the identified node. Therefore, in the generic analysis no DER was considered and every 'end node' has no further connections to other nodes. A sensitivity analysis was used to find the path between an end node and the distribution node. The sensitivity analysis describes the voltage change resulting from a changed power injection. The method used in PowerFactory was 'Sensitivity to a Single Busbar'. With this method, the effects on the voltage of the injections of ΔP and ΔQ at the selected busbar are calculated for the whole network (i.e. for all buses and branches) ([3]). A power change at the end node has the greatest effect compared to the nodes before and becomes smaller along the path until 0 at the slack (in the models the medium voltage node of the transformer station). Therefore the path can be found by analysing the sensitivity dvdP ([%/kW]) in the feeders.

Table 1: Example of an indicator matrix

feeder	Z_K	...	Δu
	$[\Omega]$	...	$[p.u.]$
feeder 1	0.35	...	2
...	0.5	...	2.4
feeder n	0.2	...	1

2.4 Statistical tools

This chapter provides a short overview on the statistical tools used in this work.

2.4.1 Linear regression model

In the next chapter many indicators will be defined that could be used in the analysis on feeder or network level. Once the programming of the indicators is done, the calculations for 247 feeders can be executed easily. However, to improve clustering results and to identify the most relevant indicators, the relation between indicators should be analysed on redundant information to reduce the number of indicators. The relation between the indicators will be analysed with a linear regression model. Indicators, that could be calculated derived by a combination of other indicators could be removed from the indicator matrix. The indicator matrix on feeder level is a 247x$N_{feeder-indicators}$ matrix and a 34x$N_{network-indicators}$ matrix on network level. An example of an indicator matrix on level can be seen in table 1. A linear regression model with constant term can be described by the following equation:

$$Y_j = ß_0 + ß_1 \cdot x_1 + ... + ß_i \cdot x_i + \varepsilon_i \tag{1}$$

j ... indicator index n ... number of indicator i = 1,...,n \j
β_i ... coefficients

ε ... error term

Y is a column and X are the remaining columns of the indicator matrix. To identify indicators, that could be calculated by others, the coefficient of determination R^2 will be used. The R^2 coefficient describes how well the model describes the set of observations. R^2 of 1 would mean perfect fit, therefore a threshold will be defined to distinguish between relevant indicators and others that can be inferred by the relevant ones.

After the linear model analysis, the indicators will be normalized to the highest value in order to avoid distortion in the significance due to indicators with large numerical values

2.4.2 Principal component analysis

Once the indicators are selected, matrices will be used to store the values for each feeder or network. With the linear regression model the number of indicators will be reduced and the remaining indicators will be stored in reduced indicator matrices. An important task during the work will be to visualize the distribution of feeders or networks. The reduced indicators matrices will be used for that. However, these matrices have still many indicators (> 3) and can therefore not be easily visualized. In order to reduce the number of dimensions and be able to visualize the observations in 3D diagrams, a PCA was used. In a PCA an orthogonal transformation is used to convert a set of observations of variables that might be correlated into a set of values of linearly uncorrelated variables. A PCA returns the principal component coefficients, also known as loadings. PCA returns a matrix, each column containing coefficients for one principal component. The columns are in order of decreasing component variance [11]. If for example 9 indicators are used, feeders and network are points in a 9-dimensional space. The indicators however may not be orthogonal. The the three most relevant principal components will be used to plot the points in the PCA-space, knowing that (a small) part of the information is omitted.

2.4.3 Clustering

After the indicators are reduced with a linear regression model, they are split up into two groups. The first group contains as previously explained indicators, that need a network simulation program to be calculated (electrical indicators) and the second group indicators, that can be calculated without network simulation. After that step, the target is to group networks or feeders by their indicator values. To solve this kind of problems, a clustering algorithm will be used. 2 clustering algorithms have been analysed.
A favourable approach would be to obtain the optimal cluster size by the cluster algorithm itself. Therefore, an agglomerative (bottom-up) hierarchical clustering algorithm will be used. The hierarchical clustering can be achieved by three commands in Matlab: ‘pdist’, ‘linkage’ and ‘dendrogram’ (for visualization, see [11]). With ‘pdist’ the euclidean distance between pairs of objects is calculated. After that on the output of ‘pdist’ the command ‘linkage’ is applied. ‘Linkage’ returns a matrix that encodes a tree of

hierarchical clusters. Finally, this matrix is used with the command 'dendrogram'. The hierarchical clustering algorithm starts with defining a cluster for each point. After that step, clusters with the closest distance are merged. This process is continued, until all points are merged to a single cluster. This option allows to 'set' a specific cluster size. An sample plot of a dendrogram can be seen in figure 2. The number of clusters can be defined by cutting the dendrogram at a specific height. The number of intersections between the horizontal cut and the vertical lines gives the number of clusters.

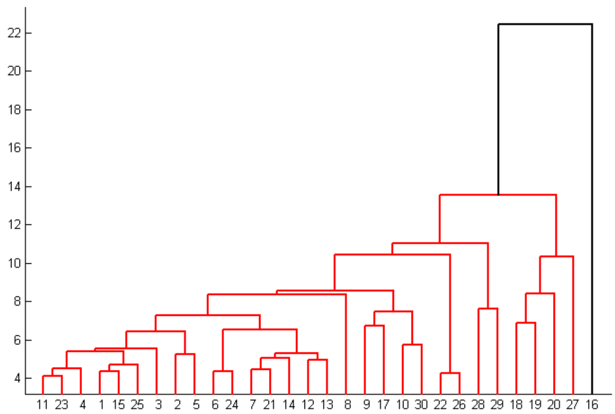


Figure 2: Dendrogram

The y-axis indicates the euclidean distance between the clusters. On the x-axis the elements are placed at y=0 equidistant. These points are the leaves of the dendrogram. The number of elements to be drawn as a leaf of the dendrogram can be defined. For example if the number of elements is 50 and 30 leafs are defined, then the 50 elements will be merged to 30. 'Dendrogram' returns a vector T of size nx1 where n is the number of elements. The vector T contains the cluster indices of the elements at the bottom of the dendrogram. Another parameter of dendrogram is 'threshold'. This parameter can be used to cut the dendrogram at a specific height (distance between the clusters). After that clusters with a lower distance than the threshold are merged to a new cluster. This option allows determining the cluster size depending on a definable/acceptable distance between points. For example, if the dendrogram in figure 2 is cut at the height 12, the objects could be grouped in 3 clusters. Object 16 has a high distance to other objects, as it can not be grouped with any other object and can be considered as outlier. Selecting a 'threshold' of height 8 would result in 9 clusters.
Secondly, the 'K-means' algorithm has been used (see [11]). This algorithm

selects randomly cluster centres from the dataset and calculates the cluster for each point. 'K-means' returns a nx1 vector with the cluster index of each point. The results of 'K-means' have been compared to the results of the hierarchical clustering. To do so, a measure is needed: a function in Matlab was used, that can be applied on both algorithms: silhouette. This function returns the silhouette value for each point. The silhouette value is a measure of how similar that point is to points in its own cluster compared to points in other clusters, and ranges from -1 to +1. It is defined as:

$$S(i) = (min(b(i,:),2) - a(i))./max(a(i), min(b(i,:))) \tag{2}$$

where a(i) is the average distance from the ith point to the other points in its cluster, and b(i,k) is the average distance from the ith point to points in another cluster k [11]. The mean of all S(i) is a criterion how good elements fit to the assigned cluster. Values above 0.5 indicate a reasonable structure and values above 0.75 a strong structure [14].

3 Introduction of suitable indicators for characterising LV-grids

As mentioned in the introduction, the 34 LV-networks of this study are fully equipped with Smart Meters and the models in PowerFactory can be validated with snapshots. In PowerFactory, many characteristic parameters are calculated within a load-flow calculation or short circuit calculation (see 2.3). More than that, the script language DPL can be used to implement calculations of new indicators. This script language allows to write with a DDE-connection directly to excel. Indicators will be written to excel, to complete the table of indicators discussed in this chapter macros will be used. Median, max and min values, for example, were calculated using excel macros to reduce the calculations in PowerFactory to a minimum. At first, the path search algorithm will be explained with an example. Next, indicators for low voltage networks or feeders were introduced and discussed in [8], [10], [12], [13]. Some of these indicators have to be adapted for the analysis and additional indicators will be introduced too. An important issue is the conversion of indicators on feeder level to the network level. Some indicators are only available at network level. On the other side indicators on feeder level have to be aggregated in a reasonable method transformed to network level. Different approaches are needed for specific indicators. This will be discussed together with the indicators.

3.1 Path search

Many indicators that will be discussed in this chapter refer to the 'end node'. The 'end node' is the node with the lowest voltage in a feeder and can change with every PSS, depending on the loading in the feeder. The programmed DPL scripts can easily identify this node, by sorting all nodes of a feeder by the voltage. The path of a feeder indicates the topological order of all nodes from the 'end node' to the transformer station. This path is necessary to draw voltage, impedances and other indicators over the distance for graphical comparisons between them. These diagrams can be used to analyse and compare different feeders across LV-networks. Examples of such diagrams will be given in the next chapter. The path is also essential to find the correct switch that connects the feeder to the distribution station. The total current through this switch will be needed to calculate the equivalent sum impedance.

The approach to find the correct path starts with the identification of the node with the lowest voltage. After that, a load-flow calculation and a sensitivity analysis is executed. These calculations return among others, the voltage sensitivity on real power change, dvdP. This value is 0 at the used slack (in the generic analysis the medium voltage node of the transformer station) and becomes higher with increasing distance to the transformer station. The highest value is at 'end node' the sensitivity was calculated for, which is already known by searching for the node with the lowest voltage in the feeder in the set of all nodes of that feeder. The next step is to sort all nodes in the set by dvdP. To select the next closer node to the transformer 3 conditions have to be fulfilled. Firstly, inside an iteration the next node with a lower dvdP value from the set of nodes is selected. This next node has a lower dvdP value and therefore could be the next closer node to the transformer station. Nevertheless, this condition is not satisfactory. In many feeders branching exist. This means, that there is more than one possible 'end node' and a possibility to select a node which is not connected to the 'end node'. Secondly, the next node has to be connected to the 'end node'. This is examined if the next node has a common element (cable, transformer or switch) with the 'end node'. If this is true, the third condition is checked. The third condition ensures that the next node is closer to the transformer station. After a load-flow calculation the property b:dist of nodes is available. This property returns the distance to the transformer station in meters. If all three conditions are met, the next node is added to a new set of the path from the 'end node' to the transformer station. Then the iteration on the set of nodes continues for the last selected node to the path. The iteration ends at the LV-node of the transformer station. The switch between the highest node of the feeder and the transformer station is stored in another set. The set of switches is needed to calculate the equivalent sum impedance. After the path search, a set of the path is available in the correct order. Interesting simulation results from short-circuit or load-flow calculations can now be exported together with the path information. The path is also essential to find the equivalent load location, which will be introduced in section 3.3.

Figure 3 shows feeder 1 of Neukirchen. In the figure the names of the nodes, cables and metering points can be seen.

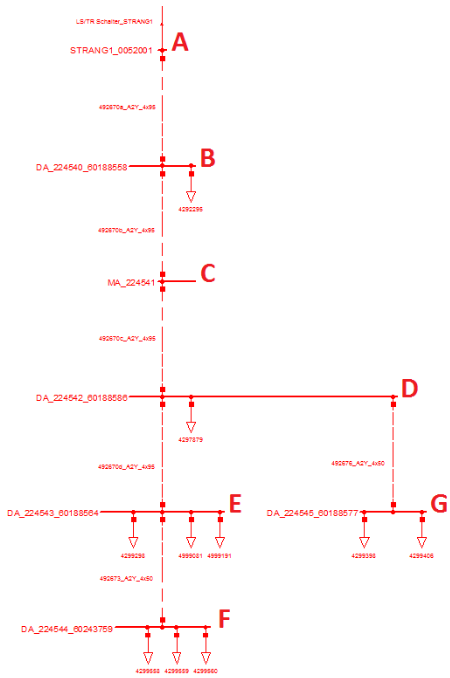


Figure 3: Example of path finding

To ease the traceability of the example, the nodes were marked with letters from A to F in table 2. The sensitivity at all nodes of feeder 1 can be seen in table 2. The end node can be identified by the lowest voltage in the feeder (without DER), which is at node F. This node is the first node of the set of the path. To find the node closer to the transformer station, the next node in the table above the 'end node' is analysed. The node E is connected, and closer to the transformer station and is added to the set of the path. Next, the neighbours of E are analysed. The next node in the table has a lower dvdP values, however, is not connected to the actual node. It is a potential 'end node' of the feeder. It would have been selected as 'end node' if the voltage drop caused by the loads on the node G would be higher than at node F. Therefore, the next node in the list is analysed. Node D has a lower dvdP and is connected to the actual node. From now on there are no branching and the iteration continues upwards until A, which is connected to the distribution node. The distribution node itself is not assigned to any feeder.

Table 2: Example of path finding (dvdP at the nodes)

Marker	Node	p.u./MW
A	STRANG1_0052001	0.01626453
B	DA_224540_60188558	0.06154764
C	MA_224541	0.1104487
D	DA_224542_60188586	0.1647812
G	DA_224545_60188577	0.1648128
E	DA_224543_60188564	0.2118454
F	DA_224544_60243759	0.3583738

3.2 Voltage ranges

3.2.1 Definition of voltage ranges

The voltage at the distribution node and the minimal voltage in the network and feeder, respectively, are measured. The voltage drop is the difference of these two values. On network level, another voltage could be of interest: The highest voltage difference between the 'end nodes'. This could characterize the alikeness of feeders in the same network. It is estimated that this indicator will change for every snapshot. In the generic analysis no DER was simulated. Therefore the highest voltage in the network is at the distribution node. These indicators can be defined as followed.

The voltage range is obtained by the comparison between the highest and the lowest voltage and can be examined at different levels. Firstly, in a unsymmetrical loaded LV-grid there can be a spread between the phases L1,L2 and L3, defined as NVR (node voltage range). In a symmetric loaded LV-grid the line-neutral voltages at every node should be equally the same and therefore NVR is always 0. Secondly, the voltage range can be analysed at a higher level, the feeder. Here the lowest line-neutral voltage inside a feeder (=zone) and the highest line-neutral voltage inside the same feeder are compared.

Consequently, the next level is the overall LV-grid which leads to the definition of GVR, found by comparing the highest and lowest line-neutral voltage in the LV-grid. These indicators could be relevant to investigate the benefits of a tap changer used at the distribution transformer, as they describe how much of the voltage band is used for distribution but also unsymmetrical loading. The advantages of a tap controller could be limited by a high GVR.

3.2.2 Calculation of voltage ranges

Figure 4 shows the principle of the indicator u_{maxmin}. This indicator measures the dispersion between 'end nodes' of feeders. One case could be that this indicator is as high as u_{max}-u_{min}. The opposite case would be if the voltage drop in all feeders is the same. Then u_{maxmin} would be 0, if the voltages were exactly equal. In figure 4 u_{maxmin} is slightly lower as the voltage drop at feeder 5 (u_{max}-u_{min}).

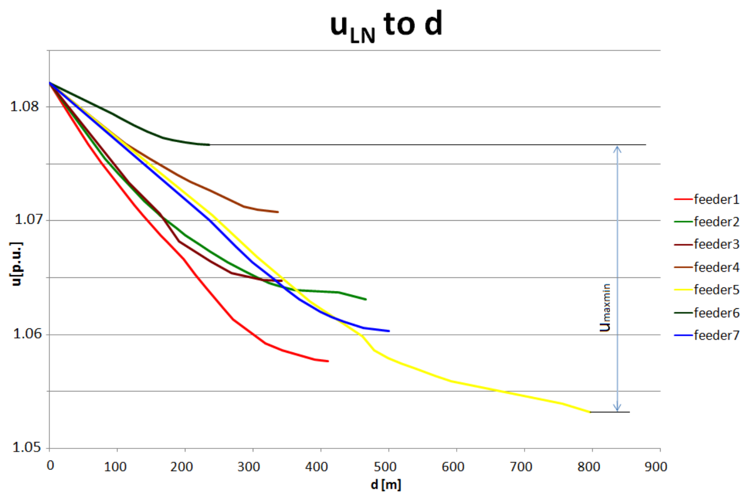


Figure 4: Principle of u_{maxmin}

The discussed indicators can be calculated as follows:

$$NVR = max(|U_{node}{}^{Ph_i} - U_{node}{}^{Ph_j}|, Ph_{i,j} = L1...L3) \tag{3}$$

$$FVR = max(|U_e{}^{Ph_i} - U_f{}^{Ph_j}|, Ph_{i,j} = L1...L3|e, f = 1...N) \tag{4}$$

$$GVR = max(|U_m{}^{Ph_i} - U_n{}^{Ph_j}|, Ph_{i,j} = L1...L3|m, n = 1...M) \tag{5}$$

N...number of nodes in the feeder
M...number of nodes in the LV-grid

NVR...node voltage range
FVR...feeder voltage range
GVR...grid voltage range

$$u_{maxmin} = max(u_{min_i} - u_{min_j}) \tag{6}$$

i,j ... indices for the end node in each feeder

3.3 Equivalent load location ε

The equivalent load location, ε, is used for estimating the voltage drop in networks [10]. ε is the location where the concentration of the total loading of a feeder causes the same voltage drop compared to the end node in normal operating state.

3.3.1 Definition of ε

The equivalent load location can be calculated easily for uniform loading and cable type according to [8].

$$\varepsilon = \frac{1}{I_{tot} \cdot l_N} \cdot \sum_{i=1}^{N} l_{ci} \cdot I = \frac{1}{N \cdot I \cdot l_N} \cdot \sum_{i=1}^{N} l_{ci} \cdot I = \frac{1}{N \cdot l_N} \cdot \sum_{i=1}^{N} l_{ci} \tag{7}$$

l_{ci} ... cumulated length from the distribution node N ... number of nodes in the feeder In figure 5 the principle of ε is illustrated. This indicator shows the visualization of the formula 7.

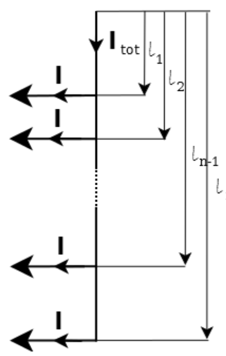


Figure 5: principle for uniform loads and cable type

This indicator makes two simplifications to allow an illustration like in formula 7. In rural areas the cable cross section might be reduced along a feeder. This can be observed in the provided network models. As a result the cumulated length of a section contains different impedances. Therefore the impedances of the section could to be used instead of the distance. This means that the length has to be split up by cable types along the path to the end node. The second simplification of uniform currents inside the feeder can't be used in this work, as the use of PSS data is projected. Instead of a uniform I in formula 7 I_i has to be used and the formula could not be reduced. In summary, a general calculation method to reach a simple formulation like in [10] and [8] becomes more complicated. The aim of this measure is to indicate the location of the equivalent load where the voltage drop would be as high as in the end node. As pointed out before, the length of the cables will not be used itself. In the hitherto definition, the equivalent load could fall to any place along the cable length (between nodes). In the adaptation of this indicator, only nodes will be allowed as load locations.

3.3.2 Calculation of ε

The equivalent load location can be found by simulations in PowerFactory. To find the equivalent load location, the path information from the transformer station to the 'end node' of a feeder is necessarily needed. The total active and reactive power of a feeder is summed up and equally distributed on the phases. In the next step the loads are switched off and an equivalent load is created. This equivalent load is placed initially at the first node of the feeder. Then a load-flow is calculated. The voltage at that node is compared to the voltage of the 'end node' in normal state. If the voltage is higher, the equivalent node will be connected to the next node of the set of the path towards the 'end node'. After that another load-flow calculation is executed. If the voltage is lower than the voltage of the 'end node' in normal state, the algorithm stops and selects the node with the closer value to the normal state as 'ε node'. In conclusion, the equivalent load is placed along the path. After each placement a load-flow calculation is executed and the voltage of the node is compared to the voltage of the 'end node' in normal state.

Formula 8 would indicate the electrical load location:

$$\varepsilon = \frac{Z_{k\varepsilon}}{Z_{k,endnode}} \tag{8}$$

$Z_{k\varepsilon}$ ... short circuit impedance of the equivalent load location node
$Z_{k,endnode}$ ... short circuit impedance of the 'end node'
In theory, also the ratio of other parameters (distance, R or X, etc.) of the nodes could be used instead.

In the analysis on feeder level there can be only one equivalent load location. On network level, there is a equivalent load location for each feeder. Therefore the minimal, maximal and median of ε were implemented for the analysis. Nevertheless, all 3 will only be used if all the information of these indicators are not redundant, which will be proven by a linear regression analysis. The equivalent load location can be calculated for PSS data by symmetrizing the loading of each Smart Meter.

For the 'ε node', also the indicators $Z_{k\varepsilon}$, R_{ε}, $S_{k\varepsilon}$" and the distance to the transformer station could be used as indicators on feeder level. On network level $max(\varepsilon)$, $min(\varepsilon)$, $median(\varepsilon)$, $min(d_{\varepsilon})$, $max(Z_k)$, $median(Z_k)$, $min(S_k")$ and $median(S_k")$ could be analysed. The equivalent load location depends on the loading of the feeder and therefore the location could change with PSS.

3.4 Equivalent sum-impedance

3.4.1 Definition of $\mathbf{Z}_{\Sigma}$

The Equivalent sum-impedance Z_{Σ} (complex) was introduced in [13]. For homogeneous loads and one cable type the formula can be given as:

$$\Delta\underline{U} = \underline{I} \cdot (R^{'} + jX^{'}) \cdot [l_1 + (l_1 + l_2) + ... + (l_1 + l_2 + ... + l_{N-1} + l_N)] \quad (9)$$

N ... number of nodes in the feeder l_i ... distance between nodes A formula for this indicator was also given for different cable diameters in [13], but for uniform loads. With the usage of PSS, non-uniform loads will be considered in a feeder. Therefore this indicator has to be adapted for further purposes. In a first step the formula was abstracted for a general case. In a second step the equivalent sum-impedance can be found by ohm's law.

3.4.2 Calculation of $\mathbf{Z}_{\Sigma}$

The calculation of Z_{Σ} will be explained with an example in figure 6: In this figure a feeder is connected to a grid. The feeder draws a total current $\underline{I}_{tot}$. As there is a branching on node4, the lowest voltage in the feeder could be at the nodes node6 or node7. This depends mainly on the loading and the cable cross section. Assume that the lowest voltage is at node7 (high loading at Load4 and/or small cable cross section of Cable L5).

Formula 10 is obtained:

$$\Delta\underline{U} = \underline{I}_{tot} \cdot \underline{Z}_1 + (\underline{I}_{tot} - \underline{I}_1) \cdot (\underline{Z}_2 + \underline{Z}_3) + (\underline{I}_{tot} - \underline{I}_1 - \underline{I}_2 - \underline{I}_3 - \underline{I}_5) \cdot \underline{Z}_5 \quad (10)$$

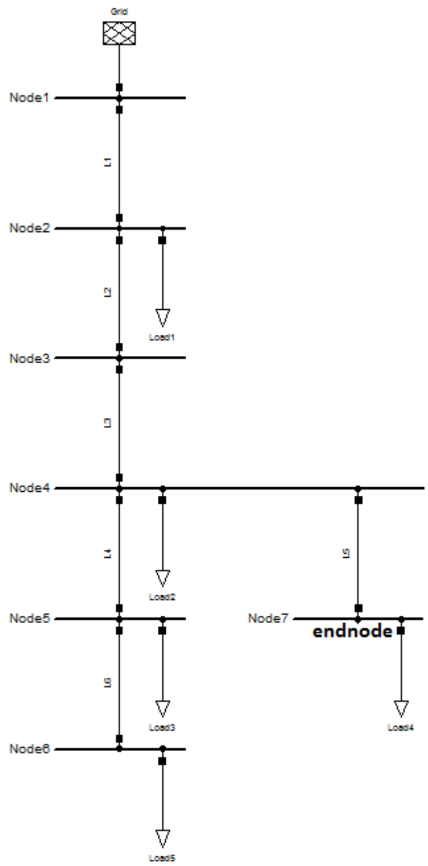


Figure 6: principle for different loadings and cable types

Next, both sides are divided by I_{tot}. Formula 11 is obtained:

$$Z_\Sigma = |\frac{\Delta \underline{U}}{\underline{I}_{tot}}| = |\underline{Z}_1 + (1 - \frac{\underline{I}_1}{\underline{I}_{tot}}) \cdot (\underline{Z}_2 + \underline{Z}_3) + (1 - \frac{\underline{I}_1}{\underline{I}_{tot}} - \frac{\underline{I}_2 - \underline{I}_3 - \underline{I}_5}{\underline{I}_{tot}}) \cdot \underline{Z}_5| \quad (11)$$

The equivalent sum impedance is the quotient of the voltage difference between the transformer station and the end node of a feeder divided by the total current flowing to that feeder. In this formula the impedances of the cables are added with the ratio of the current transported through that specific cable compared to the total current of the feeder. If only load 4 was switched on, the currents $\underline{I}_1$, $\underline{I}_2$, $\underline{I}_3$, and $\underline{I}_5$ would become 0 and the equivalent sum impedance would be $\underline{Z}_1 + \underline{Z}_2 + \underline{Z}_3 + \underline{Z}_5$. The equivalent sum-impedance can be calculated for feeders without a connection to neighbour feeders through switch. In meshed systems this indicator has to be calculated in a different way and the validity of this indicator is not straightforward. In general, the equivalent sum impedance can be calculated for any node of a feeder only by calculating Δu for the selected node.

Again, like the previous indicators, there is only one Z_Σ of each 'end node' (symmetrical case). For the analysis on network level $max(Z_\Sigma)$, $min(Z_\Sigma)$ and $median(Z_\Sigma)$ were implemented. The equivalent sum-impedance could also be computed for snapshot data for all three phases. The node with the lowest voltage on a phase is selected as the end node. With PSS, 3 $Z_{\Sigma A,B,C}$ could be calculated on feeder level. The equivalent sum impedance in the unsymmetrical case can be found as:

$$Z_{\Sigma A,B,C} = |\frac{\underline{U}_{NA,B,C} - \underline{U}_{endnodeA,B,C}}{\underline{I}_{totA,B,C}}| \quad (12)$$

The implementation of the unsymmetrical indicators and their transformation to network level will be decided after the study with sample PSS-data, if necessary. In unsymmetrical conditions the equivalent sum impedance must e carefully interpreted since it does not consider any couplings between phases.

3.5 Number of neighbour nodes

3.5.1 Definition of NON

Another indicator that could be of interest is the number of neighbours of nodes (NON). This indicator could be relevant to distinguish between urban and rural feeders. For example if a node supplies 12 one-family homes each with an own cable and end node, then the number of neighbours of the supplying node would be 13 (12 inferior nodes + 1 superior node). And if all 12 families would live in the same residential building the number of neighbours would be 2 for the supplying node. The number of neighbours of every ‘end node’ is 1. As the feeder ends at some point, the minimal number of neighbours is always 1. Therefore the information of the ‘end nodes’ itself contains no information. The maximal number of neighbours on the other hand can contain information about the area.

3.5.2 Calculation of NON

The calculation of this indicator in DPL runs simultaneously with the calculation of DTN, as the same nodes are handled. In figure 7 a first example can be seen. The nodes are marked again with letters. Node A is the start-point of feeder 3 and has 2 neighbours. The following 2 nodes (B and C) have also 2 neighbours. Only the end node (D) has 1 neighbour.

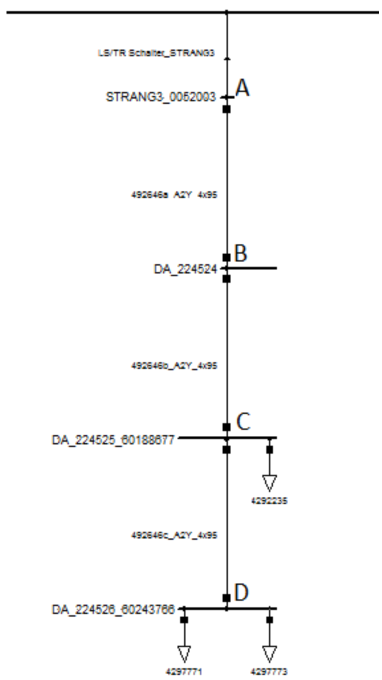


Figure 7: Average number of neighbours example 1

$\overline{NON}$ for the first example can be calculated as follows:

$$\overline{NON} = \frac{2+2+2+1}{4} = 1.75 \tag{13}$$

The values in the numerator can be split up in 3 groups. Node A has always 2 neighbours, and the ‘end node’ (D) always 1 neighbour. The remaining values for nodes B and C can be obtained by multiplying the number of nodes between node A and D by 2. The denominator is the number of nodes. This principle can also be observed on an extended network. If there was another node after node D, NON of node D would become 2. Due to the added node 'E' 1 would be added the numerator and the numerator would become 5.

$$\overline{NON} = \frac{2+2+2+2+1}{5} = 1.8 \tag{14}$$

Again, the principle would explain this equation. From the two examples formula 15 can be derived. For a hypothetical feeder which has an infinite number of nodes (continuous load) the indicator $\overline{NON}$ becomes 2. Therefore the deviation of this indicator from the value 2 contains information about branching and end nodes in a feeder.

$$\overline{NON} = \frac{2+2\cdot(N-2)+1}{N} \tag{15}$$

n... number of nodes

$$\overline{NON} = \lim_{n\to\infty} \frac{2+2\cdot(N-2)+1}{N} = 2 \tag{16}$$

The second feeder example, including branching, can be calculated in a similar way. In total, there are 11 'end nodes' with only 1 neighbour. These nodes are marked with C, E, H, I, J, L, M, Q, R, S and T. Counting from the top of the feeder, there are 5 branching nodes (with number of neighbours): B (3), D (4), G (4), K (5) and O (4). The remaining nodes A, F, N and P have 2 neighbours each. $\overline{NON}$ can be calculated as:

$$\overline{NON} = \frac{11\cdot 1+3+4+4+5+4+4\cdot 2}{20} = 1.95 \tag{17}$$

The obtained formula 18 gives the same result:

$$\overline{NON} = \frac{2+2\cdot(20-2)+1}{20} = 1.95 \tag{18}$$

In general a value of 2 should be the upper limit of this indicator in theory. Nevertheless, as in some LV-networks ring switches are installed, there are 'end nodes' in some feeders with 2 neighbour nodes. Therefore the average could become higher than 2 if the switch status is not considered. This indicator is found by a VBA macro calculating the average of all NON values of the nodes of a feeder.

$$\overline{NON} = \frac{\sum_{i=1}^{n} NON_i}{N} \tag{19}$$

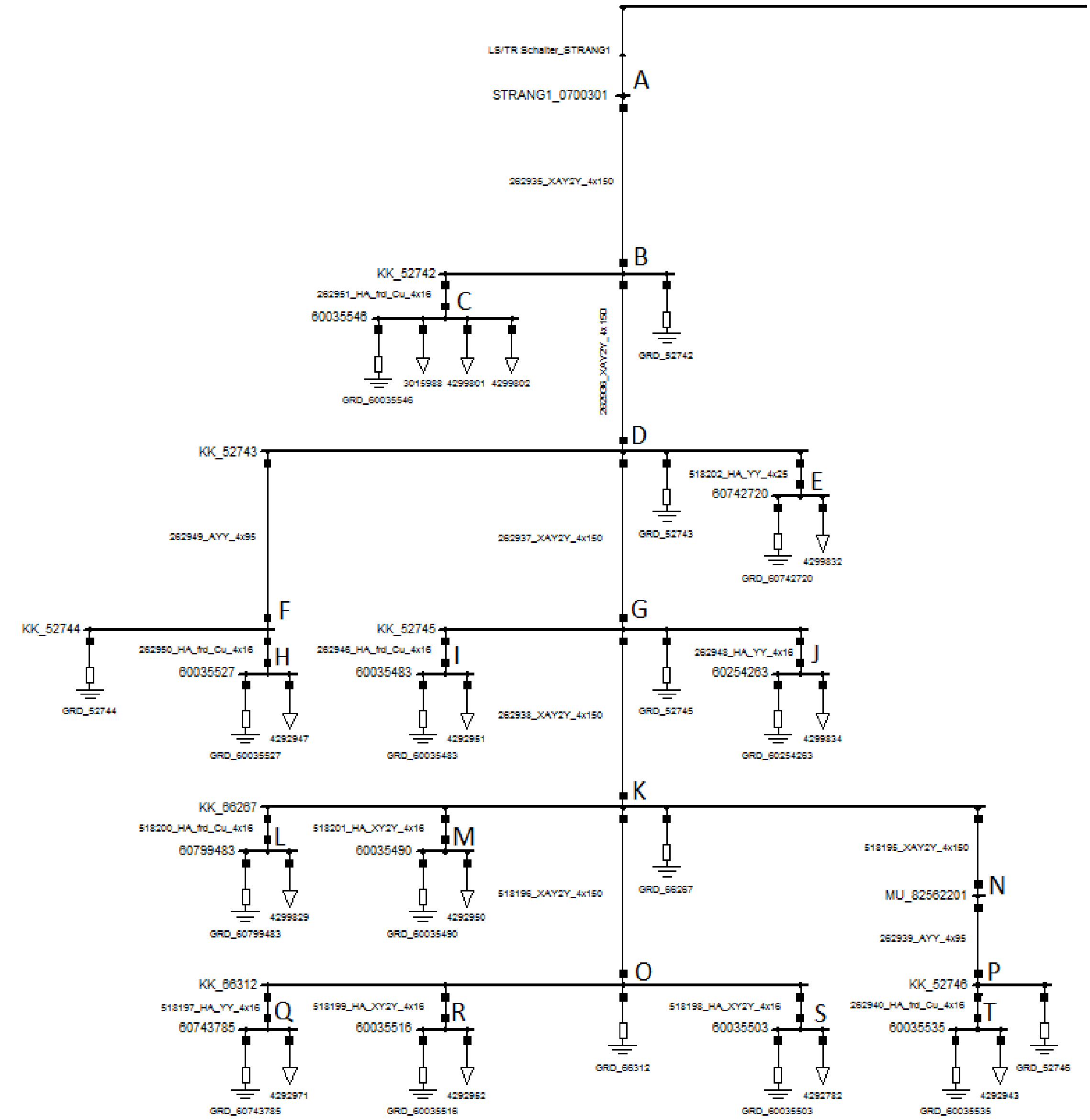


Figure 8: Average number of neighbours example 2

i... node i of the feeder
N... number of nodes of the feeder

On feeder level, $\overline{NON}$, $max(NON)$ and $min(NON)$ could be used as indicators. On network level, $median(\overline{NON})$ of the feeders could be used together with again $max(NON)$ and $min(NON)$. With a script the number of neighbour nodes of each node in a feeder is calculated. In the script all elements with 2 ports that are connected to a node are collected. Then the node on the other port of the connections are stored in a set. The length of the set is equal to the number of neighbour nodes.

3.6 Distance to neighbours

3.6.1 Definition of DTN

The distance to the neighbours contains information that could be important to characterize networks or feeders [13]. The indicators average distance of neighbour nodes ($\overline{DTN}$), maximal and minimal $\overline{DTN}$ will be calculated on feeder level for every node of a feeder. On the higher level of the network $\overline{DTN}$ of all nodes will be used.

3.6.2 Calculation of DTN

A node can have several neighbours. The actual node handled will be named center node, as all distances will be calculated from this node. The distance between the center node and its neighbours is summed up and divided by the number of neighbours. This gives the average distance between the center node and its neighbour nodes. To find the correct number of neighbours, parallel cables have to be modelled as 1 cable (with parallel lines) and not as separate element. The neighbour nodes are accessed by common elements connecting the nodes. If parallel lines are modelled as several single lines, the neighbour node would be counted several times. The distance of a node to the transformer station can be accessed in DPL by O:b:dist [3]. O is an object of type node. To find the distance between two nodes ($node_a$ and $node_b$), the difference of $node_a : b : dist$ and $node_b : bist$ can be used.

The calculation for a node is:

$$\overline{DTN_n} = \frac{1}{NON} \sum_{i=1}^{NON} abs(node_a : b : dist - node_i : b : dist) \tag{20}$$

a ... nodes of a feeder $node_a$ actual node, $\overline{DTN}$ is calculated for
$node_i$... neighbour nodes of $node_a$

NON ... number of neighbours of $node_a$
:b:dist ... distance of a node to the transformer station

The formula on feeder level is:

$$\overline{DTN_f} = \frac{1}{N}\sum_{i=1}^{N}\overline{DTN_n} \tag{21}$$

N ... number of nodes in the feeder
And on network level:

$$\overline{DTN_n} = \frac{1}{M}\sum_{i=1}^{M}\overline{DTN_f} \tag{22}$$

M ... number of nodes in the network
The indices of DTN_x were used to support understanding. This indicator will be only used on feeder and network level, therefore the meant index is unambiguous of the context. Therefore the index will not be used and the indicator will be labelled consistent $\overline{DTN}$.

3.7 Maximal load

3.7.1 Definition ML

This indicator is used to describe how many consumers are connected to a node. For example, in urban areas many consumers could be connected to the same node in residential buildings. In opposite, the number of consumers in rural areas is expected to have lower values due to the predominance of single homes.

3.7.2 Calculation ML

In PowerFactory, all object connected to a node are retrieved and the number of loads is counted. It is equivalent to the load in kW due to the homogeneous distribution. This information is stored in excel. A VBA macro searches for highest value of this indicator on feeder or network level. A formulation of this indicator is shown in formula 23.

$$ML = max(\sum_{a=1}^{k} N_{loads}^{a}) \tag{23}$$

a...actual node k...total nodes of a feeder of network

3.8 Power ratio

3.8.1 Definition of PR

At this point a new indicator shall be introduced, that could be suitable in Smart Grid studies. For that, DER have to be implemented as loads with negative loading, which is already applied within the ISOLVES project.
The power ratio can be defined at every node of the feeder or network. Basically, the total power transported to a node is either consumed by the loads connected to that node or is transmitted through that node to the neighbour nodes in the LV-network. In any case, the total power has to be 0 at every node (Kirchhoff). After a loadflow calculation in PowerFactory, two parameters available at every node are P_{load} and P_{flow}. The first parameter (:m:Pload) is the total active power consumption and/or production at the node by loads or DER. The second measure (:m:Pflow) is the power flowing to the node. In conclusion, the power ratio PR is the quotient of consumed to 'received' or 'transmitted' power. If there is no generation the ratio is between zero and one, where one means that the node is an end node' where the total power is consumed by loads or the following loads are switched off. Zero means that there are no loads connected to the node. If the feed-in at a node is greater than the loading, power is produced at that node and the ratio becomes negative. Consequently the factor -1 would mean that there is only feed-in at a specific node. On both feeder and network level, the median of the nodes of a feeder and the median of the feeders respectively could be used.

3.8.2 Calculation of PR

The power ratio can be simply calculated for any node with following formula 24:

$$PR = \frac{P_{load}}{P_{flow}} \tag{24}$$

As an example consider the feeder in figure 9. In the generic analysis all loads are assumed with 1kW and 0.1kVA. As there are 8 loads in this feeder, the total active power consumed is 8KW. The loading on the nodes can be seen in 3

Consequently, nodes with no loading have a PR of 0 and 'end nodes' have a PR of 1. The only node with loading which is not an 'end node' is node C. The power ratio at this node is 6/7. Nodes without loads always have the same PR for any PSS, as P_{load} is always 0. Additionally, 'end nodes' without DER have always a value of 1, if there is any power consumed at that node. Nevertheless, this value could be 0, if the Snap-Shot was taken

Table 3: Power Ratio

Marker	Node	P_{load}	P_{flow}	PR
		[kW]	[kW]	[1]
A	Strang3_0064703	0	8	0
B	KK_16018	0	8	0
C	60057149	6	7	6/7
D	60057158	1	1	1
E	KK_16590	0	1	0
F	60725193	1	1	1

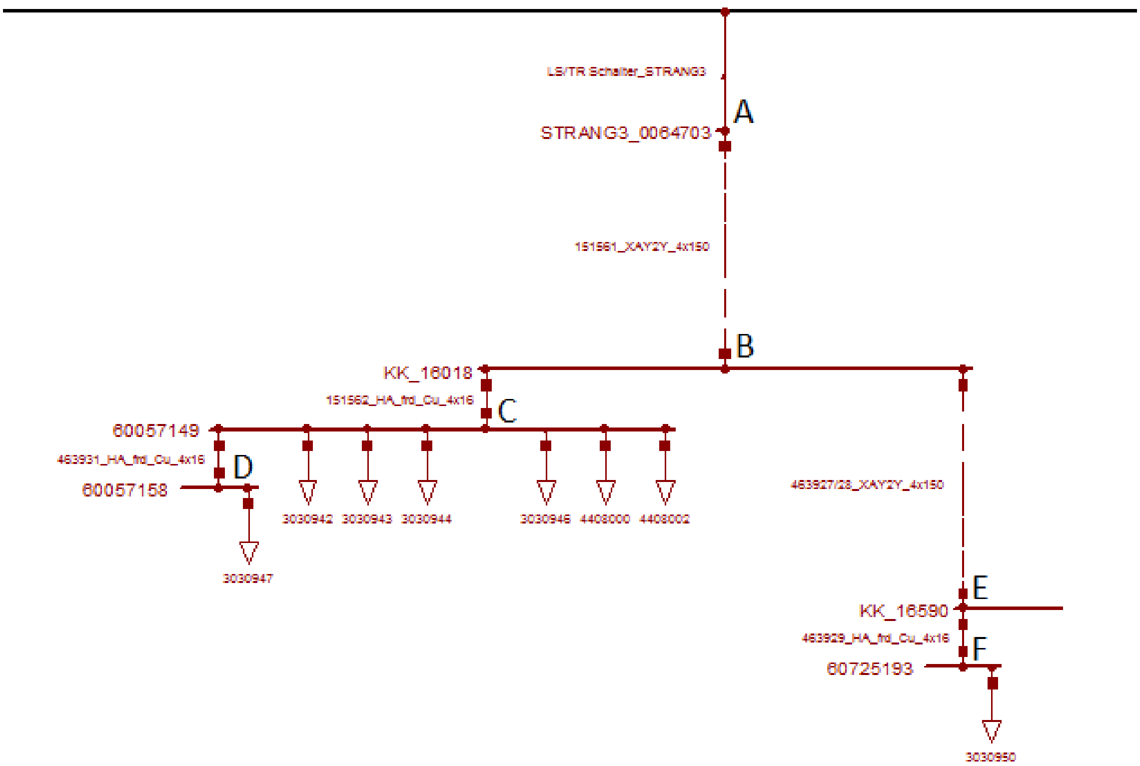


Figure 9: LV-network network 09, feeder 3

at a time when there was absolutely no loading at that node. It is expected, that nodes with a PR of 0 or 1 will not change unless new loads will be installed or a feeder is extended.

3.9 Combination of indicators

Many indicators discussed on feeder level refer to an 'end node'. These indicators on feeder level appear several times on network level. For these indicators minimal, maximal, median or average values could be used. The information content of such combined indicators will be analysed with the generic analysis. The combination of indicators could contain information too. For example in [13] the quotient of the equivalent sum impedance and number of loads was found as characterizing indicator. In this work already introduced indicators were used as well as new indicators that could assist in characterizing or classifying feeders or networks. In [9] the indicator supplied area by a transformer was stated as to be important. An important step is the definition of the supplied area. One definition would be to use cadastral land register data ([6]). However, it can be questioned for example in rural areas for agricultural households with a single family home and attached big fields. With this definition big areas, where no cables are available could be counted to the supplied area. A first reduction of the supplied are would be to use the single home area only, or to draw borders more tightly to exclude such field areas. But also areas between houses that are not supplied still would be counted to the supplied area. Instead of cadastral information, borders could be drawn around the buildings. The process of drawing borders is not defined. Borders could be drawn roughly with a straight baseline border element of a fixed length. The border could be drawn more smooth with a smaller length of the border element. At this point it could be suggested to use the cable length instead of the supplied are, which is easily determinable and available in the computer systems of DNOs. Therefore the cable length was used instead of defining a supplied area for each feeder and network. Accordingly, the compactness as introduced [8] was used. In this work, the total length of the feeder was used together with the total number or loads in the feeder. On network level, the overall cable length was used together with the overall number of loads. This indicator, and the following one will be used, to have comparable indicators on feeder and network level, as they are equally defined on both levels:

$$c = \frac{L_{tot}}{N_{loads}}[m] \tag{25}$$

L_{tot} ... total cable length on feeder or network level
N_{loads} ... number of loads on feeder or network level

Similar to the compactness in formula 25 the load distribution can be calculated. Instead of the total cable length the number of nodes with loads is used. Formula 26 shows the calculation.

$$LD = \frac{N_{nodes}}{N_{loads}}[1] \tag{26}$$

A new indicator could be derived from Z_Σ and ε. On feeder level, the multiplication of these two indicators could be used.

$$ez = Z_\Sigma \cdot \varepsilon \tag{27}$$

As in the beginning of this section already mentioned, [13] introduced the indicator related equivalent sum impedance. On network level, the median of the ratio at feeder levels could be used.

$$Z_{\Sigma-rel} = \frac{Z_\Sigma}{N_{loads}} \tag{28}$$

Finally, the ratio of R to X at the end nodes will be used. On feeder level, there is only one ratio. On network level, the median of all ratios at feeder level will be used.

$$R/X = \frac{R}{X} \tag{29}$$

The R/X ratio of the cluster centres refers to the end node. Therefore the clustering could already indicate the effects of the usage of reactive power to control the voltage drop or rise.

3.10 General indicators

Indicators, generated from simulations in PowerFactory have the disadvantage that the networks have to be modelled in a network simulation software first and have to be held up to date. Therefore electrical or non-electrical indicators that could be already available in the computer systems of the DNOs should be used and analysed too. Indicators on both feeder and network level can be used.

The transformer rating can be used to describe LV-networks [8] as well as the number of customers (N_{loads}) and feeders (N_{feeder}) [13]. Further, in [13], the average distance to neighbours ($\overline{DTN}$), cable length were examined too. The indicator $\overline{DTN}$ was described as useful for the classification of networks.

The cable length as indicator was not found to contain significant information for classifying networks. Nevertheless it will be used in combination with the transformer rating (TR).

1. Transformer rating TR
2. Number of cables (cables) N_{cables}
3. Number of loads (loads) N_{loads}
4. Number of feeders (feeders) $N_{feeders}$
5. Total cable length L_{tot}
6. Number of nodes with loads N_{nodes}
7. Transformer rating/Loads [on network level] $\frac{TR}{N_{loads}}$
8. Transformer rating/Total cable length [on network level] $\frac{TR}{L_{tot}}$
9. Transformer rating/Max. Load [on network level] $\frac{TR}{ML}$

3.11 Summary

In this chapter the principle and the calculation of the indicators were discussed. For the generic analysis 29 indicators will be used on feeder level and 25 on network level. Electrical measures have been selected for the end nodes and the node of the equivalent load location as indicators. Besides, indicators that were already introduced and were identified as describing indicators were taken directly or slightly adaptation. Some of them are load independent and could be labelled as properties of the feeder and network respectively: transformer rating, cable length, $\overline{NON}$, $\overline{DTN}$, maximal number of loads at a single node or N_{loads}. Some indicators are expected to change for different loading conditions. Of course, the voltage ranges Δu and u_{minmax} depend on the loading. At this point the relevance of the indicators for the classification was not discussed. In section 5 these indicators will be examined with a principal component analysis. After that step the relevance of the indicators will be seen and a reduction to a minimum of indicators representing the same information is objected. Analysing only 'properties' of the grid could lead to a similar classification as generic simulations. This would allow to characterize or classify the networks or feeders without much time and effort for utilities in their existing information systems. Nevertheless, clustering by electrical indicators could also result in different groupings if PSS would be used.

4 Descriptive Statistics of the LV-feeders and networks

The indicators discussed in the previous chapter were calculated for all 34 LV-networks with 247 feeders in total in DPL and VBA. In this chapter the number of indicators will be reduced to a minimum. This will be done by using the linear regression model described in 2.4.1. After that, the remaining measures will be analysed on feeder and network level. At the end of this chapter parts of the simulation output from PowerFactory will be presented. Exemplary, the values of the non-electrical indicators on network level can be seen in table 4. As these indicators could be found easily without a network simulation software their usability for a classification should be proved. In [8] p.26 a classification by transformer rating was discussed. Suggested classification levels were 100/160/250/400/630kVA. The corresponding column in table 4 shows that the networks could be classified roughly in 5 groups only using TR. However, the suitability of this parameter alone for classifying networks can be questioned. An answer to this question is provided by the cluster analysis. The values in table 4 have different ranges and units, for this reason each indicator has been normalized to the maximum value among all the feeders or networks. By doing this, the dispersion between feeders can be better compared for each indicator.

Table 4: Network properties

	N_{cables} [1]	TR [kVA]	N_{loads} [1]	$N_{feeders}$ [1]	L_{tot} [m]	N_{nodes} [1]	c [m/1]	LD [1]	$max(Load)$ [1]	$Median(\overline{NON})$ [1]	$max(NON)$ [1]	$Median(\overline{DTN})$ [m]	$max(\overline{DTN})$ [m]	$min(\overline{DTN})$ [m]	$\frac{TR}{Loads}$ [kVA/1]	$\frac{TR}{L_{tot}}$ [kVA/m]	$\frac{TR}{ML}$ [kVA/1]
network 01	137	630	182	9	9992	79	54.90	0.43	13	1.96	7	49.91	457.5	0	3.46	0.06	48.46
network 02	170	400	220	7	6173	129	28.06	0.59	8	2	4	34.89	242	5	1.82	0.06	50
network 03	185	400	240	7	6497	148	27.07	0.62	6	1.97	7	27.49	120	4	1.67	0.06	66.67
network 04	143	630	166	8	9223	89	55.56	0.54	13	1.98	7	61.51	323.33	11	3.80	0.07	48.46
network 05	130	630	179	9	5356	84	29.92	0.47	17	2	6	33.59	140	3	3.52	0.12	37.06
network 06	177	400	224	7	9225	115	41.18	0.51	8	2	6	43.79	444.25	10.5	1.79	0.04	50
network 07	103	250	188	6	3700	76	19.68	0.40	42	1.94	6	26.77	210	11	1.33	0.07	5.95
network 08	89	250	74	7	4842	50	65.43	0.68	8	1.86	4	44.83	333	10	3.38	0.05	31.25
network 09	124	400	120	6	8924	76	74.37	0.63	6	1.91	6	53.29	412.33	2	3.33	0.04	66.67
network 10	116	250	133	6	5179	73	38.94	0.55	8	1.95	7	39.81	121.4	4	1.88	0.05	31.25
network 11	182	400	252	9	8832	127	35.05	0.50	18	1.94	6	41.46	213	11.5	1.59	0.05	22.22
network 12	151	400	220	10	6908	99	31.40	0.45	15	1.95	9	37.83	304	11	1.82	0.06	26.67
network 13	77	400	87	8	4568	49	52.51	0.56	7	1.98	6	54.44	258	17.5	4.60	0.09	57.14
network 14	80	400	117	8	4740	53	40.51	0.45	9	1.90	4	51.20	371.5	11.5	3.42	0.08	44.44
network 15	104	250	106	6	6560	62	61.89	0.58	5	1.90	6	55.80	334.33	7	2.36	0.04	50
network 16	83	400	118	7	4339	51	36.77	0.43	17	2	7	58.08	144.5	16	3.39	0.09	23.53
network 17	96	250	114	8	4860	62	42.63	0.54	11	1.94	6	43.31	173.5	15	2.19	0.05	22.73
network 18	125	400	110	7	11543	71	104.94	0.65	17	1.95	7	69.71	581	0	3.64	0.03	23.53
network 19	87	250	93	5	5350	49	57.53	0.53	6	1.96	5	49.54	315.5	14	2.69	0.05	41.67
network 20	36	160	33	3	3051	19	92.45	0.58	3	1.91	4	71.86	184	14	4.85	0.05	53.33
network 21	136	630	159	9	6028	91	37.91	0.57	8	1.95	9	35.78	242	11	3.96	0.10	78.75
network 22	110	400	176	12	5997	63	34.07	0.36	26	1.94	12	40.43	318	0	2.27	0.07	15.38
network 23	57	630	54	4	3085	43	57.13	0.80	3	1.93	6	40.62	149.5	15	11.67	0.20	210
network 24	114	800	223	11	5627	90	25.23	0.40	12	1.86	9	38.79	380	7.5	3.59	0.14	66.67
network 25	36	400	101	4	1858	21	18.40	0.21	21	2.05	5	47.70	155	5	3.96	0.22	19.05
network 26	147	400	163	5	7294	92	44.75	0.56	9	2	6	50.85	208	14.5	2.45	0.05	44.44
network 27	37	630	31	2	2202	23	71.03	0.74	3	1.74	4	81.91	160	20	20.32	0.29	210
network 28	76	400	84	2	3485	46	41.49	0.55	8	1.99	6	39.69	212.5	14.5	4.76	0.11	50
network 29	173	800	415	12	6534.54	132	15.75	0.32	45	1.85	6	37.06	242	0.01	1.93	0.12	17.78
network 30	128	800	570	11	3890	93	6.82	0.16	11	1.92	5	25.38	309	0	1.40	0.21	72.73
network 31	155	400	184	9	6376	97	34.65	0.53	7	1.95	6	37.57	114	0	2.17	0.06	57.14
network 32	178	400	178	8	9458	117	53.13	0.66	13	1.96	8	42.89	314	0	2.25	0.04	30.77
network 33	153	400	149	9	10005	88	67.15	0.59	8	1.97	5	41.60	491	0	2.68	0.04	50
network 34	76	160	75	6	3658	46	48.77	0.61	6	1.98	5	41.40	264	12	2.13	0.04	26.67

4.1 Indicator selection on the basis of LR

In table 4 one part of the measures was presented. To improve the clustering algorithms, redundant information in the indicator matrices should be reduced to improve clustering results. In a first step, a linear regression model was used to iteratively investigate if one of the indicators can be replaced by a linear combination of the others. The coefficient of determination has been used to decide if the reduction is admissible: if $R^2 > 0.9$, the indicator has been removed from the set. A lower value than 0.9 of R^2 would reduce the information content significantly, which is not targeted. In general, a lower number of indicators would result in better clustering, but less indicators would at the same time reduce some information that could be of relevance. The linear regression model is applied to all 4 indicator matrices (electrical and non-electrical on feeder and network level) in the same way. At first, the R^2 values are calculated for each column. After that the column with the highest R^2 is removed. At this point it is necessary to calculate the residuals for the remaining indicators again, before removing the next one, this process is repeated several iterations. If no more measures with R^2 greater than 0.9 are left, the smallest acceptable set of indicators is obtained. The reduced indicator matrices contain a reduced set of indicators, that can not further reduced.

4.1.1 Linear regression on feeder level

Table 5 shows the residuals for the complete indicator matrix on feeder level. The first indicator, that could be removed from the indicator matrix would be $Z_{k-endnode}$ with a R^2 of 0.99877. The indicator with the lowest R^2 is $Median(PR)$. It is together with Δu the only electrical indicators that are below the threshold of 0.9. On the other side many non-electrical indicators are below R^2=0.9.

Table 5: Linear Regression Model - Feeder (first iteration)

Indicator	ε	N_{loads}	Z_{Σ}	d_{ε}	$Z_{k-\varepsilon}$	$S''_{k-\varepsilon}$	$dvdP_{\varepsilon}$
R^2	0.9677	0.8390	0.9876	0.9550	0.9895	0.9708	0.9892
Indicator	N_{cables}	L_{tot}	ML	N_{nodes}	$d_{endnode}$	$S''_{k-endnode}$	$Z_{k-endnode}$
R^2	0.9826	0.9655	0.6331	0.9670	0.9741	0.9763	0.9988
Indicator	$dvdP_{endnode}$	$dvdQ_{endnode}$	$\overline{NON}$	$max(NON)$	$\overline{DTN}$	$max(\overline{DTN})$	$min(\overline{DTN})$
R^2	0.9984	0.9732	0.7292	0.7096	0.9314	0.8186	0.8300
Indicator	c	LD	$\varepsilon \cdot Z_{\Sigma}$	$Median(PR)$	$Z_{\Sigma-rated}$	R/X	Δu
R^2	0.7685	0.7791	0.9898	0.5629	0.8481	0.8631	0.8791

At the beginning of this section the iteration process was explained. At the end of the iterations, the number of measures was reduced to 18 indicators. The remaining indicators were divided in two groups, electrical and non-electrical indicators. The first group, electrical indicators have to be calculated in PowerFactory. The second group of indicators could be obtained without a detailed network model. Instead, the indicators could be calculated from the already available data in DNOs databases. Table 6 shows the final indicators used for clustering on feeder level and their R^2. As expected, reducing a column effects the others. This can be seen by comparing this table with table 5.

Table 6: Header of reduced indicator matrices - Feeder

electrical	ε	d_ε	$S^{"}_{k-\varepsilon}$	Z_Σ	$Z_{\Sigma-rated}$	$dvdQ_{endnode}$	R/X	$Median(PR)$	Δu
R^2	0.7544	0.8362	0.6289	0.8483	0.7001	0.8634	0.7004	0.4985	0.8189
non-electrical	N_{loads}	ML	N_{nodes}	$\overline{NON}$	$max(NON)$	$max(\overline{DTN})$	$min(\overline{DTN})$	c	LD
R^2	0.8078	0.5923	0.8184	0.6907	0.6865	0.7314	0.4565	0.5322	0.6709

4.1.2 Linear regression on network level

On network level, the pre-selection of the indicators is, as already mentioned, of higher importance as there are only 34 'measures'. Therefore the analysis on feeder level was done first, to obtain the relevance of indicators that are available on both levels. In total 25 indicators were analysed:

Table 7: Linear Regression Model - Network

indicator	N_{loads}	c	L_{tot}	N_{nodes}	LD
R^2	0.9783	0.9931	0.9766	0.9880	0.9877
indicator	ML	TR	$N_{feeders}$	$\overline{NON}$	$\overline{DTN}$
R^2	0.8425	0.9402	0.9475	0.8667	0.9756
indicator	$max(\overline{DTN})$	$min(\overline{DTN})$	TR/N_{loads}	TR/L_{tot}	TR/ML
R^2	0.8875	0.9179	0.9871	0.9935	0.9708
indicator	u_{maxmin}	Δu	$Median(\varepsilon)$	$Median(S_{k\varepsilon}$	$Median(Z_\Sigma)$
R^2	0.8955	0.9040	0.8484	0.9769	0.9411
indicator	$Median(dvdQ)$	$Median(R/X)$	$max(d_{endnode})$	$Median(PR)$	$Median(Z_{\Sigma-rel})$
R^2	0.9333	0.9343	0.9363	0.8530	0.7890

The first row of table 7 for example shows indicators, that are also available on feeder level. On network level the values of N_{loads}, L_{tot}, N_{nodes} are the sum of all values on feeder level of a certain network. Nevertheless, the 3 indicators $\overline{NON}$, $\overline{DTN}$ and $Median(PR)$ had to be calculated again on network level as the average of the feeder indicators would be an average of average values. In the even rows the R^2 values of the indicators with the first linear regression analysis can be seen. Most values are above the threshold, meaning that the lack of redundancy is high. The number of indicators was reduced similar to the reduction on feeder level with a linear regression analysis. The first indicator that was withdrawn was TR/N_{loads} with a value of 0.9935. After that the R^2 values were again calculated for the remaining indicators. Consequently, also on network level a R^2 of 0.9

was used as criteria to withdraw an indicator. The remaining indicators can be seen in table 8. Again, the indicators are separated in electrical and non-electrical indicators. On network level, the highest distance to an 'end node' is part of the reduced indicator matrix. In the generic analysis the 'end node' is always the node with the smallest voltage due to the fact that loads are uniform. With PSS data, the node with the lowest voltage could be at a different 'end node' and the distance to this node would change. Therefore this indicator was also used in the electrical reduced indicator matrix.

Table 8: Reduced Indicator Matrix Header- Network

non-electrical	TR	LD	ML	$\overline{NON}$	$\overline{DTN}$	$max(\overline{DTN})$	$min(\overline{DTN})$	$\frac{TR}{N_{loads}}$	$max(d_{endnode})$
R^2	0.7844	0.7752	0.4987	0.6309	0.8292	0.7748	0.6735	0.8293	0.8038
electrical	$umaxmin$	Δu	$Median(\varepsilon)$	$Median(Z_{\Sigma})$	$Median(dvdQ)$	$Median(R/X)$	$Median(PR)$	$Median(Z_{\Sigma-rel})$	$max(d_{endnode})$
$R\hat{2}$	0.7762	0.8466	0.7738	0.8608	0.8592	0.8474	0.6490	0.6220	0.8038

4.2 Statistical analysis of LV-networks and Feeders

In this section the reduced indicator matrices will be presented and discussed for LV-networks and feeders. The introduced measures have different units and ranges, therefore they will be normalized by dividing each indicator column by the maximum value. This is needed as there are some indicators with a range between 0 and 1 and others with ranges above 1000. The division of every column by its maximum is invariant to the linear regression model. The coefficients β_i would change however, only R^2 was used which is not effected.

4.2.1 Feeder indicator statistics

In figure 10 the boxplot of the electric and non-electric indicators on feeder level can be seen. The central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to 5th and 95th percentiles. Outliers are plotted individually. The measures can be split up in 2 groups. The distribution of the first group is equally in a specific area, for example ε, $Median(PR)$ or LD. The second group of indicators has many outliers and could be used to find them ($min(\overline{DTN})$, or c).

In table 9 the min, median and maximal values of the non-normalized electrical measures can be seen that will be used for the clustering. There are 9 indicators left. The matrix containing the values for all feeders will be named reduced indicator matrix. The ranges were calculated for each indicator.

In table 10 the min, median and maximal values of the non-normalized non-electrical measures can be seen. It can be seen that the values cover a

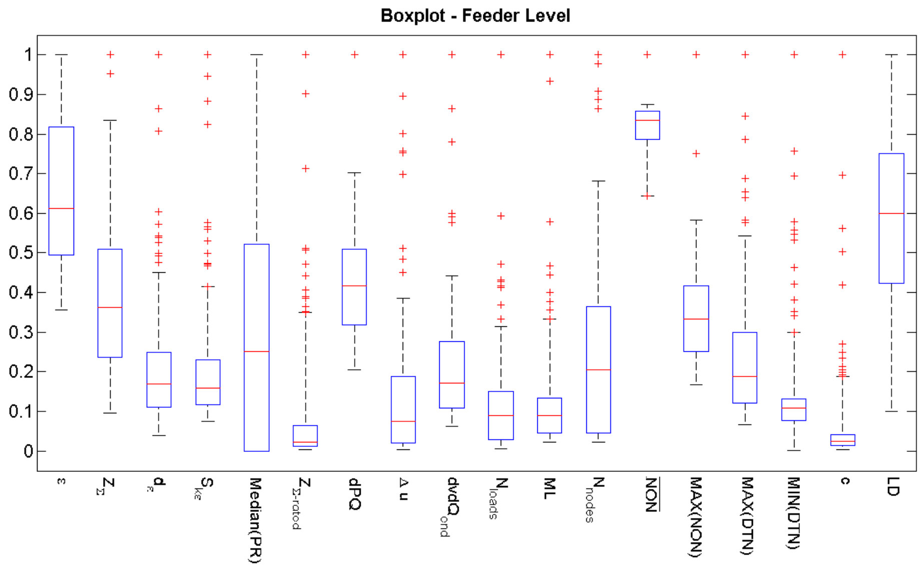


Figure 10: Boxplot Indicators - Feeder Level

Table 9: Ranges of electrical indicators - feeder level

electrical	ε	d_ε	$S^{"}_{k-\varepsilon}$	Z_Σ	$Z_{\Sigma-rated}$	$dvdQ_{endnode}$	R/X	$Median(PR)$	Δu
unit	[1]	[m]	[MVA]	[Ω]	[Ω/1]	[p.u./MVA]	[1]	[1]	[p.u.]
min	0.2439	10	0.6527	0.0005	0.0000	0.0758	0.3720	0	0.0000
median	0.612	242	1.9898	0.0726	0.0041	0.3373	2.1294	0.25	0.0068
max	1	1439	12.5252	0.2006	0.1829	1.9672	5.1036	1	0.0913

wide range. The values were calculated for each indicator, therefore none of the rows is characteristic for a feeder.

Table 10: Ranges of non-electrical indicators - feeder level

non-electrical	N_{loads}	ML	N_{nodes}	$\overline{NON}$	$max(NON)$	$max(\overline{DTN})$	$min(\overline{DTN})$	c	LD
unit	[1]	[1]	[1]	[1]	[1]	[m]	[m]	[m/1]	[1]
min	1	1	1	1.5	2	10	0	0.56863636	0.03846154
median	18	4	9	1.9474	4	109	20.5	43.4333	0.6
max	204	45	44	2.33333333	12	581	190	1776.10004	1

4.2.2 Network indicator statistics

In figure 11 the boxplot of the electric and non-electric indicators on network level can be seen. Again, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to 5th and 95th percentiles. Outliers are plotted individually. Compared to feeder level, there are less outliers. The area of most 25th and 75th percentiles is smaller than on feeder level. Only $min(\overline{DTN})$ has a higher range on network level.

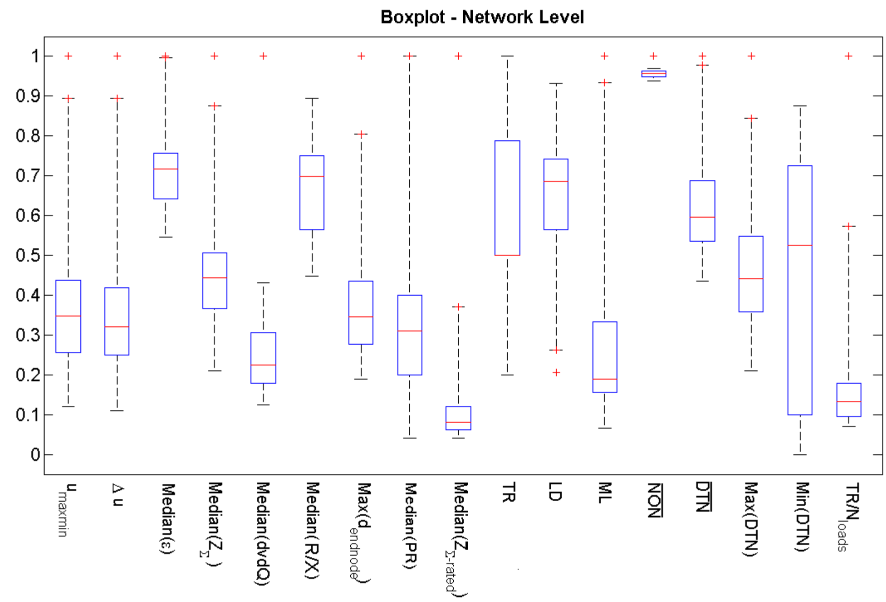


Figure 11: Boxplot Indicators - Network Level

In table 11 the min, median and maximal values of the non-normalized electrical measures can be seen that will be used for the clustering. There are, again 9 indicators left. The matrix containing the values for all networks will be named reduced indicator matrix.

Table 11: Ranges of electrical indicators - network level

	u_{maxmin}	Δu	$Median(\varepsilon)$	$Median(Z_{\Sigma})$	$Median(dvdQ)$	$Median(R/X)$	$max(d_{endnode})$	$Median(PR)$	$Median(Z_{\Sigma-rel})$
unit	[p.u.]	[p.u.]	[1]	[Ω]	[p.u./MVA]	[1]	[m]	[1]	[Ω/1]
min	0.0094	0.0099	0.4566	0.0198	0.1858	1.2608	435	0	0.0007
median	0.0282	0.0293	0.6230	0.0753	0.3449	2.1474	796	0.31	0.0040
max	0.0810	0.0913	0.8677	0.1695	1.5349	3.0781	2307	1	0.0494

In table 12 the min, median and maximal values of the non-normalized non-electrical measures can be seen. These indicators will be used for the clustering.

Table 12: Ranges of non-electrical indicators - network level

	TR	LD	ML	$\overline{NON}$	$\overline{DTN}$	$max(\overline{DTN})$	$min(\overline{DTN})$	TR/N_{loads}	$max(d_{endnode})$
unit	[kVA]	[1]	[1]	[1]	[m]	[m]	[m]	[kVA/1]	[m]
min	160	0.1632	3	1.9090909	26.926889	114	0	1.3297872	435
median	400	0.54575	8.5	1.9609335	45.80484	257	10.5	2.6863679	796
max	800	0.7963	45	2.05	76.729848	581	20	20.322581	2307

4.2.3 Available simulation output

In this section exemplary diagrams obtained from the DPL scripts and VBA macros can be seen for network 01. These diagrams are available for every network.

Figure 12 shows the dependency of the initial short circuit power from the distance. All lines start from the transformer station up to the end node. At Feeder 3, where the lengths of the cables is shorter compared to other feeders, it can be seen that S_k" has a 1/x characteristics due to the constant cable cross-section. For larger cable segments, this observation is distorted since line segments are drawn between nodes. At first sight the special case of feeder 8 is obvious. The reason is that the distance to the customers is longer than usually. To reduce losses and the voltage drop, the voltage is transformed up to about 980V and transformed down again close to the first customers. The vertical lines are the nodes on the primary and secondary side of the transformers with different S_k". The distance between these is 0.

Figure 13 shows the voltage drop from the transformer until the 'end node'. The different gradients are the results of different cable cross sections and reduced currents, as there are branches inside a feeder. In figure 14 the equivalent sum impedance Z_{Σ} can be seen. Z_{Σ} was calculated for each path (transformer station to end node) on every nodes. For the calculation the current entering the feeder was used together with the voltage difference between the transformer station and the actual node. The impedance jump at the transformers in feeder 8 can be seen as vertical lines again. The equivalent sum impedance of feeder 1, 4, 6 and 9 are close to each other,

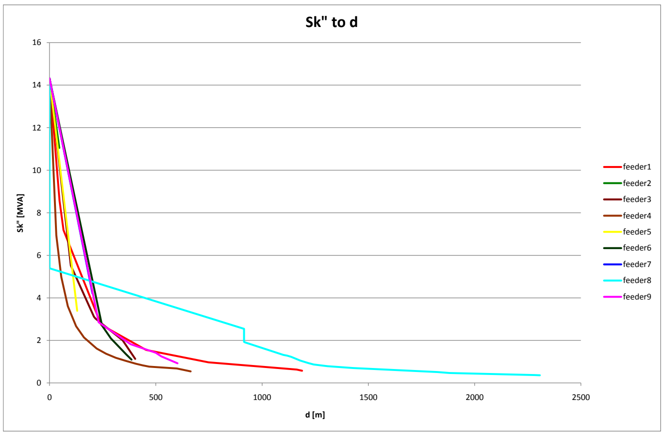


Figure 12: Initial short circuit power

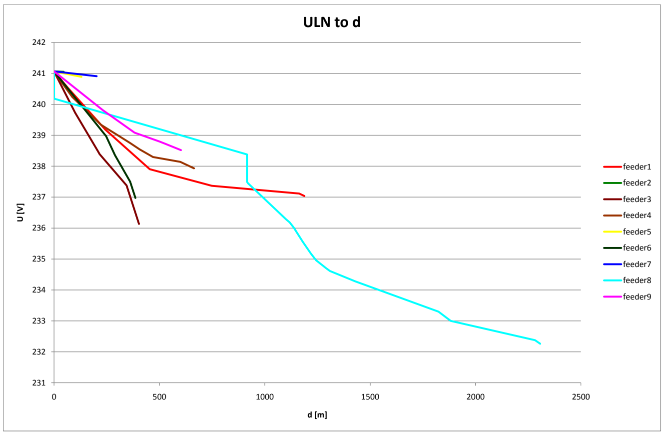


Figure 13: Voltage drop

even the distances are different. This shows a certin level of uniformity. The next figure (figure 15, shows the dependency between Z_Σ andZ_k.

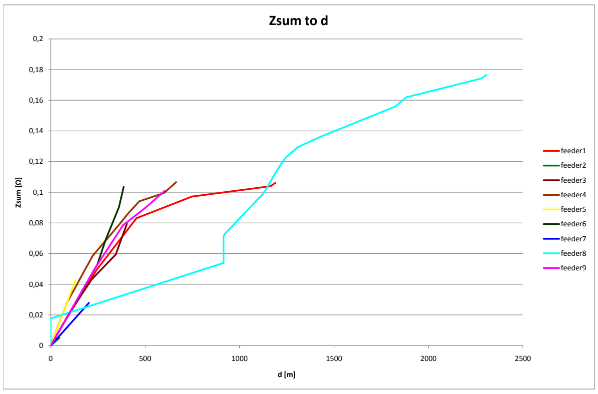


Figure 14: Z_Σ to d

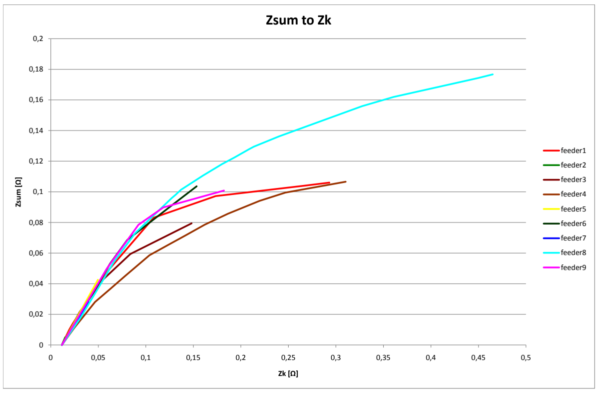


Figure 15: Z_Σ to Z_k

5 Classification of feeders and LV-networks

In the previous chapters a large amount of data was generated with indicators for all LV-networks and feeders. In this chapter, this data will be used to classify LV-networks and feeders. First, the visualization of the data will be discussed and afterwards the process of finding an appropriate cluster size will be addressed. After that the data will be clustered with the selected hierarchical clustering algorithm. In the next step the clusters will be analysed and outliers discussed. The clusters will be marked in diagrams according to table 13. At the end of this chapter 2 Power Snap-Shots (PSS) will be used to analyse 2 feeders of the LV-network of network 01.

Table 13: Colour and marker of clusters

Cluster	Colour	Marker
1	blue	diamond
2	magenta	5-pointed star
3	cyan	plus
4	red	circle
5	blue	asterisk
6	green	six-pointed star
7	black	x
8	green	square
9	magenta	point
10	cyan	square
11	red	square
12	blue	square

The clustering results can be found in appendix A where clusters of each feeder and network can be seen. In appendix B the network diagrams of the selected cluster centres on feeder and network level can be found followed by the network and feeder outliers. In this chapter the indicators values of the mentioned feeders and networks will be presented.

5.1 Analysis on feeder level

In this section the graphical visualization of feeders will be described and the investigations to find the optimal cluster size discussed. Then the clustering results will be analysed. For each cluster the most characteristic feeder (cluster center) will be discussed. A criteria to find outliers will be defined and applied with electrical and non-electrical measures.

5.1.1 Data preparation - Principal component analysis

As described in section 2.4.1, the principal components analysis generates components from the original measures that are orthogonal to each other in order of importance to explain the data variance. There are as many components as original indicators, and each component is a linear combination of all original indicators. However, to visualize the clustering results only the 3 most important components (the first three) will be used. In figure 16 the cumulated sum of the explained variance can be seen. The first 3 indicators explain more than 80% of the variance for both electrical and non-electrical indicators, meaning in a simplified way, that about 20% of the information explaining the variance of the data is omitted if only the first 3 components are used. Figure 17 and 18 show the plotted electrical and

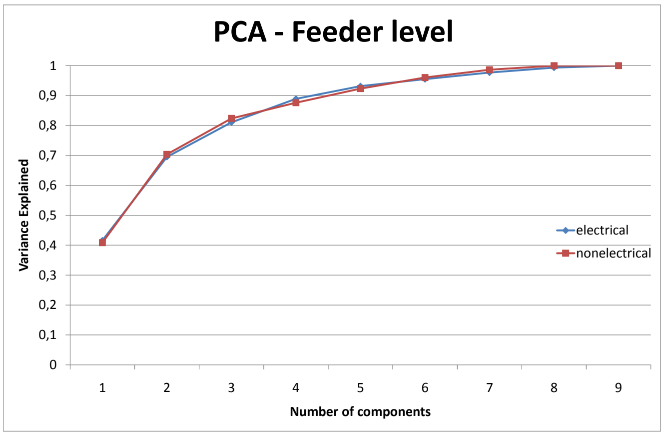


Figure 16: PCA - Feeder level

non-electrical reduced indicator matrices on feeder level. In the figures each point represents a feeder. The lines are the original indicators described in the PCA space. Measures that are parallel to one axis would correspond to the component. Component 1 of the electrical PCA corresponds to PR, the other components are a combination of other indicators. Non-electrical components (1, 2 and 3) correspond to LD, N_{nodes} and $max(\overline{DTN})$. Instead of a single 3D-plot, the points were plotted in 2D for 3 combinations of the 3 most relevant components. In both figures it can be seen that components 1-2 and 1-3 suggest a clustering in at least two groups. The usage of only Components 2-3 would not allow such a grouping. Additionally, it can not be said at the moment, that the similarities in the dimensions 1-2 and 1-3 in both figures result from the same feeders. Nevertheless, this will be analysed after the hierarchical clustering.

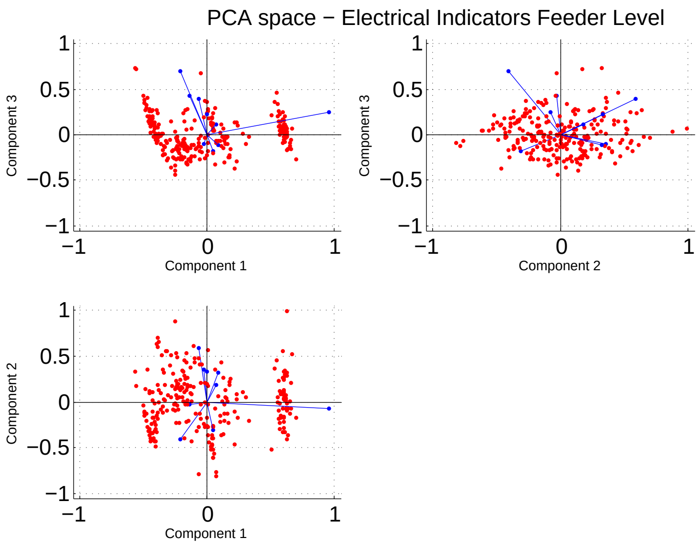


Figure 17: Principal components of electrical indicators

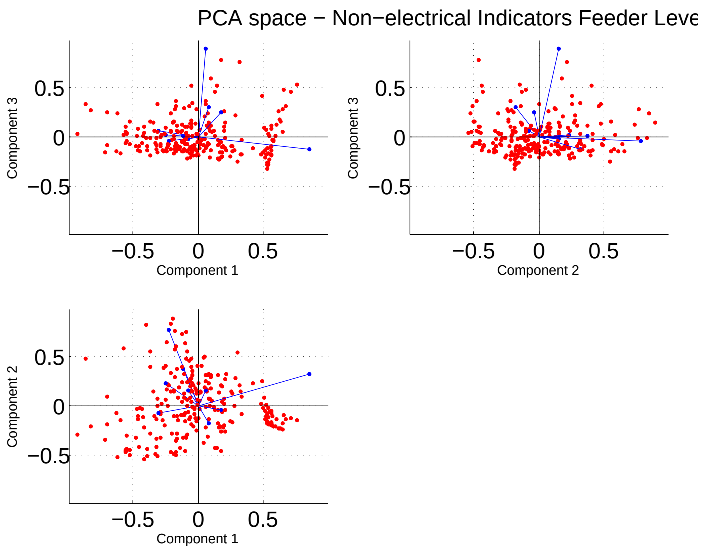


Figure 18: Principal components of non-electrical indicators

5.1.2 Clustering on feeder level

In section 2.4.3 two clustering algorithms were discussed. They have been investigated to find the optimal cluster size. To find the optimal cluster size, the reduced indicator matrices were clustered with the 2 algorithms and a given cluster size. The results were then evaluated with the Matlab function 'silhouette' (see [11], or 2.4.3). Figure 19 shows the results for electrical and non-electrical measures. For electrical indicators, it can be seen that the highest score is found for 2 clusters. However, this solution has only a rating of 0.5 (1 would mean perfect clustering) and 2 clusters is not a reasonable size to cluster the feeders. However there are several local maxima since the score would increase for large number of clusters. Hierarchical clustering has a local maximum for cluster size 5 and 12. K-means has local maxima at 9 and 14. However, defining 'silhouette' as quality criteria would consequently mean to select the cluster size 2, which is from a technical point of view not reasonable. For this reason, this metric to quantify the clustering quality is not suitable (similar results were obtained for the clustering according to non-electrical indicators)

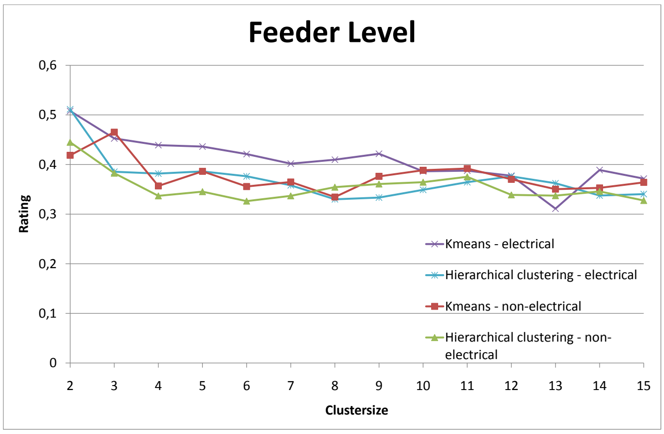


Figure 19: Rated clustering algorithms - Feeder level

As already mentioned, hierarchical clustering will be used to obtain a reasonable cluster size by using a suitable metric. To find an appropriate cluster size 2 distance criteria based on the same metric were analysed. The euclidean distance between two clusters as provided by the hierarchical clustering algorithm is proposed as quality metric. A threshold of 10% and 25% of the total distance were used to find the appropriate cluster size. Figure

20 shows the distance between the feeders in increasing order (circles) after normalization to the maximal distance. The red lines show the 10% and 25% thresholds. The number of clusters can now be found by counting the point above the threshold (and adding 1). Figure 22 shows the results in a dendrogram for the lower threshold and figure 23 for the higher threshold. The lower threshold would result in a cluster size of 30, which is still very large. The higher level is proposed and means that observations (feeders) are grouped within one cluster only of the distance between them is smaller than 25% of the maximal distance between observations. In the dendrogram each cluster is drawn in a different colour. The higher threshold leads to a cluster size for electrical indicators of 9. The same was done for non-electrical indicators. The thresholds can be seen in figure 21 and the visualization in a dendrogram for the threshold in figure 24. The cut in the dendrogram shows that a cluster size of 11 was found for non-electrical indicators.

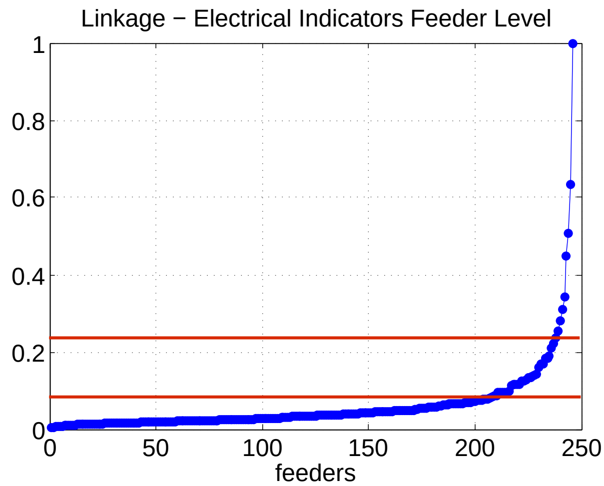


Figure 20: Linkage Thresholds - Electrical indicators

The cluster size 9 was therefore obtained from the clustering with electrical indicators. The results can be seen in figure 25 using the first 3 PCA components and the markers according to table 13. The circles will be used to illustrate outliers and will be explained in the next section. It can be seen that the feeders are at different positions depending on the used components. In all figures the same feeder was selected as center of the circles.

Next the feeders were clustered by non-electrical measures. Again it can be seen that the position of the feeders changes depending on the components. In the figure concentric circles were drawn to illustrate outliers. The figure shows, that feeders of the same clusters have different distances to each other

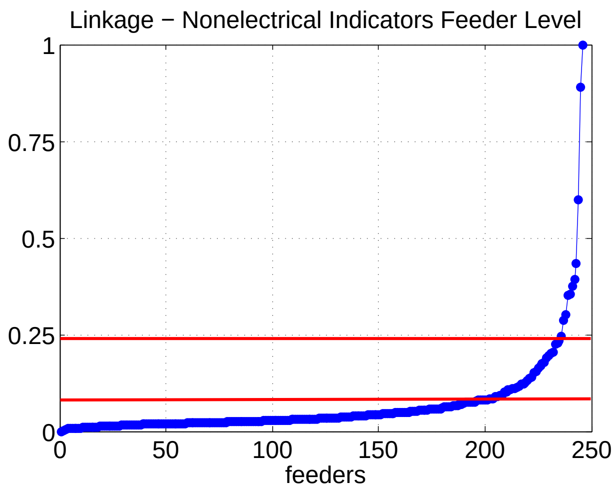


Figure 21: Linkage thresholds - Non-electrical indicators

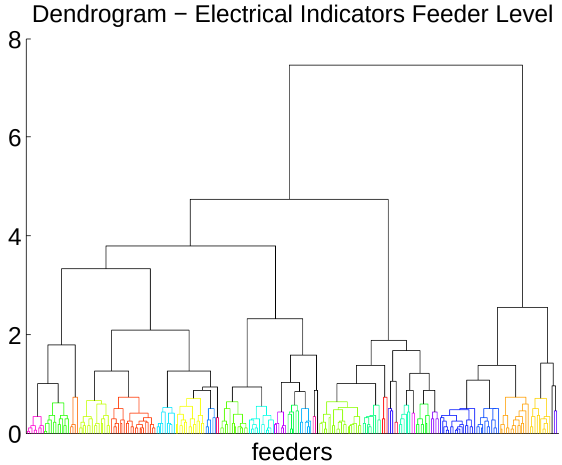


Figure 22: Linkage thresholds= 10% - Electrical indicators

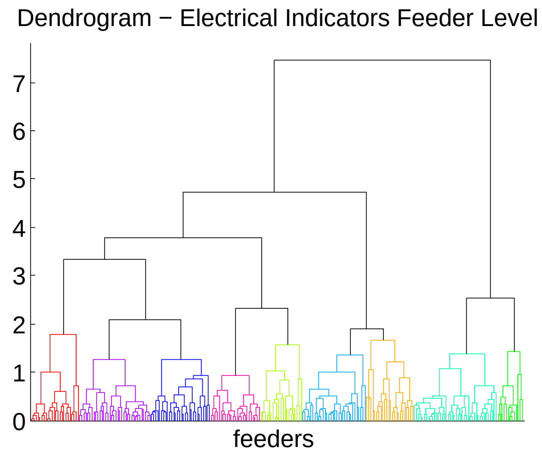


Figure 23: Dendrogram threshold=25% - Electrical indicators

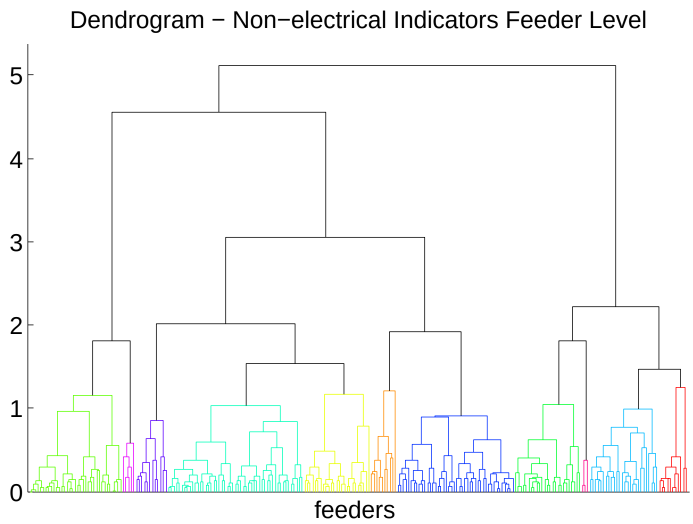


Figure 24: Dendrgoram threshold=25% - Non-electrical indicators

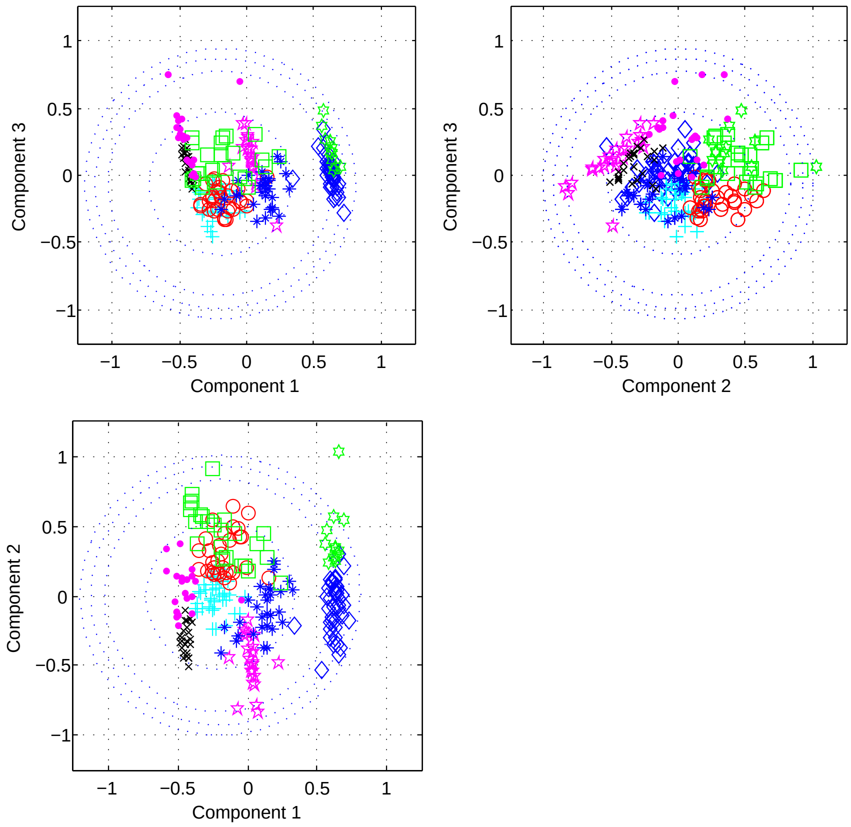


Figure 25: Clustered feeders by electrical indicators

depending on the principal component 2D-view .

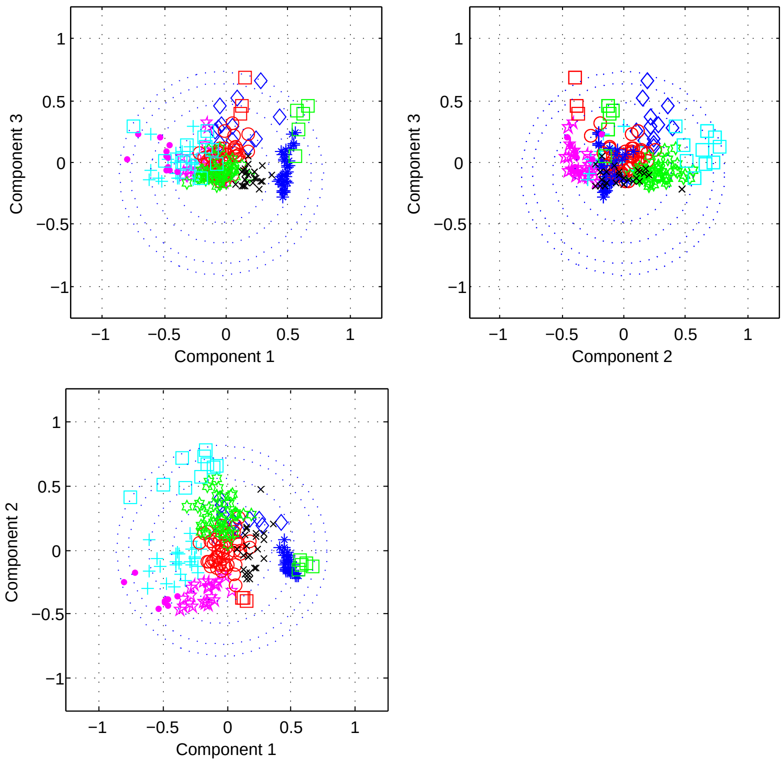


Figure 26: Clustered feeders by non-electrical indicators

To describe the clusters, the center of each cluster needs to be identified. To find the most characteristic feeder of a cluster, the median of each electrical indicator column of all cluster members was calculated. Then the feeder with the most similar (using euclidean distance) indicator values was selected as cluster center. The cluster centres can be seen in table 14. The most similar feeder was found by calculating the distance with the function 'pdist'. The feeder with the smallest distance to the hypothetical median feeder was selected. In the following, the clusters will be shortly discussed, the network diagram of each cluster is shown in appendix B. Due to the high number of feeders, this analysis only considers and compares cluster centres. In the appendix the network diagram of each feeder can be seen. Additionally the available simulation output for the selected feeders can be seen in figures 28, 29, 30 and 31. Additionally the sensitivity dvdP/dvdQ along the feeders can be seen in figure 32.

In figure 27 the used cable technology of each cluster can be seen. There are feeders with only underground cables, overline cables and feeders with both types of cables. It can be seen that overline cables and mixed cables are concentrated in the cluster 4. Additionally, the mixed cables in cluster 4 are mostly overline cables, therefore this cluster can be seen as feeders with overline cables. In table 14 it can be seen that cluster 4 has a high Z_{Σ} and at the same time a low R/X. In the other clusters mostly underground cables are used.

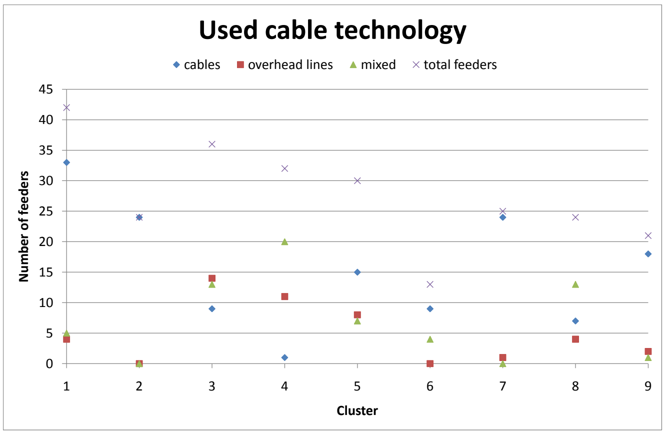


Figure 27: Used cable technology

Cluster 1

This cluster is characterized by branching and an average load density. In this feeder only underground cables are used and therefore R/X is large. Additionally, the service cable at the end of the feeder has a cable cross section of only $16mm^2$. In figure 32 it can be clearly seen, that R/X at the end of the node has a significant gradient change due to the mentioned reduction. The equivalent load location has an average value and also the short circuit power and the equivalent sum impedance. The voltage drop gradient is higher at the beginning due to the branching.

Cluster 2

The second cluster has a rather short distance (200m). The power is only distributed to a single node with few loads. Therefore the equivalent load location is 1. The strong sizing (there is only cable with a cross section of

$240mm^2$) is leading to the smallest equivalent sum impedance and voltage drop and the highest initial short circuit power. Due to the small cable impedance compared to other feeders, the 'end node' has the smallest R/X ratio.

Cluster 3

The third cluster can be characterized by 2 load concentrations in the middle and the end of the feeder. The equivalent load location is 0.54 shortly after the middle of the middle of the feeder and slightly higher than cluster center 1, although there is no branching of the feeder. At the same time, the first cable segment has a larger cross-section. The feeder length has an average value (about 300m - as cluster 1). In this feeder cables and overhead lines are used. The equivalent sum impedance and the voltage drop (%0.8) are comparable to cluster 1. Only the median of the power ratio (0.1) is lower compared to cluster 1 (1) due to the fact that this feeder is mixed.

Cluster 4

The fourth cluster has an average density and a equally distribution of loads in the feeder. There are several branching. Due to the branching in the first part of the feeder and to the reduction in the cable/line cross section the smallest epsilon appears in this feeder (0.45). The voltage diagram and the sensitivity diagram show both a first steep curve (before branching), and then a steep again due to the change of cross-section (larger voltage drop gradient) and R/X increases. This feeder has a small initial short circuit power at the 'end node' (about 800kVA second smallest among cluster centres) combined with a high number of loads (second highest number of loads) leads to the second largest voltage drop among cluster centres (almost 1.8%) and a corresponding a large equivalent sum impedance (second highest). This is a 'strongly loaded' feeder in terms of the used voltage band.

Cluster 5

The fifth cluster can be characterized by the high load density (two nodes with in total 25 loads). It can be characterized as residential area where the loads are predominantly located at the end of the feeder which results in an equivalent load location of 0.6. The voltage drop and the initial short circuit power have average values compared to the other cluster centres. The compactness has the smallest value due to the high number of loads and rather short cable length. In this feeder the same voltage drop occurs as in cluster 1 and 3 but on a significantly shorter distance.

Cluster 6

This cluster contains mostly outliers, (see appendix and table). However, this feeder itself is not an outlier. It has an average load density and the largest voltage drop (almost 3% and a large equivalent sum impedance). In the feeder cables and overhead lines are used and branching can be seen in the network diagram in the appendix. In the figure 29 the steepest segment of the diagram shows a reduction of the line cross section to $50mm^2$. It has the highest number of loads and rather weak cables/lines. In terms of voltage band, this cluster center is also a ‘strongly loaded feeder’.

Cluster 7

The seventh cluster center has a similar structure of the center of cluster 2 but with a slightly higher equivalent sum impedance and a voltage drop due to the service cable at the end of the feeder. This feeder is significantly smaller (about 100m) compared to cluster center 2. The equivalent load location is at the ‘end node and the R/X ratio is significantly larger than for the center of cluster 2. This feeder is characterized by distribution to only one node.

Cluster 8

This cluster has a homogeneous load distribution and is also a cluster with mixed cables and overhead lines. Despite the branching at the beginning of the feeder, the equivalent load location is 0.6 due tot the fact that the cross-section is reduced at a rather far point from the beginning. The voltage drop is the second highest of the selected centres (2.2%) and this feeder has the highest equivalent sum impedance. This is the longest and most poorly sized feeder (because of Z_Σ) considering the number of loads and has a rather large R/X ratio.

Cluster 9

The last cluster center has a low load density. It is similar to the center of cluster 2 and 7 but significantly longer (almost 500m). On the other side it has a significantly larger equivalent sum impedance and is similar to cluster center 6 and 8 (both poorly sized). This cluster is characterized by a very small load supplied by a long feeder and is therefore a poorly sized feeder with a small voltage drop due to the very low number of loads.

Table 14: Cluster centres by electrical indicators

		ε	Z_Σ	d_ε	$S_{k\varepsilon}$	$Median(PR)$	$Z_{\Sigma-rel}$	R/X	Δu	$dvdQ_{end}$	Nr. feeders
	Feeder	[1]	[Ω]	[m]	[MVA]	[1]	[Ω/1]	[1]	[$p.u.$]	[$p.u./MVA$]	[1]
1	network 31 feeder 4	0.5476	0.0621	249	2.3566	1	0.0025	2.1393	0.0090	0.3259	42
2	network 01 feeder 7	1	0.0279	202	4.2046	0.5	0.0070	1.0584	0.0006	0.1661	24
3	network 02 feeder 2	0.5849	0.0735	286	1.9471	0.1	0.0039	2.6310	0.0085	0.3225	36
4	network 08 feeder 2	0.4575	0.0829	232	1.7183	0.12	0.0023	1.7260	0.0175	0.6390	32
5	network 05 feeder 9	0.6103	0.0546	201	3.2351	0.5	0.0022	2.6719	0.0081	0.1777	30
6	network 10 feeder 1	0.5730	0.1173	448	1.1734	1	0.0029	2.6468	0.0284	0.5478	13
7	network 30 feeder 2	1	0.0371	113	3.8103	0	0.0124	1.7696	0.0007	0.1406	25
8	network 15 feeder 1	0.6408	0.1577	683	0.9821	0.2	0.0066	2.5840	0.0228	0.5908	24
9	network 15 feeder 5	0.8290	0.1181	483	1.3232	0	0.0591	2.0459	0.0014	0.3988	21

In table 15 the cluster centres by using non-electrical measures can be seen. The cluster centres were found by the same method used for the cluster centres of table 14. The analysis with non-electrical measures could be seen as simpler clustering because the power consumption and the network

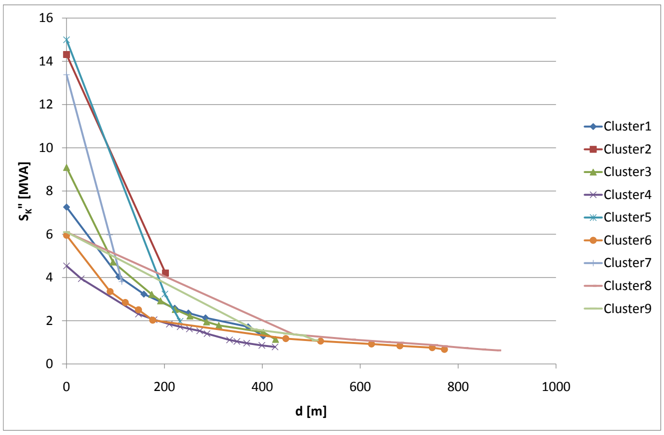


Figure 28: Initial short circuit power

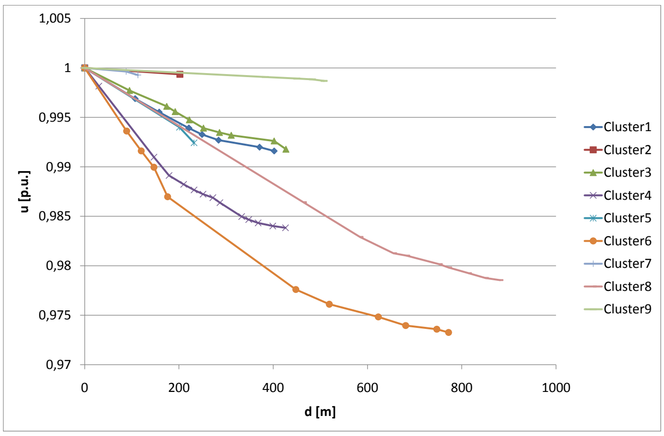


Figure 29: Voltage drop

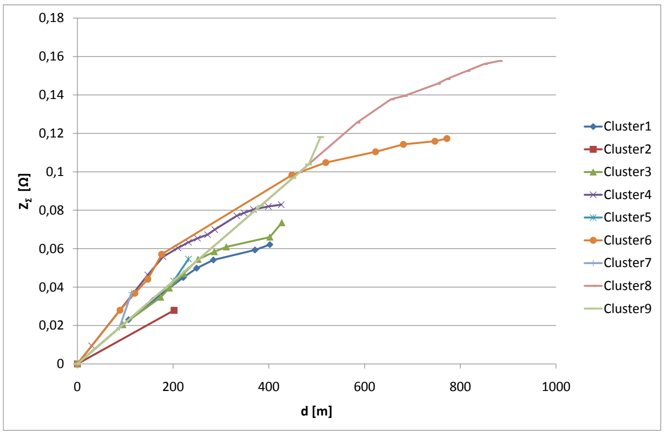


Figure 30: Equivalent sum impedance to distance

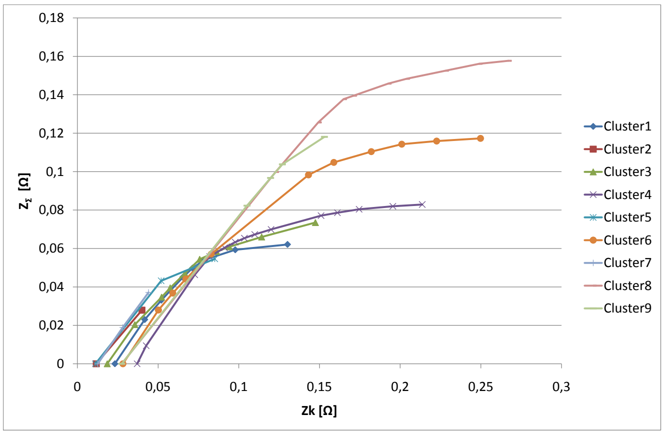


Figure 31: Equivalent sum impedance to short circuit impedance

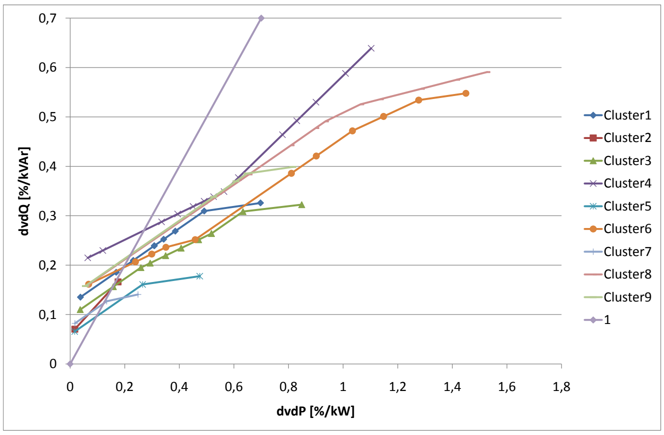


Figure 32: Sensitivity

topology are not considered. **cluster 1 (non-electrical)** The first cluster center, network 04 feeder 6, has many 'end nodes' and the loads are mainly at these. There are only 12 feeders that can be described similar. In these feeders there are no nodes with names starting with 'DA_', which means that no overhead lines are used. Therefore this cluster describes a rural area. The feeders of this cluster were based on the electrical clustering in the clusters 1, 6 and 8. In this cluster there are 4 electrical and 2 non-electrical outliers. The outliers will be discussed in the next section.

cluster 2 (non-electrical) The second cluster (network 31 feeder 5) is similar to the cluster centres 2 and 7 of the electrical clustering and indeed, both centres are part of this cluster. The cluster 2 of non-electrical clustering mostly contains feeders from cluster 2 and 7 of the electrical clustering and several from clusters 1 and 3. In network 31 feeder 5 the loads are also at the end of the feeders. Here also cables are used. In this cluster there are 3 electrical outliers: network 01 feeder 2, network 29 feeder 10 and 11. This cluster group is characterized by a low load density compared to the other cluster centres. There is only one feeder from electrical cluster 9 that was assigned to this cluster (network 19 feeder 4). Compared to cluster 1 a lower compactness value can be seen together with a lower number of loads. Therefore the supplied area is smaller. The total cable length is nearly 255m. This cluster describes feeders that supply several customers that are next to each other and about 250m away from the transformer station. $max(\overline{DTN})$ has the lowest value of all centres. 25 feeders were grouped in this cluster.

cluster 3 (non-electrical) The third cluster has the second highest number of loads. The cable length of network 01 feeder 3 is above 1000m. The highest number of loads at a node is 13 which means that apartments are supplied. Still the cable length suggest that this feeders are mostly in rural areas and additionally supply single homes. The 26 feeders of this cluster were in the following groups of the electrical clustering: 1, 2, 3, 5 and 8.

cluster 4 (non-electrical) The next cluster center (cluster 4) is network 20 feeder 1 for 51 feeders. It is characterized by a medium number of loads, a low ML and a cable length of slightly more than 1500m. Comparing the number of loads and the compactness of this cluster with cluster 1 leads to the conclusion that this cluster could be a down-scaled version of cluster 1. Feeders in this cluster were categorized in the electrical analysis mostly in groups 1, 3, 4 and 5.

cluster 5 (non-electrical) Cluster 5 is characterized by 1 single load at the end of the feeder. This group has a LD of one. This means for the median, that the load is 165m from the distribution node away and there are no branchings in the feeder. This group is mostly formed by feeders of the electrical clusters 2, 7, and 11. This group has only 2 electrical and 2 non-electrical outliers and consists of 35 feeders.

cluster 6 (non-electrical) The next cluster contain 44 feeders and network 03 feeder 7 was selected as median. This feeder is characterized by the highest number of loads and a low ML. This means that there are many nodes and that the supplied buildings are mostly single customers. This cluster has similarities with clusters 1, 3 and 4. In many feeders overhead lines can be found, but not only. The main difference to cluster 4 is that even the cable length is longer, $max(\overline{DTN})$ is shorter.

cluster 7 (non-electrical) network 05 feeder 4 (cluster 7) was selected as median of 25 feeders in cluster 7. The feeders of this classification were in the electrical clustering mainly in the groups 1, 3 and 5. This can be seen in the appendix B. Therefore this cluster could be compared with clusters 4 and 6. ML, $max(NON)$ are similar for all 3 cluster centres. However, the number of loads of cluster 4 is two times higher and of cluster 6 even more than 4 times. Multiplying the number of loads with the compactness gives a total cable length of 320m. This is half the length of cluster center 4. Therefore cluster 7 can be seen as down-scaled version of clusters 4 and 6. More than 50% of the feeders in this cluster have overhead lines.

cluster 9 (non-electrical) network 05 feeder 9 was also selected as cluster center as in the electrical case. This cluster contains 3 feeders of network 16 (2, 3, and 44) and 4 feeders of network 22(2, 3, 9, 11). This cluster supplies buildings with a high number of customers. This is indicated by ML and N_{nodes}. The cluster center has a total cable length of 250m. There are 11 feeders in this cluster. Half of the feeders were in group 7 of the electrical clustering.

cluster 8, 9 and 11 (non-electrical) - outliers Clusters 8, 10 and 11 (network 09 feeder 6, network 09 feeder 5 and network 12 feeder 2) are all outlier clusters. In the dendrogram in figure 24 these clusters are the purple, orange and pink ones. Cluster 8 contains of feeders that were in cluster 9 of the electrical clustering. Again this cluster contains of feeders with mostly 1 load at the end of the feeder, or several feeders each at an ‘end node’. The total cable length can be directly seen, as there is only 1 load and is equal to c (1237m). Most feeders in cluster 10 were in cluster 4 and 6 of the electrical clustering. The feeders of cluster 11 were all part of cluster 2 of the previous clustering. network 12 feeder 2 has many low indicator values. only $Min(DTN)$ is much higher than in the other cluster centres.

Table 15: Cluster centres by non-electrical indicators

		N_{loads}	ML	N_{nodes}	$\overline{NON}$	$max(NON)$	$max(\overline{DTN})$	$Min(\overline{DTN})$	c	LD	Nr. feeders
	Feeder	[1]	[1]	[1]	[1]	[1]	[m]	[m]	[m/1]	[1]	[1]
1	network 04 feeder 6	23	3	15	1.9630	6	323.33	25	139.6522	0.6522	12
2	network 31 feeder 5	6	4	2	1.8	3	76.33	25	42.3333	0.3333	25
3	network 01 feeder 3	44	13	11	2	6	106	25	25	0.2500	26
4	network 20 feeder 1	16	3	9	1.9474	4	173.5	25	94.5	0.5625	51
5	network 17 feeder 2	1	1	1	1.6667	2	82.5	25	165	1	25
6	network 03 feeder 7	35	4	23	1.9677	5	118	8	32.9143	0.6571	44
7	network 05 feeder 4	8	3	6	2	4	94	19	39.8750	0.75	25
8	network 09 feeder 6	1	1	1	1.6667	2	315.5	25	1237	1	5
9	network 05 feeder 9	25	17	2	1.75	3	100.5	27	10.36	0.08	11
10	network 09 feeder 5	64	6	43	1.9851	6	269	2	69.9687	0.6719	10
11	network 12 feeder 2	2	2	1	1.5	2	288	144	144	0.5	3

5.1.3 Outliers on feeder level

In this section the outliers will be identified. The 90 and 95 percentiles outliers will be shown and the 95% outliers discussed. To find a comparable static distance to the feeders, the following approach was selected: First, the median of each indicator column was calculated to find the cluster median. To calculate the global median, all feeders with either their electrical or non-electrical property were used. Next, the feeder with the smallest distance to the calculated hypothetical median was selected as global median. Then the distance of all feeders to this feeder was calculated and the 50, 75 90 and 95 percentiles calculated. Finally the feeders with distances above these levels were found (see circles on figures 25 and 26). The result of this approach can be seen in table 16. Figures 25 and 26 show the principle of finding outliers. As these figures were plotted by using the first 3 PCA components, only 80% of the variance can be seen. If a point is outside of any drawn circle in any 2D-combination of components, it becomes an outlier of the specific percentile. network 30 feeder 5 is characterized by a single branching with a short path and a longer path. Along the path to the ‘end nodes’ the loads are distributed equally. network 05 feeder 8 is characterized by more branchings but again equally distributed loads along the paths. The supplying line is split up in 2 paths and these again in 2 and 3 paths. There are 5 'end nodes'. Both feeders can be found in the appendix.

Table 16: Overall median electrical/non-electrical feeders

	ε	Z_Σ	d_ϵ	$S_{k\epsilon}$	$Median(PR)$	$Z_{\Sigma-rel}$	R/X	Δu	$dvdQ_{end}$
Median Feeder - Electrical indicators	[1]	[Ω]	$[m]$	$[MVA]$	[1]	[Ω/1]	[1]	$[p.u.]$	$[p.u./MVA]$
network 30 feeder 5	0.6811	0.0747	271	2	0.185	0.0034	2	0.0106	0.4366
	N_{loads}	ML	N_{nodes}	$\overline{NON}$	$MAX(NON)$	$max(\overline{DTN})$	$MIN(DTN)$	c	LD
Median Feeder - Non-electrical indicators	[1]	[1]	[1]	[1]	[1]	[m]	[m]	[m/1]	[1]
network 05 feeder 8	20	4	11	2	4	90	16	30	0.55

The outliers by the highest percentiles can be seen in table 17.

Table 17: Feeder outliers by electrical indicators

outlier criterion		electrical outliers	ε	Z_Σ	d_ϵ	$S_{k\epsilon}$	$Median(PR)$	$Z_{\Sigma-rel}$	R/X	Δu	$dvdQ_{end}$
90%	95%	feeder	[1]	[Ω]	$[m]$	$[MVA]$	[1]	[Ω/1]	[1]	$[p.u.]$	$[p.u./MVA]$
X	X	network 01 feeder 2	1	0.0050	46	11.0496	0.5	0.0025	0.5614	0.0000	0.0814
X	X	network 01 feeder 8	0.4162	0.1767	1243	0.8681	0.08	0.0050	2	0.0381	1.1786
X		network 03 feeder 5	0.5011	0.1284	461	1.2085	1	0.0039	3.0562	0.0250	0.5215
X	X	network 04 feeder 6	0.4590	0.1639	758	0.9542	1	0.0071	3.1457	0.0231	0.7241
X		network 07 feeder 1	0.6165	0.1123	420	1.4099	1	0.0017	2.3987	0.0466	0.4765
X	X	network 09 feeder 5	0.3627	0.1780	869	0.8175	1	0.0028	2	0.0731	1.9672
X		network 11 feeder 8	0.5271	0.1296	498	1.1668	1	0.0041	3.6388	0.0257	0.4574
X		network 12 feeder 7	0.5217	0.0160	42	6.6342	1	0.0005	1.8176	0.0032	0.1442
X	X	network 14 feeder 2	0.8667	0.1829	718	0.9685	0	0.1829	2.2273	0.0010	0.4747
X	X	network 18 feeder 5	1	0.1649	441	0.9552	0	0.1649	3.0781	0.0010	0.3277
X	X	network 18 feeder 7	0.6922	0.1911	578	0.8330	1	0.0119	3.7201	0.0187	0.4694
X		network 20 feeder 3	0.4566	0.1911	459	0.8288	0	0.0174	1.2573	0.0127	1.6986
X		network 21 feeder 4	0.2966	0.0861	338	2.0375	1	0.0036	3.6264	0.0125	0.4508
X	X	network 23 feeder 1	0.2439	0.0283	92	5.8425	1	0.0026	3.5832	0.0019	0.1900
X	X	network 23 feeder 4	1	0.0026	48	12.5252	0.5	0.0026	0.4129	0.0000	0.0758
X	X	network 26 feeder 3	0.5878	0.0724	578	1.0681	1	0.0009	1.8173	0.0637	0.8696
X	X	network 28 feeder 1	0.5209	0.1151	528	1.2781	1	0.0027	2.7064	0.0913	0.5910
X	X	network 29 feeder 3	0.9996	0.0027	25	11.8530	0.35	0.0000	0.3720	0.0015	0.0836
X		network 29 feeder 10	1	0.0267	122	4.4312	1	0.0024	1.3767	0.0018	0.1367
X		network 29 feeder 11	0.4842	0.0270	64	7.2130	1	0.0021	2.1918	0.0022	0.1220
X		network 31 feeder 9	1	0.1303	114	1.2054	0.5	0.1303	4.0504	0.0007	0.1946
X		network 32 feeder 1	0.3566	0.0537	258	2.4318	1	0.0041	3.4304	0.0042	0.3292
X		network 32 feeder 7	0.5209	0.0931	775	1.5021	1	0.0016	2.3033	0.0348	0.5470
X		network 33 feeder 8	0.5861	0.1811	420	1.1049	1	0.0259	2.3808	0.0074	0.5893
X		network 33 feeder 9	0.6871	0.1694	1439	0.8480	0	0.0106	2.2763	0.0159	0.6906

The outliers in table 17 can be seperated in 5 groups. The first group consists of feeders, that have a single line without branchings and few loads at the end ($\varepsilon = 1$). However, their indicators Z_Σ for example are different. Feeders of this type are: network 01 feeder 2, network 14 feeder 2 (also outlier with non-electrical clustering), network 18 feeder 5, network 23 feeder 4, network 29 feeder 10 (1 single line that supplies 3 parts similar to this feeder group), network 29 feeder 11 (similar to the previous with sightly more loads), network 31 feeder 9 and network 33 feeder 8.

The second group of feeders contain transformers to supply isolated areas. Transformers are used to reduce transmission losses and the voltage drop over a distance. Near the distribution node the voltage is transformed to 980V and near the first customers back to 230V. The voltage drop on this type of feeders can be seen in figure 13 for network 01 feeder 8, which was also identified as outlier. Feeders of this type are also network 32 feeder 7 and network 33 feeder 9. After the transformer there are several branchings.

The next groups of feeders have many branchings and a few number of loads along the paths. Examples of this type are: network 03 feeder 5, network 04 feeder 6, network 09 feeder 5, network 11 feeder 8, network 18 feeder 7, network 21 feeder 4, network 32 feeder 1, network 23 feeder 1, network 20 feeder 3, network 28 and network 26 feeder 3. Some of the feeders start with a line with a certain length. From the end of the first line the power is distributed by branchings along several paths to the customers. These could even be seen as own outlier group.

The last group of outliers supply residential areas of high density. Examples of these outliers are: network 07 feeder 1, network 29 feeder 3 and network 12 feeder 7.

Table 18 shows the outliers by the highest percentiles 90, and 95 with non-electrical indicators.

Table 18: Feeder outliers by non-electrical indicators

outliers criterion		non-electrical outliers	N_{loads}	ML	N_{nodes}	$\overline{NON}$	$MAX(NON)$	$max(\overline{DTN})$	$MIN(DTN)$	c	LD
90%	95%	feeder	[1]	[1]	[1]	[1]	[1]	[m]	[m]	[m/1]	[1]
X		network 02 feeder 1	59	4	40	1.9821	4	242	16.5	41.28814	0.6780
X		network 04 feeder 1	1	2	1	1.5	2	212	106	212	1
X		network 04 feeder 8	1	1	1	1.6667	2	121	25	893	1
X	X	network 06 feeder 5	1	2	1	1.6667	2	304.35	25	1776.1	1
X	X	network 07 feeder 1(+EL)	68	42	7	1.9091	5	210	25	13.3088	0.1029
X	X	network 08 feeder 6	1	1	1	1.6667	2	256	25	999	1
X	X	network 09 feeder 5 (+EL)	64	6	43	1.9851	6	269	2	69.9687	0.6719
X	X	network 09 feeder 6	1	1	1	1.6667	2	315.5	25	1237	1
X		network 11 feeder 2	96	18	38	2	4	147.33	15	20.8438	0.3958
X	X	network 12 feeder 2	2	2	1	1.5	2	288	144	144	0.5
X	X	network 14 feeder 2 (+EL)	1	1	1	1.6667	2	371.5	25	743	1
X		network 15 feeder 4	59	4	37	2.0175	6	334.33	7	55.4576	0.6271
X	X	network 18 feeder 3	10	2	8	1.95	5	581	0	333.6	0.8
X		network 18 feeder 4	8	1	8	1.9474	6	400	16	216.625	1
X		network 22 feeder 2	26	26	1	1.6667	2	185	40	26.9231	0.0385
X		network 22 feeder 4	16	1	16	1.9524	12	49.5	13	37.0625	1
X	X	network 24 feeder 9	2	2	1	1.5	2	380	190	190	0.5
X		network 25 feeder 3	75	21	5	2.1	5	220	25	7.6	0.0667
X	X	network 26 feeder 3 (+EL)	85	7	44	1.9861	6	150	21	39.8706	0.517647
X		network 26 feeder 5	1	1	1	1.5	2	208	104	416	1
X		network 29 feeder 3 (+EL)	88	20	5	1.8333	2	12.51	0.01	0.5686	0.0568
X	X	network 29 feeder 12	87	45	5	1.875	5	70.5	25	3.747126	0.057471
X	X	network 30 feeder 7	204	11	25	2.0313	5	309	0	3.7011	0.122549
X		network 32 feeder 7 (+EL)	60	3	37	2.0189	8	314	0	49.5833	0.616667
X		network 34 feeder 2	2	2	1	1.5	2	264	132	132	0.5

Table 18 shows outliers by non-electrical clustering. Some of the feeders were already identified as outliers by electrical clustering. These feeders are marked with (+EL) in table. There are only 6 feeders that can be found by both indicator groups.

5.2 Analysis on network level

In this section the same approach as in the previous was done. The principal component analysis is followed by the study of the optimal cluster size for the 2 different cluster algorithms.

5.2.1 Data preparation - PCA

As described in the previous section, the principal components were used to plot feeders and networks as points. Again, there are as many components as original indicators, nevertheless, each could be a linear combination of all original indicators and the 3 most relevant will be used to mark the points. The usage of all PCA-components explains the total variance of the original indicator matrix. However, to visualize the clustering results only the first 3 most important components will be used to plot figures. In figure 33 the cumulated sum of explained variance can be seen. The first 3 indicators explain compared to feeder level less than 80% of the variance for both electrical and non-electrical measures, therefore on network level, there is a higher remaining variance than on feeder level. In the next two figures

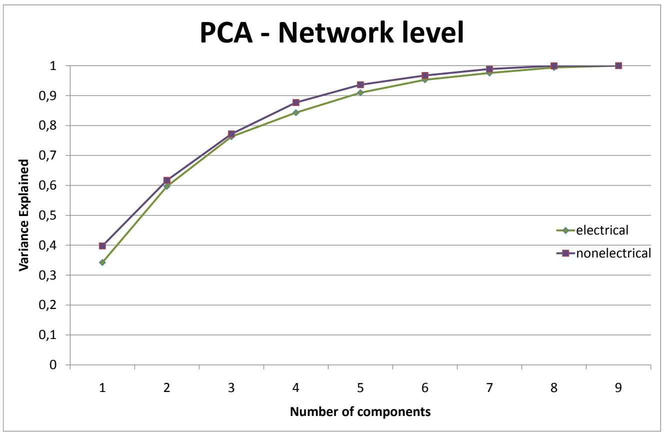


Figure 33: PCA - Network level

17 and figure 18 the distribution of the networks depending on the electrical and non-electrical indicator matrix can be seen. Compared to feeder level, there is not an area with high density of points. And the distribution in both figure are more diverse than compared to feeder level.

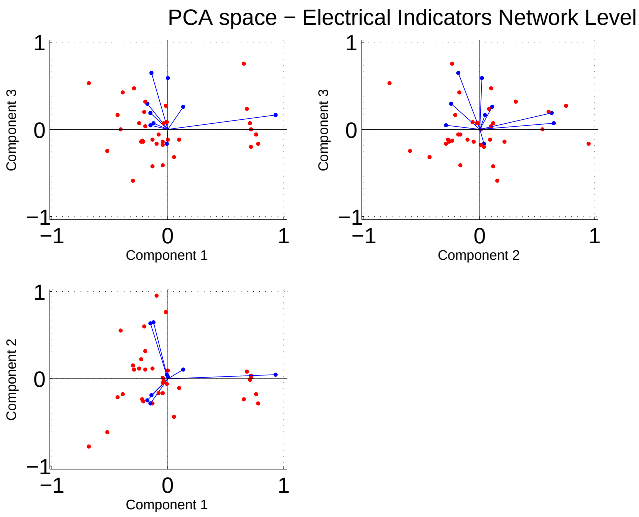


Figure 34: Principal components of electrical indicators

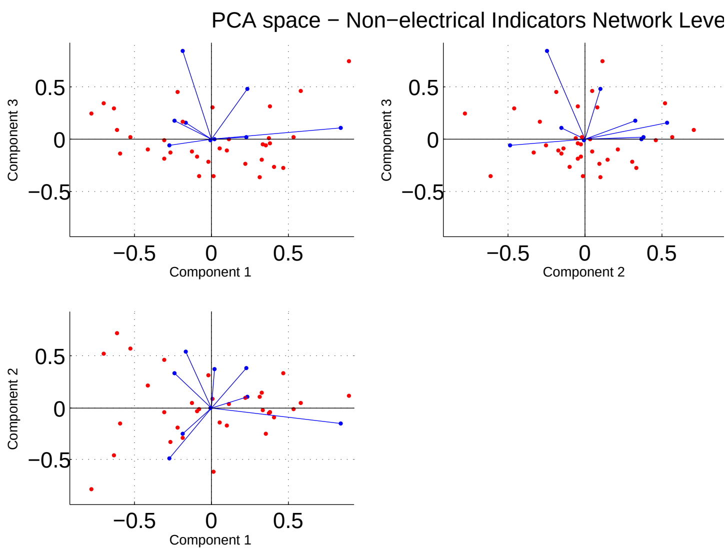


Figure 35: Principal components of non-electrical indicators

5.2.2 Clustering results network level

In this section the optimal clustering size was also analysed with ‘silhouette’, which also lead not to a clear cluster size propose. Therefore the same solution as on feeder level was used.

In figure 36 and 36 the threshold criterium can be seen. In these figures the effect of the variation of the threshold can be seen more easily than on feeder level. The number of clusters can easily be seen by counting the points above any defined threshold plus 1.

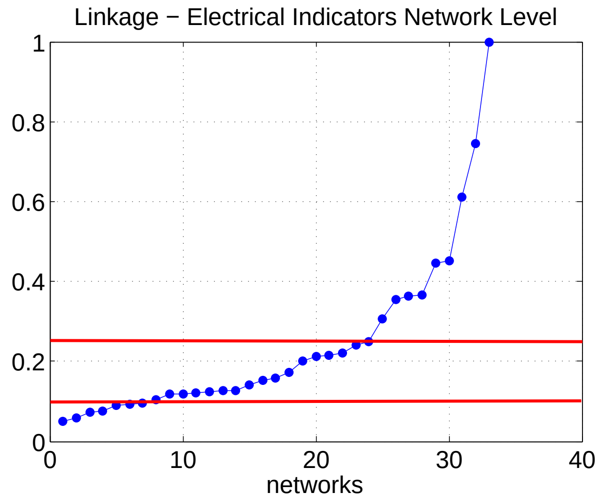


Figure 36: Linkage Thresholds - Electrical indicators

In figure 38 the distribution depending on electrical indicators of the networks can be seen. Again, the global median network was found by calculating the median of every measure. The network with the smallest distance to that point was then selected as global median. After that the circles indicating the 50%, 75%, 90% and 99% distances of the maximal distance to the center were drawn. The outliers can be identified analysing the points in the 2D-plots.

Table 20 shows the selected median of each cluster. The median was found by calculating the median of the indicators using only the networks in the same cluster. Clusters with only 1 network were selected directly as ‘median value. The points with same colour and symbol in figure 38 are in the same cluster. In total, 12 clusters were defined.

Finally, the clustering of networks with non-electrical measures was done. The points with same colour and symbol in figure 39 are in the same cluster. In total, 12 clusters were defined. Compared to the clustering with electrical

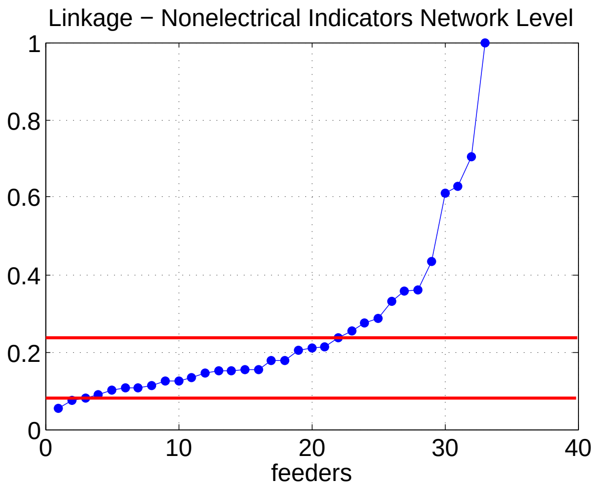


Figure 37: Linkage Thresholds - Non-electrical indicators

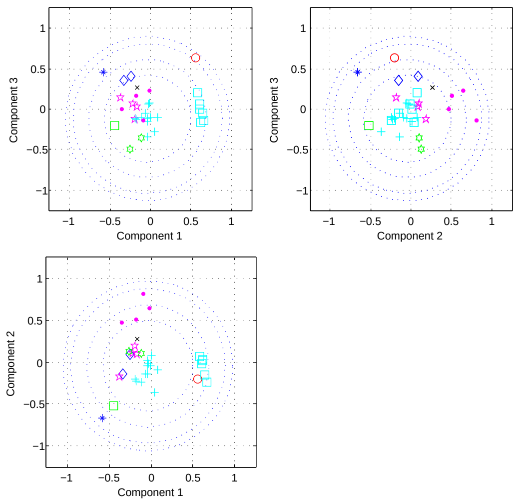


Figure 38: Clustered networks by electrical indicators

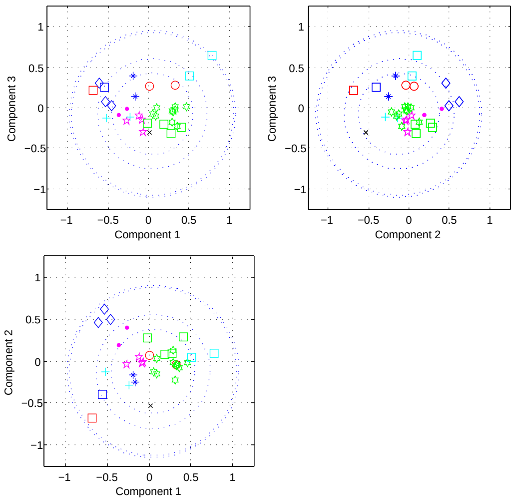


Figure 39: Clustered networks by non-electrical indicators

indicators, a completely different distribution of the points can be seen. The circles show again the distance from the global median (50%, 75%, 90% and 99%).

Table 19: Cluster centres by electrical indicators

	u_{maxmin}	Δu	$Median(\varepsilon)$	$MedianZ_{\Sigma}$	$Median(dvdQ)$	$Median(R/X)$	$max(d_{endnode})$	$Median(PR)$	$Median(Z_{\Sigma-rel})$
Network	[p.u.]	[p.u.]	[1]	[Ω]	[p.u./MVA]	[1]	[m]	[1]	[Ω/1]
network 01	0.0381	0.0381	0.6496	0.1008	0.2741	2.3103	2307	0.14	0.0050
network 02	0.0397	0.0398	0.5956	0.0843	0.4858	2.0076	905	0.25	0.0034
network 17	0.0286	0.0290	0.6551	0.0694	0.3826	2.0136	741	0.33	0.0048
network 18	0.0218	0.0228	0.6681	0.1407	0.4694	3.0781	1857	1	0.0130
network 20	0.0127	0.0165	0.5082	0.1695	1.5349	1.3809	1045	0	0.0174
network 25	0.0355	0.0359	0.8049	0.0198	0.2565	1.5714	438	0.065	0.0007
network 28	0.0226	0.0913	0.4566	0.1023	0.6618	2.3327	849	0.33	0.0024
network 08	0.0173	0.0175	0.8677	0.0809	0.4457	1.5213	623	0.04	0.0494
network 11	0.0671	0.0691	0.5810	0.0909	0.4574	2.2920	884	0.33	0.0041
network 31	0.0280	0.0287	0.5430	0.0621	0.3259	2.1393	796	1	0.0025

5.2.3 Outliers on network level

The network network 08 consists of 7 feeders and feeder 6 is an outlier with non-electrical clustering. There are 4 more feeders that are similar to feeder 6. Only feeder 2 and 3 supply many loads and have many branchings. Only $Median(Z_{\Sigma-rel})$ of this network is similar to the median network network 17. This can be seen in table 21 and table 22. The last network in 22 is network 20. It consists of only 3 feeders where one is already an electrical outlier. No

Table 20: Cluster centres by non-electrical indicators

	TR	LD	ML	NON	DTN	$max(DTN)$	$Min(DTN)$	TR/N_{loads}	$max(d_{endnode})$
Network	[kVA]	[1]	[1]	[1]	[m]	[m]	[m]	[kVA/1]	[m]
network 33	400	0.5906	8	1.9573	57.1822	491	0	2.6846	1656
network 03	400	0.6167	6	1.9737	33.4405	126	4	1.6667	796
network 22	400	0.3580	26	1.9500	46.1044	318	0	2.2727	919
network 21	630	0.5723	8	1.9858	46.5248	328	17	3.9623	838
network 24	800	0.4036	12	1.9274	44.1765	380	7.5	3.5874	675
network 06	400	0.5134	8	1.9888	44.7553	304.35	16.5	1.7857	775
network 07	250	0.4043	42	1.9717	35.6300	210	12.5	1.3298	540
network 34	160	0.6133	6	1.9753	48.2807	264	12	2.1333	585
network 09	400	0.6333	6	1.9535	55.8751	315.5	2	3.3333	1581
network 23	630	0.7963	3	1.9344	43.6461	149.5	15	11.6667	747
network 29	800	0.3181	45	1.9362	34.0155	242	0.01	1.9277	497
network 30	800	0.1632	11	1.9630	26.9269	309	0	1.4035	435

Table 21: Overall median electrical/non-electrical networks

	u_{maxmin}	Δu	$Median(\epsilon)$	$Median(Z_{\Sigma})$	$Median(dvdQ)$	$Median(R/X)$	$max(d_{endnode})$	$Median(PR)$	$Median(Z_{\Sigma-rel})$
Median Network - Electrical indicators	[p.u.]	[p.u.]	[1]	[Ω]	[p.u./MVA]	[1]	[m]	[1]	[Ω/1]
network 17	0.0286	0.0290	0.6551	0.0694	0.3826	2.0136	741	0.33	0.0048
	TR	LD	ML	NON	DTN	$max(DTN)$	$Min(DTN)$	TR/N_{loads}	$max(d_{endnode})$
Median Network- Non-electrical indicators	[kVA]	[1]	[1]	[1]	[m]	[m]	[m]	[kVA/1]	[m]
network 12	400	0.45	15	2	44	304	11	2	709

indicator value is similar to the global electrical median network. Network network 18 is an outlier by both indicator groups and will be discussed next.

The network network 18 is a 90% percentile outlier by both indicator groups. It has 7 feeders in total. Feeder 3 and 4 are also outliers of non-electrical clustering. And feeder 5 and 7 are outliers of electrical clustering. Feeder 3 contain similar transformers as already mentioned for network 01 feeder 8. Feeder 1 has many loads and is supplied by 2 parallel 4x150 cables. Probably the combination of feeders leads to the result that network 18 is an outlier network. network 27 consists of only two feeders that are completely different from each other. Feeder two consists of a single load and a 160kW DER which was not considered in the generic analysis. The second feeder consists of 2 main paths where 1-3 loads are supplied along the distance at many nodes. Compared to the indicator values of network network 12 in 21, no similarities can be seen. The last network in table 23 is network 29. This network consists of 12 feeders. 3 feeders are electrical outliers and 2 are non-electrical outliers. There are 5 feeders that have a simple topology with max. 3 short paths and a little number of loads. On the other side there are 2 feeders that supply residential buildings with a high number of loads. The other feeders have many branchings and equally distributed loads along

Table 22: Network outliers by electrical indicators

outlier criterion		electrical outliers	u_{maxmin}	Δu	$Median(\epsilon)$	$Median(Z_{\Sigma})$	$Median(dvdQ)$	$Median(R/X)$	$max(d_{endnode})$	$Median(PR)$	$Median(Z_{\Sigma-rel})$
90%	95%	network	[p.u.]	[p.u.]	[1]	[Ω]	[p.u./MVA]	[1]	[m]	[1]	[Ω/1]
X	X	network 08	0.0173	0.0175	0.8677	0.0809	0.4457	1.5213	623	0.04	0.0494
X		network 18	0.0218	0.0228	0.6681	0.1407	0.4694	3.0781	1857	1	0.0130
X	X	network 20	0.0127	0.0165	0.5082	0.1695	1.5349	1.3809	1045	0	0.0174

Table 23: Network outliers by non-electrical indicators

outlier criterion		non-electrical outliers	TR	LD	ML	NON	DTN	$max(DTN)$	$Min(DTN)$	TR/N_{loads}	$max(d_{endnode})$
90%	95%	network	[kVA]	[1]	[1]	[1]	[m]	[m]	[m]	[kVA/1]	[m]
X		network 18	400	0.6455	17	1.9848	76.7298	581	0	3.6364	1857
X	X	network 27	630	0.7419	3	1.9487	47.7179	160	20	20.3226	983
X	X	network 29	800	0.3181	45	1.9362	34.0155	242	0	1.9277	497

the feeders. Comparing the indicator values with the global median of non-electrical indicators show, that especially the measure ML is different. Also TR is 2 times higher than in network 12. The comparison of electrical and non-electrical outliers shows that the two measure groups lead to a different quantification of outliers. Only the network network 27 can be gathered as new information of the network analysis, as all other networks contain at least 1 outlier on feeder level. Therefore the analysis on feeder level could be of more interest.

5.3 Analysis of network 01 with PSS

Until the finalization of this thesis PSS data was only available for two feeders in the LV-network of network 01. Two Snap-Shots[1] for the feeders 1 and 8 were selected for an exemplary analyse.

In an optimal balanced LV-grid the current would uniformly flow between the three phases. However, most loads used at homes are single phase loads. The distribution on the 3 phases is in general not well known. The first analysis of the Snap-Shot data of network 01 leads to the observation that the loading on phase L1 is higher compared to the other two phases as the voltage drop on L1 is generally higher. In this section the question of to which extend the available voltage band is used due to unbalance will be investigated. For this real Snap-Shot data has to be used. In a first step a loadflow calculation is taken out and the line-neutral voltages at the loads are exported. In a second step the consumption of every load is distributed on the three phases equally. After that, again a loadflow calculation and the line-neutral voltages are also exported. The comparison of the two results returns the theoretically effect of ideally symmetrising loads on the usage of the allowed voltage band for every node. The effect of balancing was already investigated and was presented at the CIRED Workshop 2012 [1].

In figure 40 the voltage drop along the distance for each phase can be seen. The difference between the highest and lowest voltage in the LV-network is 3.6%. This FVR is between line (phase) A and B. After a symmetric distibution of the active and reactive power on all phases of a meter, the balancing gain can be seen. Figure 41 shows the voltage drop in an ideally symmetric case. The voltage drop from the transformer station to the end node is reduced significantly. The voltage band used would be reduced to 0.9%. This means that the balancing gain for the first Snap-Shot of feeder 1 is 2.7%. The second Snap-Shot for feeder 1 can be seen in the next figure 42. In this Snap-Shot, the voltage range is a slightly higher than in the first Snap-Shot. The voltage diagram of line A has nearly the same shape like in the previous Snap-Shot but has a little higher level. Line-neutral voltage B has also a similar shape like in figure 40 but has a slightly negative offset. Line-Neutral voltage C is mainly rising this time. In this situation the voltage range is higher than before and reaches 3.8%. Figure 44 shows a Snap-Shot of feeder 8. In this feeder the voltage is transformed at the distribution node to 980V and transmitted approximately 900 meters to a transformer which transforms the voltage down to 230V. The effects of unbalanced loading results in a used feeder voltage range of nearly 3.25%. In

[1] These two Snap-Shots were not validated yet due to some problems with the phase assignment

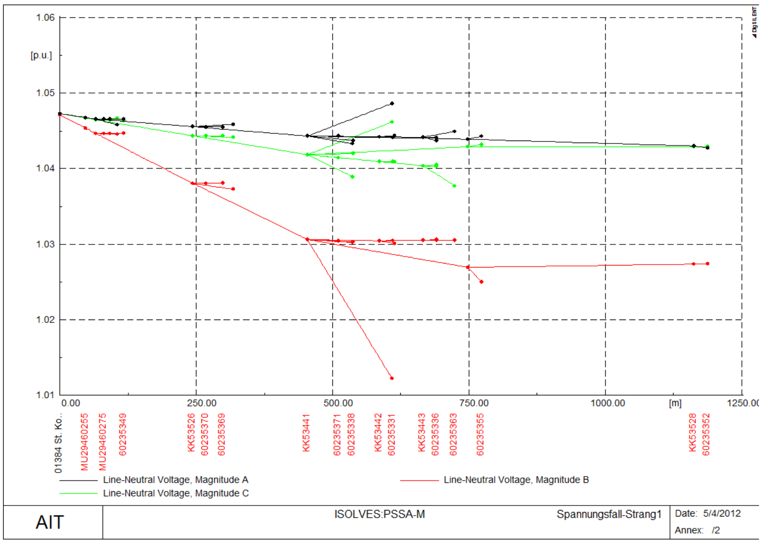


Figure 40: network 01 feeder 1 Snap-Shot 1

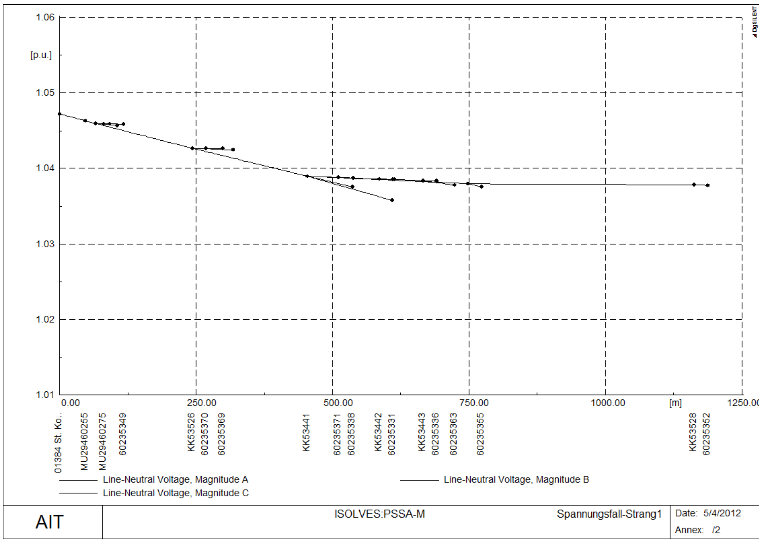


Figure 41: network 01 feeder 1 Snap-Shot 1 symmetrized

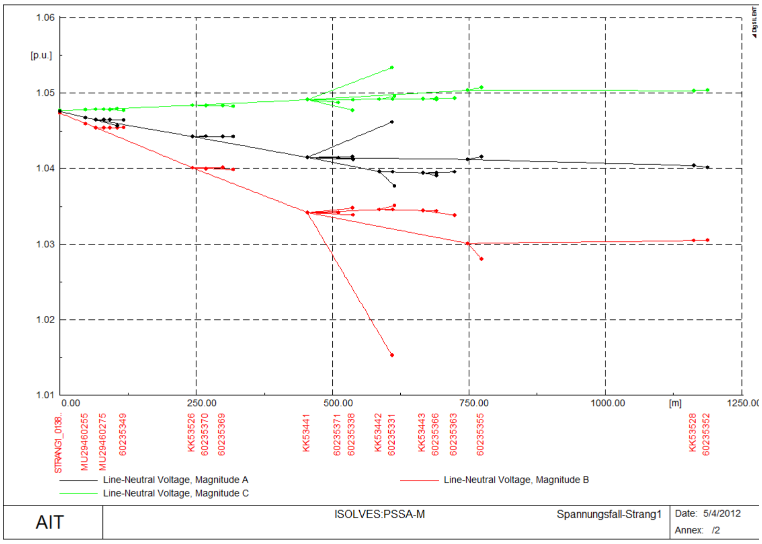


Figure 42: network 01 feeder 1 Snap-Shot 2

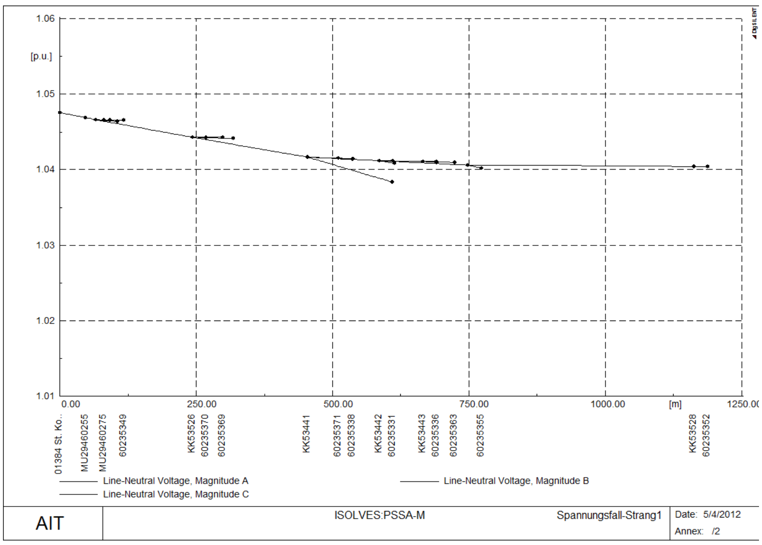


Figure 43: network 01 feeder 1 Snap-Shot 2 symmetrized

figure 45 the voltage drop of the ideally balanced case can be seen. As in the previous figure line B and C could be explained by line A and an offset, the change in the voltage drop in the symmetric case becomes more clear. The voltage range caused by the loading could be reduced to 1.5%. That means that the balancing gain is 1.75%.

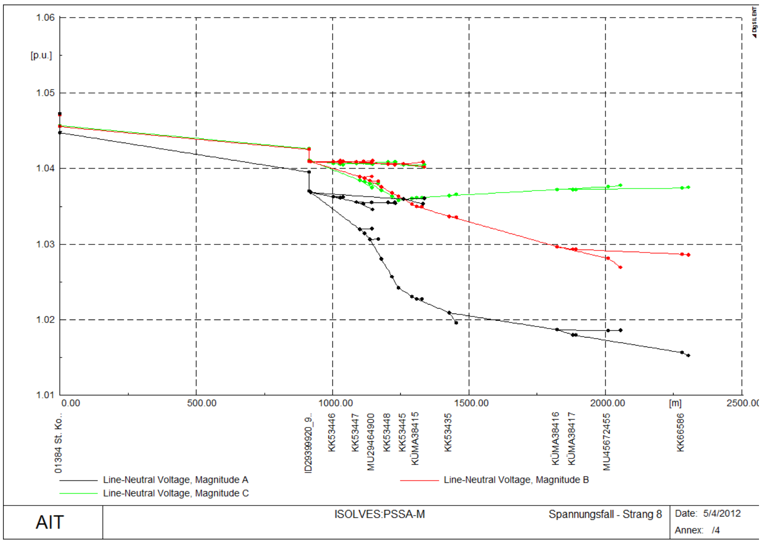


Figure 44: network 01 feeder 8 Snap-Shot 1

The results for the second Snap-Shot of feeder 8 show similar results. The voltage band used in figure 46 is nearly 2.2%. In figure 47 the balancing gain can be seen. The voltage range can be reduced to by 50% to 1.5% of the voltage.

The analysis of the symmetrized figures show, that the voltage can be graphically obtained by drawing the average of all 3 line-neutral voltages. Therefore the balancing gain could be calculated without network simulations of balanced PSS in PowerFactory models. Instead this factor can be directly calculated with the average of the measured voltages. In the generic analysis balanced loads were used to cluster feeders and networks. Also PSS data has to be balanced to calculate all indicators. Therefore BG could be used as coefficient to characterize the effects of balancing in the analysis. Low BG values would mean that the balancing has little effects on the used PSS.

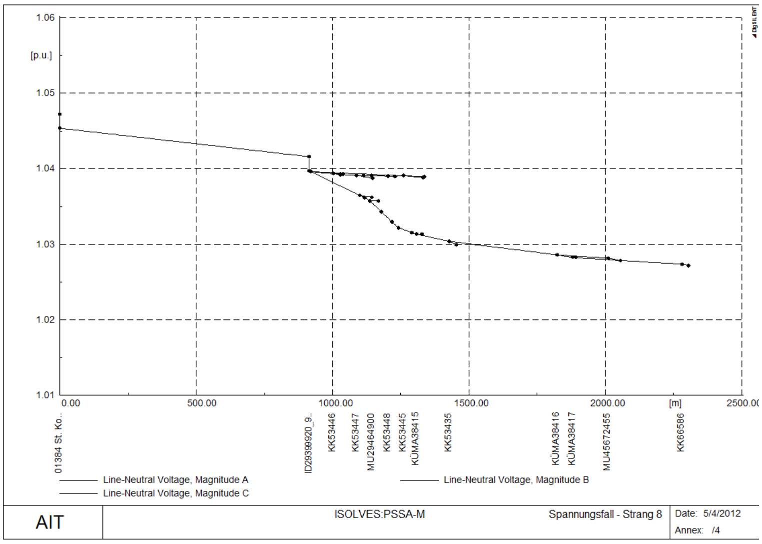


Figure 45: network 01 feeder 8 Snap-Shot 1 symmetrized

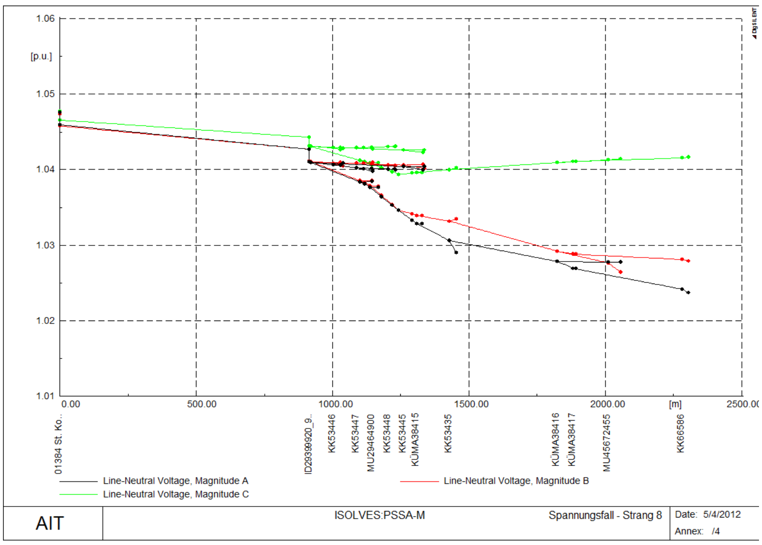


Figure 46: network 01 feeder 8 Snap-Shot 2

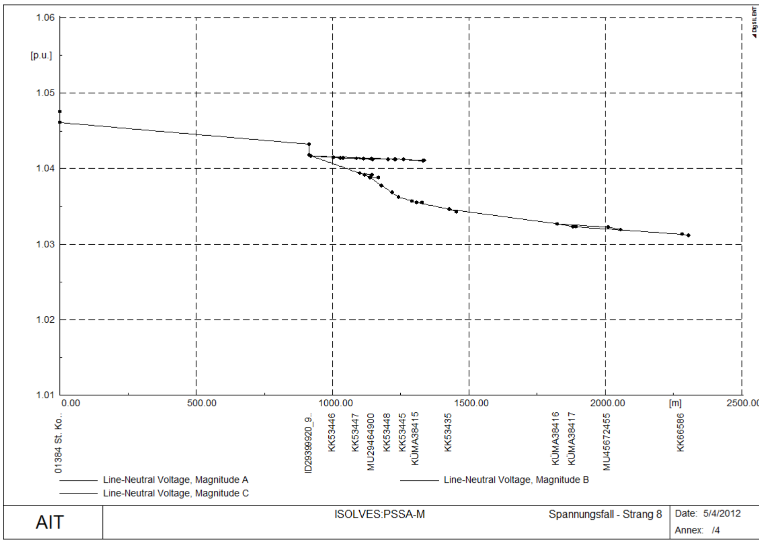


Figure 47: network 01 feeder 8 Snap-Shot 2 symmetrized

6 Outlook

The results presented in this work were obtained by considering a constant and equal consumption for all loads. In the frame of the project ISOLVES:PSSA-M, Power Snap-shots will be used once they are fully available and once the network models are validated for this purpose. The validation will consist in comparing the voltages obtained from the simulation fed with the load data from the Snap-shots to the measured voltages. With more available Snap-shots of the 34 grids in the future it is estimated that all interesting load situations (e.g. strong voltage rise or drop) can be captured and the reasons analysed. Additionally electrical indicators can be calculated again for feeders and networks and compared with the results from the generic analysis. Cases, where a feeder or network would be classified in another cluster than the same feeder or network with generic data would be of high interest. Additionally, any in PowerFactory modelled grid can be analysed with the tools programmed in this work without any adaptation. As new loads could be added to feeders or meters could be changed in the future, an external versioning system could ensure the usability of past Snap-shot data with the latest network models. The analysis of the Snap-shots is expected to help improving network planning for smart grids. In particular, future scenarios could be investigated with the Snap-shot data by adapting it according to the expected growth of renewable energy. The effects on the voltage band could for example be simulated and the installable renewable generation capacity could be determined. On the basis of such analyses, possible solutions to solve potential problems can be tested in the simulation and their effectiveness can be quantified. This would deliver valuable information to decide on investments which will be necessary to prepare the grids for the challenges of the future. The power Snap-shots could be used to investigate and quantify the effect of unbalance on the voltage band. With the help of a Pareto algorithm, a set of reconfiguration measures (phase assignment) could help reducing the part of the voltage band needed by unbalance.

7 Conclusion

In a first step, a rather large set of indicators that can be used to characterise and classify LV feeders (and networks) has been proposed and evaluated. These indicators help understand the difference and commonalities between LV feeders. The work has been focused on electrical indicators (calculated by network computations) since they provide a priori more information than non-electrical (not needing network computations, such as transformer rating, number of nodes or total cable length). The computation of the indicators proposed in this work has been automatized so that they can easily be evaluated for any network. Besides the high-level classification done in this work, some specific issues can be investigated by analysing particular indicators. For example, the R/X ratio which has been used in the overall classification is of interest when investigating the effectiveness of a reactive power-based voltage control. On the basis of a linear regression, the number of indicators has been reduced to nine electrical indicators due to the partial redundancy between some of them. In a second step, the LV feeders or networks have been clustered by a hierarchical clustering algorithm. A criterion for selecting the number of clusters has been proposed. It consists of considering that elements whose distance is smaller than 25 % of the maximal distance between any two elements can be grouped into a same cluster. The comparison of the results obtained by the clustering on the basis of the electrical and non-electrical indicators shows that the results are different and suggests that the information content of electrical indicators cannot be fully replaced by (a large number of) non-electrical indicators. The proposed approach led to a final number of clusters of 9. For each cluster, the 'center', defined as the median feeder has been identified; it can be seen as representative for the cluster. In a dedicated chapter, these cluster centres have been described and the clustering results discussed. While some of the clustering results are rather straightforward to explain, some peculiarities are more complex and result from the multidimensional character of the problem. Indeed, the difference between elements close to the border between clusters might be sometimes difficult to explain. One should keep in mind that clustering techniques are statistical tools allowing a structured description of datasets according to some characteristics. The significance of the results in terms of electrical properties highly depends on the used metrics, methods and criteria. The used methods and considerations allowed however having a better understanding on the network properties. The types of clusters which could be best interpreted were for example very short and generously sized or rather long and poorly sized feeders supplying few loads concentrated at the end, feeder with an average load density and numerous branchings resulting in a larger stress in terms of voltage band and short feeders with a high load density. Apart from this cluster analysis, outliers have been identified by defining a median feeder and looking at the similarity (measured by the euclidean distance) between feeders and this median feeder. The outliers consisting of the 5 % less common feeders do not only correspond to long and or stressed feeders but also to very short ones. The results must be carefully interpreted since they are based on several assumptions. For example, a constant active and reactive power value has been used for all loads, which is of course not the case in the reality. Nevertheless, the analysis allows characterising the networks or feeders with a focus on the basis of the topology. As previously mentioned, further analyses will be performed in the frame of the project ISOLVES:PSSA-M with the Power Snap-Shots once they are fully available.

References

[1] A. Abart, B. Bletterie, S. Kadam, et. al. Characterising lv networks on the basis of smart meter data and accurate network models (cired 2012 workshop).

[2] A. Abart M. Heidl B. Bletterie, A. Gorsek. Understanding the effects of unsymmetrical infeed on the voltage rise for the design of suitable voltage control algorithms with pv inverters.

[3] DIgSilent. Digsilent powerfactory user's manual, 2011.

[4] Energie-control Austria. 10 jahre energiemarkt-liberalisierung, June 2011.

[5] http://www.bmwfj.gv.at. 138.verordnung intelligente messgerte-einfhrungsverordnung. `http://www.bmwfj.gv.at/Ministerium/Rechtsvorschriften/kundgemachte_rechtsvorschriften/Documents/IntelligenteMessgerte.pdf`, visited April 2012.

[6] http://www.neue energieanbieter.de. `http://www.neue-energieanbieter.de/data/uploads/06_11_30_anlage_1_vdn_ar.pdf`, visited April 2012.

[7] http://www.ris.bka.gv.at. Elektrizittswirtschafts- und -organisationsgesetz elwog 1998. `http://www.ris.bka.gv.at/Dokumente/BgblPdf/1998_143_1/1998_143_1.pdf`.

[8] J. Scheffler. Bestimmung der maximal zulssigen Netzanschlussleistung photovoltaischer Energiewandlungsanlagen in Wohnsiedlungsgebieten, 2002.

[9] Heuck K. Elektrische energieversorgung, 2002.

[10] M. Bodach. Energiespeicher im Niederspannungsnetz zur Integration dezentraler, fluktuierender Energiequellen, 2006.

[11] Mathworks. Matlab help.

[12] R. Witzmann, G. Kerber. Aufnahmefähigkeit der Verteilnetzte fr Strom aus Photovoltaik. *ew Fachthema Erneuerbare Energien*, 4:50–54, 2007.

[13] R. Witzmann, G. Kerber. Statistische Analyse von NS-Verteilungsnetzen und Modellierung von Referenznetzen. *ew Fachthema Netze*, 6:22–26, 2008.

[14] Sascha Szott. Einführung in das data mining, 2008.

List of Tables

List of Figures

A Clustering Results

In this section the results of the hierarchical clustering will be presented. The distance between the furthest points (Z_{max}) is the euclidean distance calculated with 'pdist' function of Matlab. A threshold of $0.75 \cdot Z_{max}$ was used to cluster the data. The following pages show the clustering results on feeder and network level with electrical and non-electrical indicators. In each cluster the hypothetical median of all objects was found by calculating the median for each measure of all objects of the cluster. After that the nearest feeder or network to the hypothetical median was defined as most describing element of the cluster. These elements were marked bold in the clusters.

Cluster 1 Network 01 feeder 1 Network 01 feeder 3 Network 01 feeder 6 Network 01 feeder 9 Network 06 feeder 1 Network 10 feeder 3 Network 10 feeder 5 Network 10 feeder 6 Network 11 feeder 5 Network 11 feeder 7 Network 12 feeder 3 Network 12 feeder 6 Network 12 feeder 7 Network 13 feeder 1 Network 15 feeder 2 Network 16 feeder 1 Network 16 feeder 6 Network 16 feeder 7 Network 17 feeder 6 Network 18 feeder 1 Network 21 feeder 4 Network 21 feeder 7 Network 22 feeder 4 Network 22 feeder 7 Network 23 feeder 1 Network 23 feeder 2 Network 24 feeder 2 Network 24 feeder 10 Network 24 feeder 11 Network 29 feeder 2 Network 29 feeder 10 Network 29 feeder 11 Network 29 feeder 12 Network 30 feeder 11 Network 31 feeder 2 **Network 31 feeder 4** Network 32 feeder 1 Network 32 feeder 2 Network 32 feeder 5 Network 33 feeder 1 Network 34 feeder 1 Network 34 feeder 6	Cluster 3 **Network 02 feeder 2** Network 03 feeder 2 Network 03 feeder 4 Network 04 feeder 7 Network 05 feeder 1 Network 05 feeder 2 Network 05 feeder 4 Network 06 feeder 7 Network 07 feeder 3 Network 11 feeder 1 Network 11 feeder 3 Network 11 feeder 4 Network 12 feeder 8 Network 12 feeder 9 Network 14 feeder 1 Network 14 feeder 3 Network 15 feeder 3 Network 17 feeder 5 Network 17 feeder 7 Network 21 feeder 1 Network 21 feeder 5 Network 22 feeder 5 Network 22 feeder 6 Network 24 feeder 1 Network 25 feeder 1 Network 26 feeder 2 Network 29 feeder 6 Network 30 feeder 5 Network 31 feeder 1 Network 31 feeder 5 Network 31 feeder 7 Network 31 feeder 8 Network 33 feeder 4 Network 34 feeder 3 Network 34 feeder 4 Network 34 feeder 5

Cluster 4	Cluster 5
Network 01 feeder 4	Network 02 feeder 3
Network 02 feeder 1	Network 03 feeder 3
Network 02 feeder 4	Network 03 feeder 6
Network 02 feeder 5	Network 03 feeder 7
Network 02 feeder 6	Network 04 feeder 3
Network 03 feeder 1	Network 04 feeder 5
Network 04 feeder 2	Network 05 feeder 3
Network 06 feeder 4	Network 05 feeder 5
Network 07 feeder 2	Network 05 feeder 6
Network 07 feeder 4	Network 05 feeder 8
Network 08 feeder 2	**Network 05 feeder 9**
Network 08 feeder 3	Network 06 feeder 3
Network 09 feeder 2	Network 09 feeder 1
Network 11 feeder 2	Network 13 feeder 5
Network 12 feeder 4	Network 13 feeder 6
Network 13 feeder 2	Network 13 feeder 8
Network 13 feeder 3	Network 14 feeder 4
Network 15 feeder 4	Network 14 feeder 5
Network 17 feeder 1	Network 14 feeder 6
Network 19 feeder 1	Network 17 feeder 3
Network 20 feeder 2	Network 17 feeder 8
Network 24 feeder 6	Network 21 feeder 3
Network 26 feeder 4	Network 22 feeder 11
Network 27 feeder 1	Network 24 feeder 3
Network 28 feeder 2	Network 24 feeder 4
Network 29 feeder 5	Network 24 feeder 5
Network 29 feeder 7	Network 25 feeder 3
Network 30 feeder 7	Network 30 feeder 3
Network 32 feeder 4	Network 30 feeder 8
Network 33 feeder 2	Network 31 feeder 3
Network 33 feeder 3	
Network 33 feeder 6	

<table>
<tr>
<td>Cluster 2
Network 01 feeder 2
Network 01 feeder 5
Network 01 feeder 7
Network 02 feeder 7
Network 04 feeder 1
Network 07 feeder 5
Network 08 feeder 1
Network 09 feeder 3
Network 09 feeder 4
Network 12 feeder 2
Network 12 feeder 5
Network 13 feeder 7
Network 21 feeder 6
Network 22 feeder 9
Network 23 feeder 4
Network 24 feeder 9
Network 25 feeder 2
Network 26 feeder 5
Network 27 feeder 2
Network 29 feeder 1
Network 29 feeder 3
Network 30 feeder 9
Network 30 feeder 10
Network 34 feeder 2</td>
<td rowspan="2">Cluster 7
Network 04 feeder 4
Network 04 feeder 8
Network 07 feeder 6
Network 08 feeder 5
Network 08 feeder 7
Network 15 feeder 6
Network 16 feeder 2
Network 16 feeder 3
Network 16 feeder 4
Network 17 feeder 2
Network 17 feeder 4
Network 21 feeder 8
Network 21 feeder 9
Network 22 feeder 2
Network 22 feeder 3
Network 22 feeder 10
Network 22 feeder 12
Network 24 feeder 7
Network 24 feeder 8
Network 25 feeder 4
Network 29 feeder 8
Network 29 feeder 9
Network 30 feeder 1
Network 30 feeder 2
Network 32 feeder 8</td>
</tr>
<tr>
<td>Cluster 6
Network 03 feeder 5
Network 04 feeder 6
Network 07 feeder 1
Network 09 feeder 5
Network 10 feeder 1
Network 11 feeder 8
Network 18 feeder 2
Network 18 feeder 7
Network 26 feeder 3
Network 28 feeder 1
Network 31 feeder 6
Network 32 feeder 7
Network 33 feeder 8</td>
</tr>
</table>

Cluster 8	Cluster 9
Network 01 feeder 8	Network 06 feeder 5
Network 05 feeder 7	Network 08 feeder 4
Network 06 feeder 2	Network 08 feeder 6
Network 06 feeder 6	Network 09 feeder 6
Network 10 feeder 2	Network 12 feeder 10
Network 10 feeder 4	Network 13 feeder 4
Network 11 feeder 6	Network 14 feeder 2
Network 11 feeder 9	Network 14 feeder 7
Network 12 feeder 1	**Network 15 feeder 5**
Network 14 feeder 8	Network 16 feeder 5
Network 15 feeder 1	Network 18 feeder 5
Network 18 feeder 3	Network 19 feeder 2
Network 18 feeder 4	Network 19 feeder 4
Network 18 feeder 6	Network 22 feeder 1
Network 19 feeder 3	Network 22 feeder 8
Network 19 feeder 5	Network 26 feeder 1
Network 20 feeder 1	Network 29 feeder 4
Network 20 feeder 3	Network 30 feeder 4
Network 21 feeder 2	Network 31 feeder 9
Network 23 feeder 3	Network 32 feeder 3
Network 30 feeder 6	Network 33 feeder 7
Network 32 feeder 6	
Network 33 feeder 5	
Network 33 feeder 9	

Non-electrical Indicators part 1 of 3

Cluster 1

Network 01 feeder 1
Network 01 feeder 8
Network 04 feeder 6
Network 13 feeder 1
Network 15 feeder 1
Network 18 feeder 2
Network 18 feeder 3
Network 18 feeder 4
Network 18 feeder 7
Network 21 feeder 2
Network 32 feeder 5
Network 33 feeder 9

Cluster 2

Network 01 feeder 2
Network 01 feeder 5
Network 01 feeder 7
Network 04 feeder 4
Network 05 feeder 1
Network 09 feeder 3
Network 09 feeder 4
Network 12 feeder 5
Network 14 feeder 3
Network 14 feeder 6
Network 17 feeder 8
Network 19 feeder 4
Network 21 feeder 6
Network 21 feeder 9
Network 24 feeder 3
Network 24 feeder 8
Network 25 feeder 2
Network 29 feeder 1
Network 29 feeder 9
Network 29 feeder 10
Network 29 feeder 11
Network 30 feeder 1
Network 30 feeder 2
Network 30 feeder 9
Network 31 feeder 5

Cluster 4

Network 01 feeder 4
Network 01 feeder 9
Network 02 feeder 2
Network 02 feeder 3
Network 02 feeder 5
Network 02 feeder 6
Network 03 feeder 6
Network 04 feeder 2
Network 05 feeder 2
Network 05 feeder 7
Network 05 feeder 8
Network 06 feeder 2
Network 06 feeder 4
Network 06 feeder 7
Network 10 feeder 2
Network 11 feeder 4
Network 11 feeder 5
Network 11 feeder 6
Network 13 feeder 2
Network 14 feeder 1
Network 14 feeder 4
Network 14 feeder 8
Network 15 feeder 2
Network 15 feeder 3
Network 16 feeder 1
Network 17 feeder 5
Network 17 feeder 6
Network 19 feeder 2
Network 19 feeder 3
Network 19 feeder 5
Network 20 feeder 1
Network 20 feeder 3
Network 21 feeder 1
Network 21 feeder 3
Network 22 feeder 8
Network 24 feeder 1
Network 25 feeder 1
Network 29 feeder 6
Network 30 feeder 5
Network 30 feeder 11
Network 31 feeder 7
Network 31 feeder 8
Network 33 feeder 3,4,5,6,7,8
Network 34 feeder 3,5,6

<table>
<tr>
<td>Cluster 3
Network 01 feeder 3
Network 01 feeder 6
Network 05 feeder 6
Network 06 feeder 6
Network 10 feeder 6
Network 11 feeder 3
Network 12 feeder 4
Network 12 feeder 7
Network 12 feeder 8
Network 12 feeder 9
Network 13 feeder 3
Network 14 feeder 5
Network 17 feeder 4
Network 18 feeder 1
Network 22 feeder 5
Network 24 feeder 2
Network 24 feeder 10
Network 24 feeder 11
Network 25 feeder 3
Network 26 feeder 2
Network 29 feeder 3
Network 30 feeder 3
Network 30 feeder 6
Network 30 feeder 8
Network 30 feeder 10
Network 31 feeder 3</td>
<td rowspan="2">Cluster 5
Network 02 feeder 7
Network 04 feeder 1
Network 07 feeder 5
Network 07 feeder 6
Network 08 feeder 1
Network 08 feeder 4
Network 08 feeder 5
Network 10 feeder 5
Network 11 feeder 7
Network 12 feeder 10
Network 13 feeder 4
Network 13 feeder 6
Network 13 feeder 7
Network 13 feeder 8
Network 14 feeder 7
Network 15 feeder 5
Network 15 feeder 6
Network 16 feeder 5
Network 17 feeder 2
Network 17 feeder 7
Network 18 feeder 5
Network 18 feeder 6
Network 21 feeder 8
Network 22 feeder 1
Network 22 feeder 10
Network 22 feeder 12
Network 23 feeder 4
Network 24 feeder 7
Network 25 feeder 4
Network 26 feeder 5
Network 27 feeder 2
Network 29 feeder 4
Network 29 feeder 8
Network 30 feeder 4
Network 31 feeder 9</td>
</tr>
<tr>
<td>Cluster 8
Network 04 feeder 8
Network 06 feeder 5
Network 08 feeder 6
Network 09 feeder 6
Network 14 feeder 2</td>
</tr>
</table>

Non-electrical Indicators part 3 of 3

<table>
<tr>
<td rowspan="4">Cluster 6
Network 03 feeder 1
Network 03 feeder 2
Network 03 feeder 3
Network 03 feeder 4
Network 03 feeder 5
Network 03 feeder 7
Network 04 feeder 5
Network 05 feeder 5
Network 06 feeder 3
Network 07 feeder 2
Network 07 feeder 3
Network 07 feeder 4
Network 08 feeder 2
Network 08 feeder 3
Network 09 feeder 2
Network 10 feeder 1
Network 10 feeder 3
Network 11 feeder 8
Network 11 feeder 9
Network 12 feeder 1
Network 12 feeder 3
Network 12 feeder 6
Network 16 feeder 6
Network 16 feeder 7
Network 17 feeder 1
Network 19 feeder 1
Network 21 feeder 4
Network 21 feeder 5
Network 23 feeder 3
Network 24 feeder 4
Network 24 feeder 6
Network 27 feeder 1
Network 28 feeder 1
Network 28 feeder 2
Network 29 feeder 2
Network 29 feeder 5
Network 29 feeder 7
Network 31 feeder 2
Network 31 feeder 4
Network 31 feeder 6
Network 32 feeder 2
Network 32 feeder 4
Network 32 feeder 6
Network 33 feeder 1</td>
<td>Cluster 7
Network 04 feeder 7
Network 05 feeder 3
Network 05 feeder 4
Network 06 feeder 1
Network 08 feeder 7
Network 09 feeder 1
Network 10 feeder 4
Network 11 feeder 1
Network 13 feeder 5
Network 17 feeder 3
Network 20 feeder 2
Network 21 feeder 7
Network 22 feeder 4,6,7
Network 23 feeder 1
Network 23 feeder 2
Network 24 feeder 5
Network 26 feeder 1
Network 31 feeder 1
Network 32 feeder 1
Network 32 feeder 3
Network 33 feeder 2
Network 34 feeder 1,4</td>
</tr>
<tr>
<td>Cluster 9
Network 05 feeder 9
Network 07 feeder 1
Network 16 feeder 2,3,4
Network 22 feeder 2,3,9,11
Network 29 feeder 12
Network 32 feeder 8</td>
</tr>
<tr>
<td>Cluster 10
Network 02 feeder 1
Network 02 feeder 4
Network 04 feeder 3
Network 09 feeder 5
Network 11 feeder 2
Network 15 feeder 4
Network 26 feeder 3,4
Network 30 feeder 7
Network 32 feeder 7</td>
</tr>
<tr>
<td>Cluster 11
Network 12 feeder 2
Network 24 feeder 9
Network 34 feeder 2</td>
</tr>
</table>

Electrical Indicators Network Level

Cluster 1 **Network 01** Network 33	Cluster 2 **Network 02** Network 07 Network 15 Network 19
Cluster 3 Network 03 Network 04 Network 05 Network 06 Network 12 Network 13 Network 14 **Network 17** Network 21 Network 24 Network 27 Network 34	Cluster 4 **Network 18** Cluster 5 **Network 20** Cluster 6 **Network 25** Network 29 Cluster 7 **Network 28** Cluster 8 **Network 08** Cluster 9 Network 09 **Network 11** Network 26 Network 30
Cluster 10 Network 10 Network 16 Network 22 Network 23 **Network 31** Network 32	

Non-electrical Indicators Network Level

Cluster 1 Network 01 Network 18 **Network 33**	Cluster 2 Network 02 **Network 03** Network 10 Network 31
Cluster 3 **Network 22** Network 25	Cluster 4 Network 04 **Network 21**
Cluster 5 Network 05 **Network 24**	Cluster 6 **Network 06** Network 11 Network 12 Network 13 Network 14 Network 16 Network 17 Network 19 Network 26 Network 28
Cluster 7 Network 07	Cluster 8 Network 08 Network 15 Network 20 **Network 34**
Cluster 9 **Network 09** Network 32	Cluster 10 **Network 23** Network 27
Cluster 11 **Network 29**	Cluster 12 Network 30

B Cluster centers and outliers

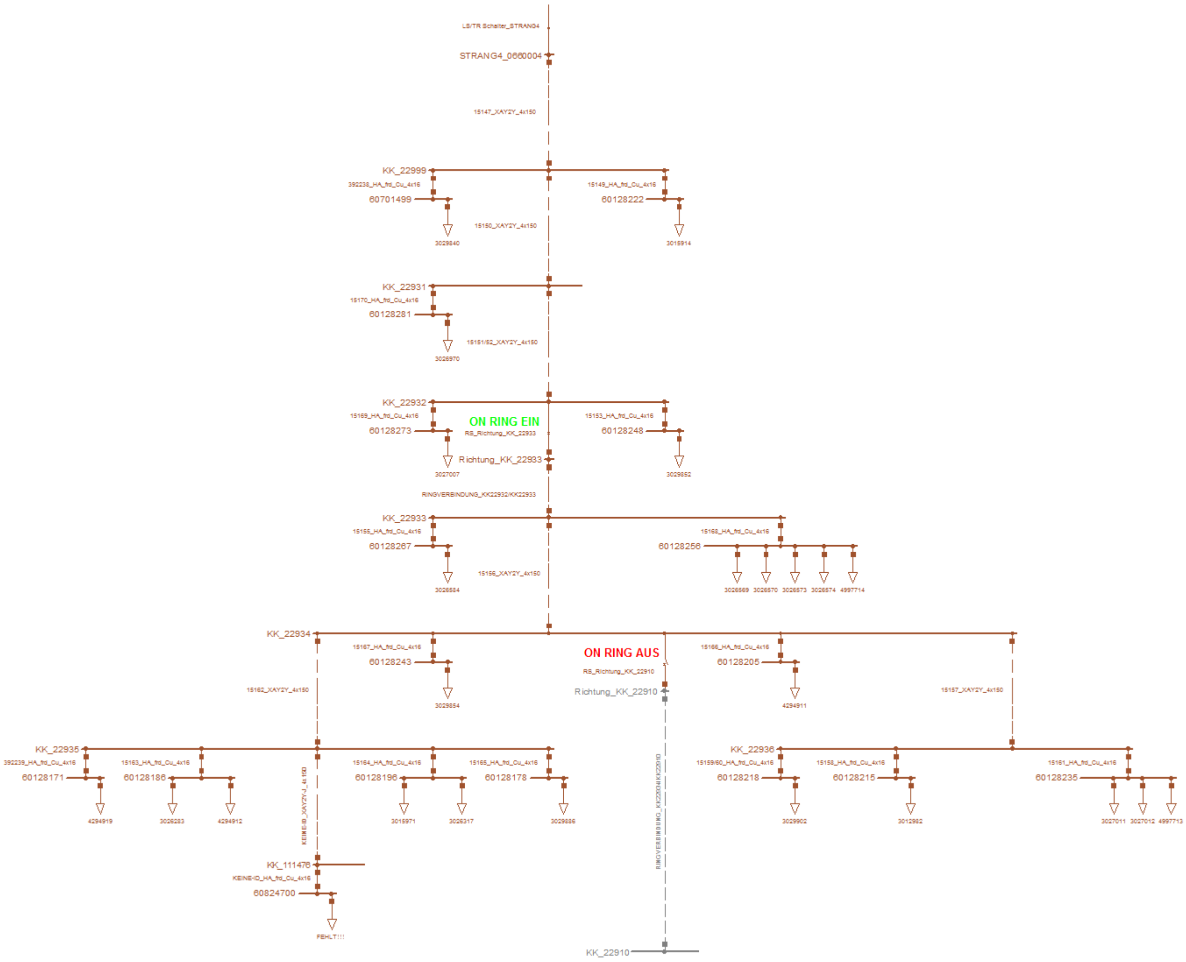
In this section the cluster centers on feeder level can be found together with the 90% outliers from the electrical and non-electrical median feeder and network additionally to the indicator values already discussed int chapter 5. Due to the definition of the number of feeder, zone and 'Strang' in PowerFactory, the assignment between feeder and 'Strang' is not consistent. Table 24 shows the assignment for the feeders in the appendix. Onyl differing values are written in column 'Strang'.

Table 24: Strang - Feeder assignment

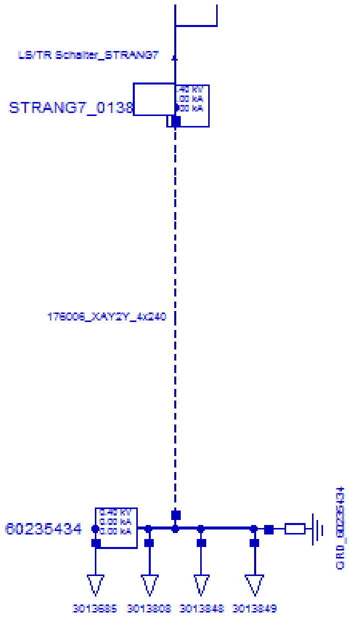
Strang	Feeder	Strang	Feeder
	network 01 feeder 2	4	network 20 feeder 3
	network 01 feeder 8	6	network 21 feeder 4
	network 02 feeder 1	12	network 22 feeder 2
	network 03 feeder 5	11	network 22 feeder 4
	network 04 feeder 1		network 23 feeder 1
	network 04 feeder 6	6	network 23 feeder 4
11	network 04 feeder 8	8	network 24 feeder 9
6	network 06 feeder 5		network 25 feeder 3
	network 07 feeder 1		network 26 feeder 3
	network 08 feeder 6		network 26 feeder 5
	network 09 feeder 5	7	network 28 feeder 1
	network 09 feeder 6	12	network 29 feeder 3
4	network 11 feeder 2	7	network 29 feeder 10
13	network 11 feeder 8	9	network 29 feeder 11
13	network 12 feeder 2	10	network 29 feeder 12
8	network 12 feeder 7	5	network 30 feeder 7
	network 14 feeder 2	11	network 31 feeder 9
6	network 15 feeder 4		network 32 feeder 1
	network 18 feeder 3		network 32 feeder 7
5	network 18 feeder 4	7	network 33 feeder 8
6	network 18 feeder 5	10	network 33 feeder 9
8	network 18 feeder 7	4	network 34 feeder 2
2	network 15 feeder 1	7	network 15 feeder 5
3	network 17 feeder 2	12	network 30 feeder 2

Cluster Centres

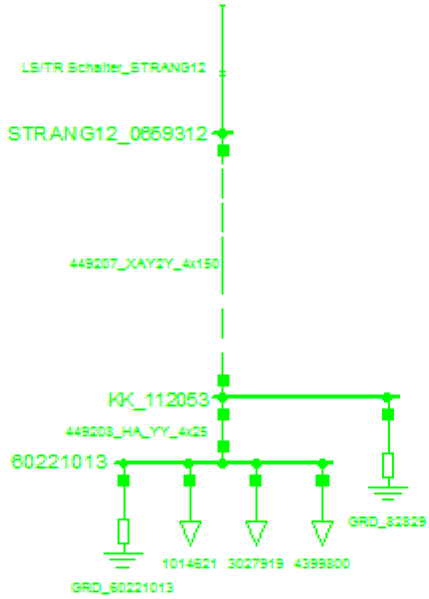
Electrical cluster 1 - Network 31 feeder 4



Electrical cluster 2 – Network 01 feeder 7

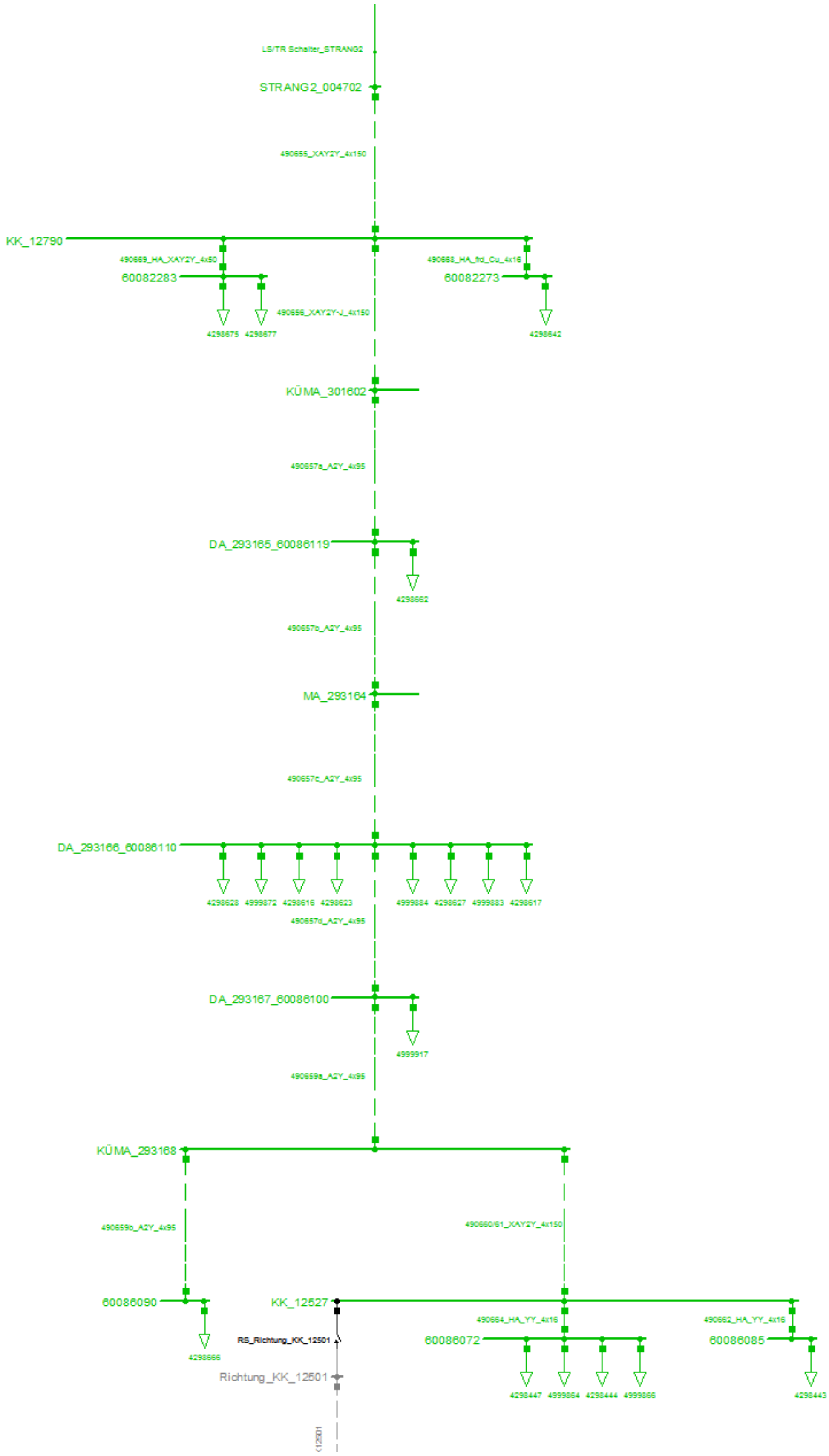


Electrical cluster 7 - Network 30 feeder 2



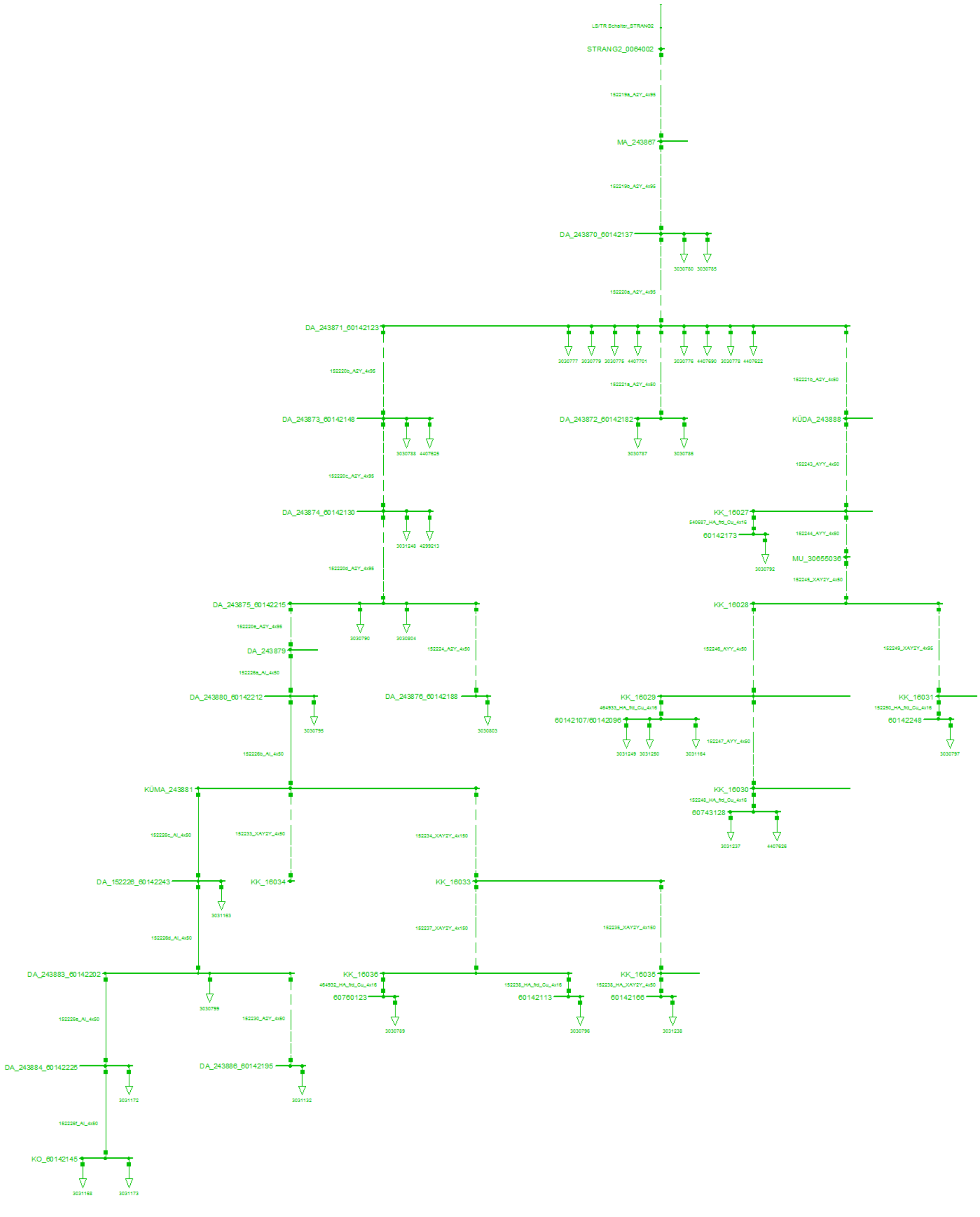
Cluster Centres

Electrical cluster 3 - Network 02 feeder 2



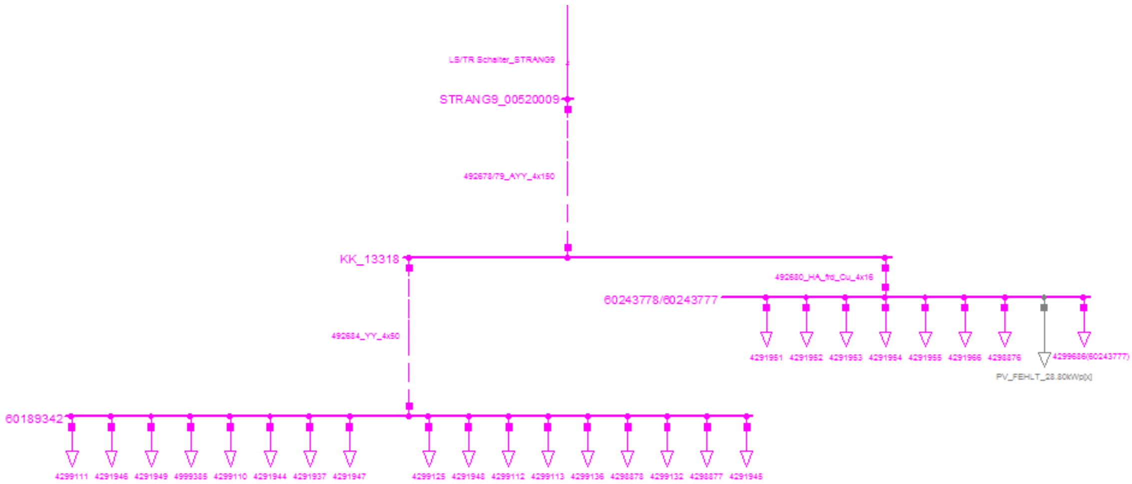
Cluster Centres

Electrical cluster 4 – Network 08 feeder 2

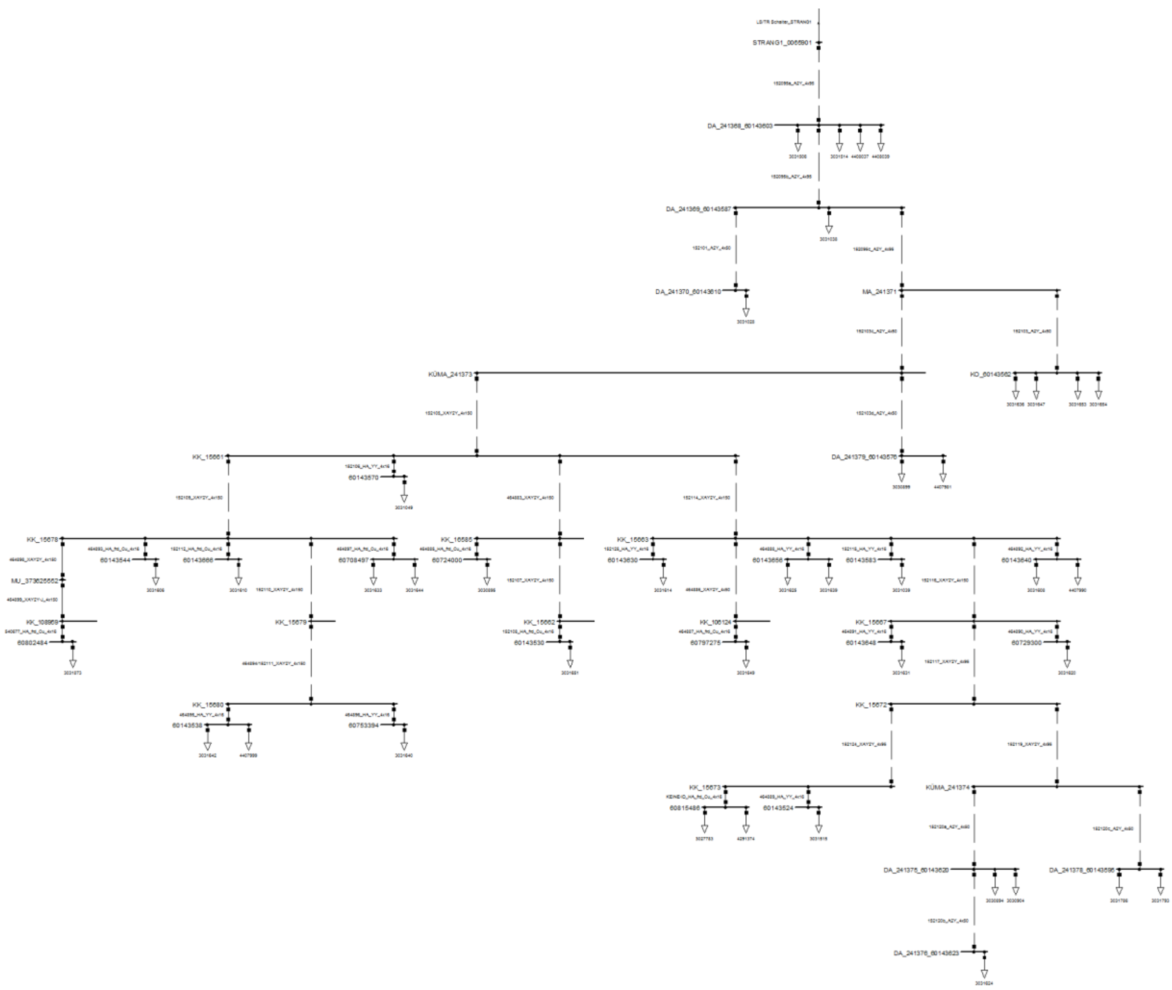


Cluster Centres

Electrical cluster 5 – Network 05 feeder 9

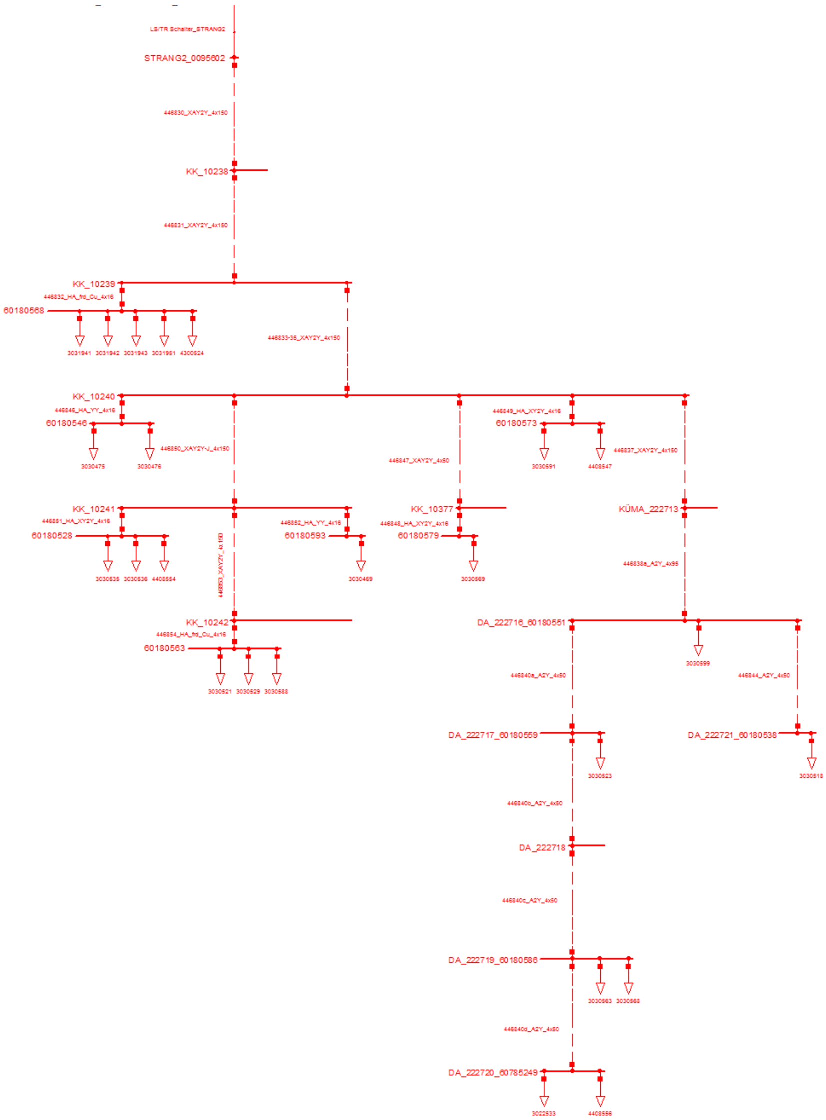


Electrical cluster 6 – Network 10 feeder 1



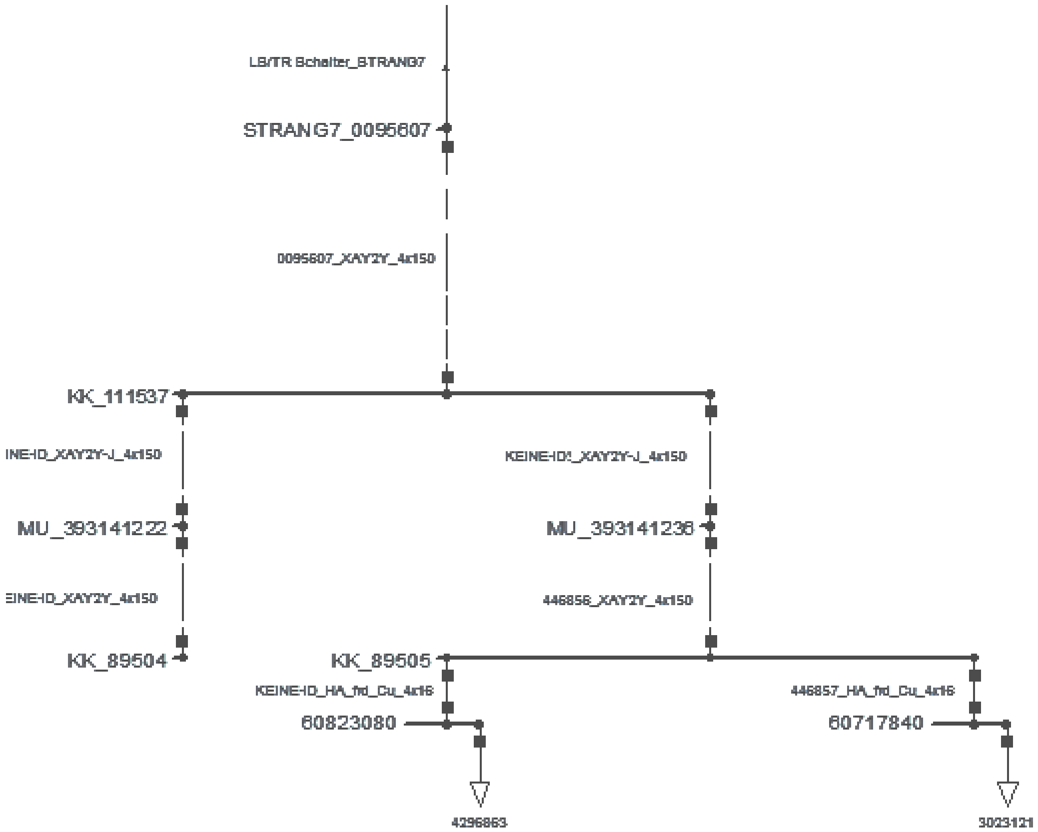
Cluster Centres

Electrical cluster 8 – Network 15 feeder 1

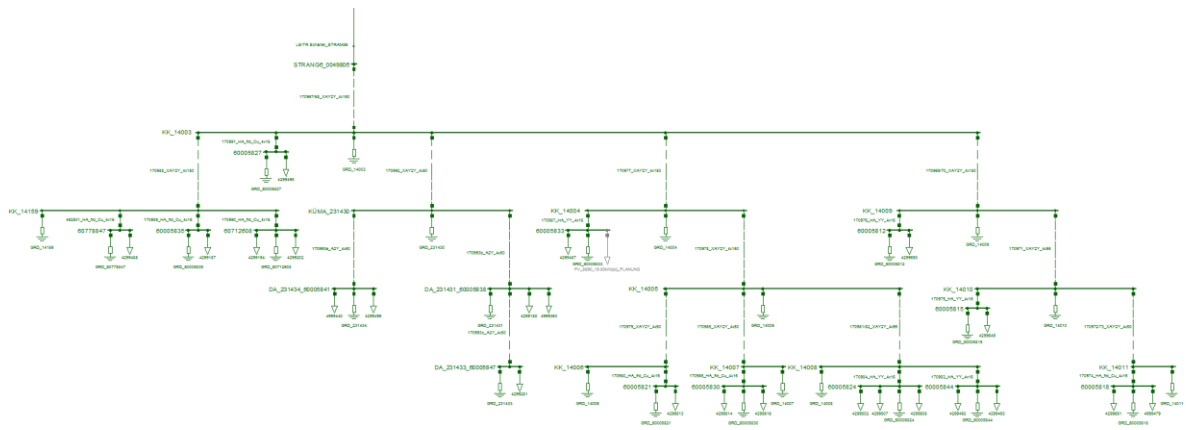


Cluster Centres

Electrical cluster 9 – Network 15 feeder 5

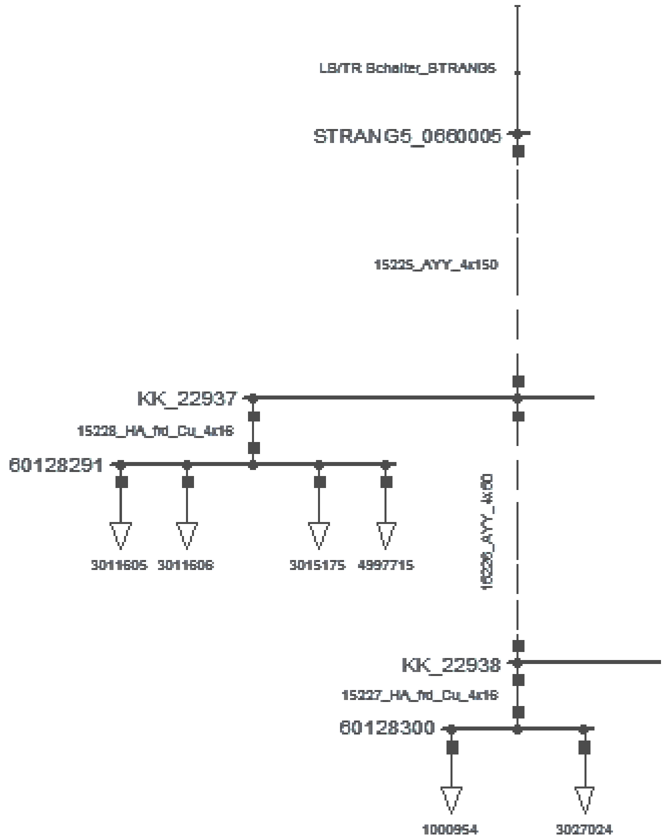


Non-electrical cluster 1 - Network 04 feeder 6



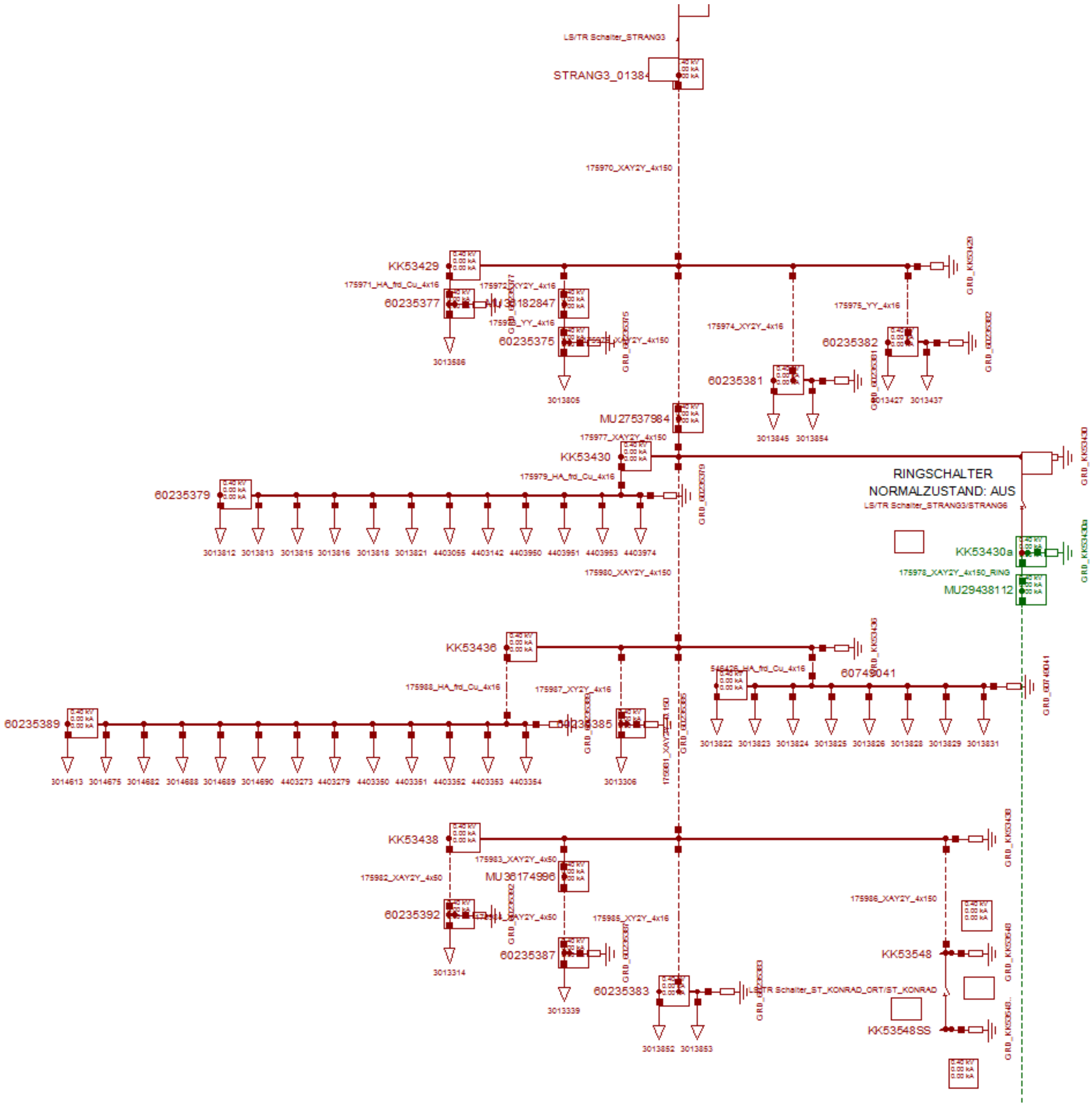
Cluster Centres

Non-electrical cluster 2 – Network 03 feeder 5



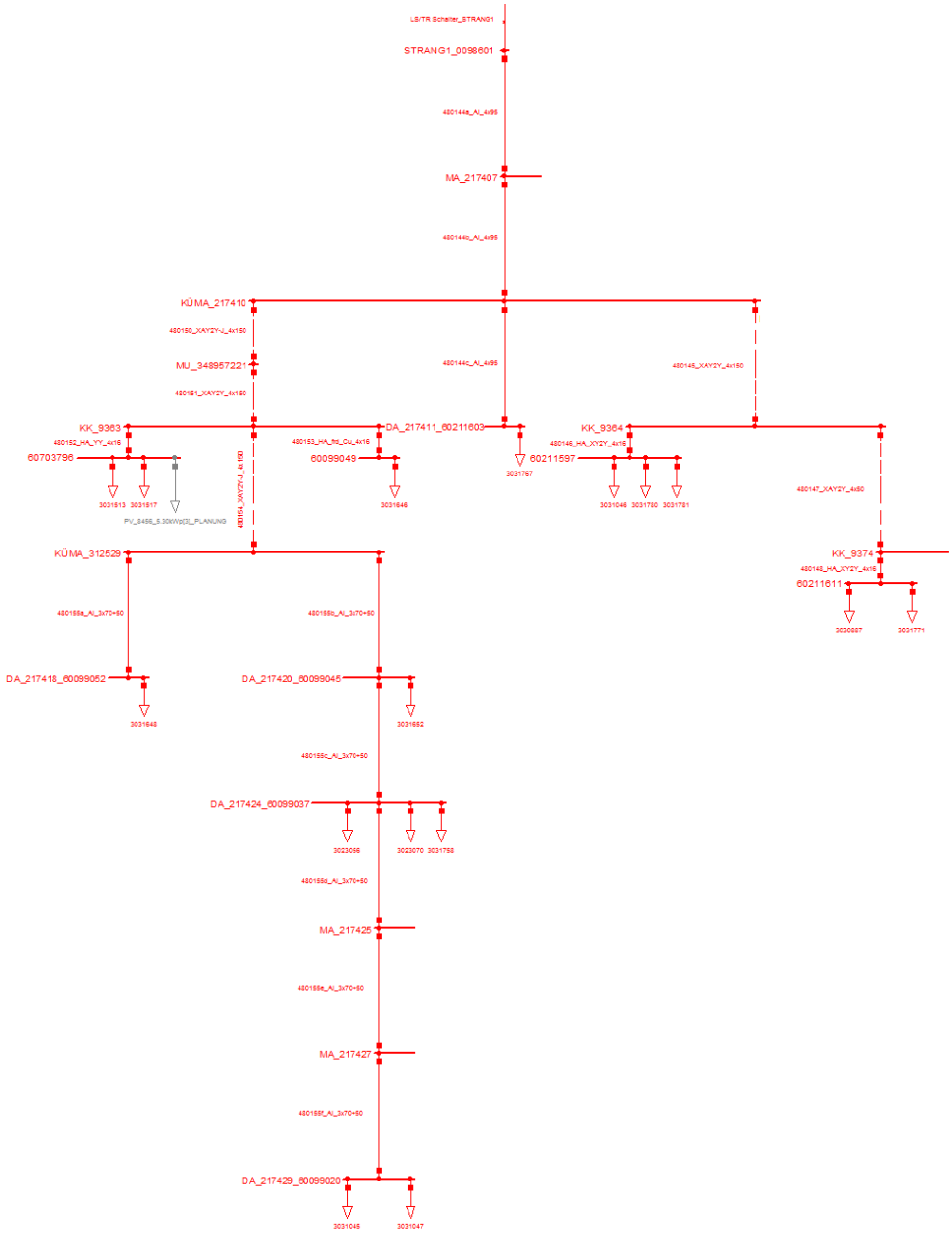
Cluster Centres

Non-electrical cluster 3 – Network 01 feeder 3



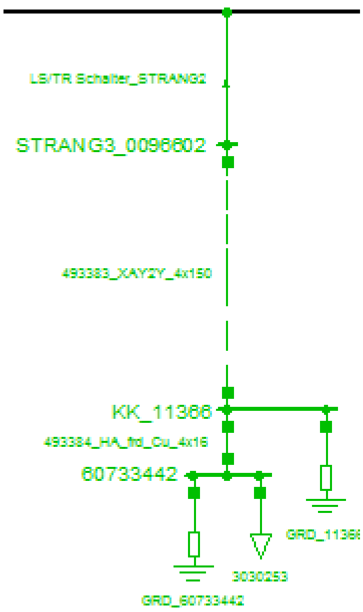
Cluster Centres

Non-electrical cluster 4 – Network 20 feeder 1

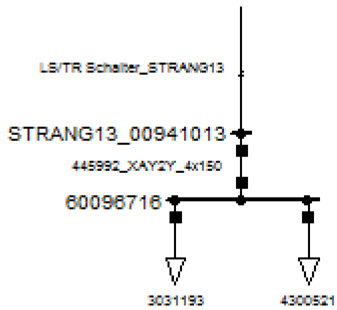


Cluster Centres

Non-electrical cluster 5 – Network 17 feeder 2

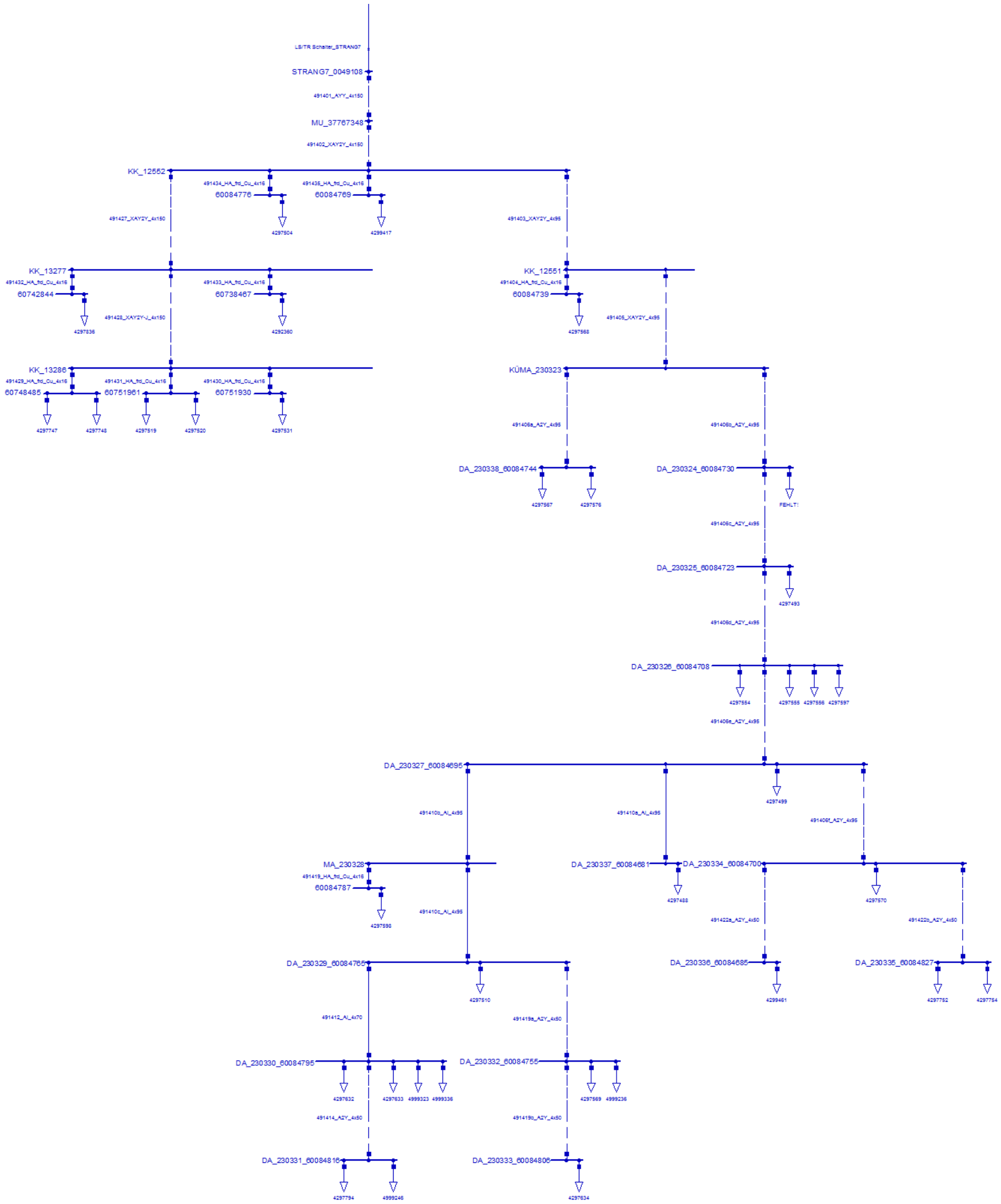


Non-electrical cluster 11 – Network 12 feeder 2



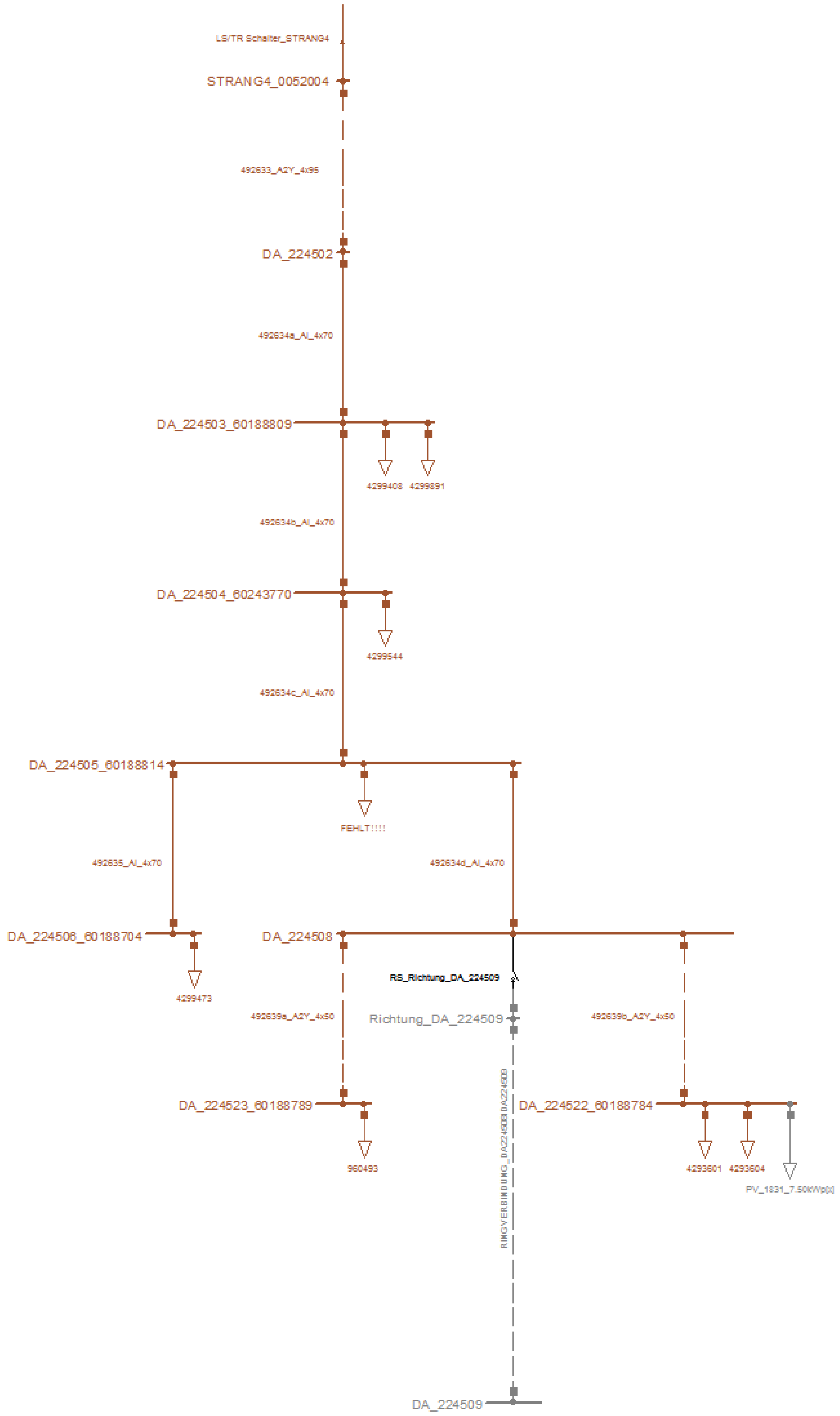
Cluster Centres

Non-electrical cluster 6 – Network 02 feeder 7



Cluster Centres

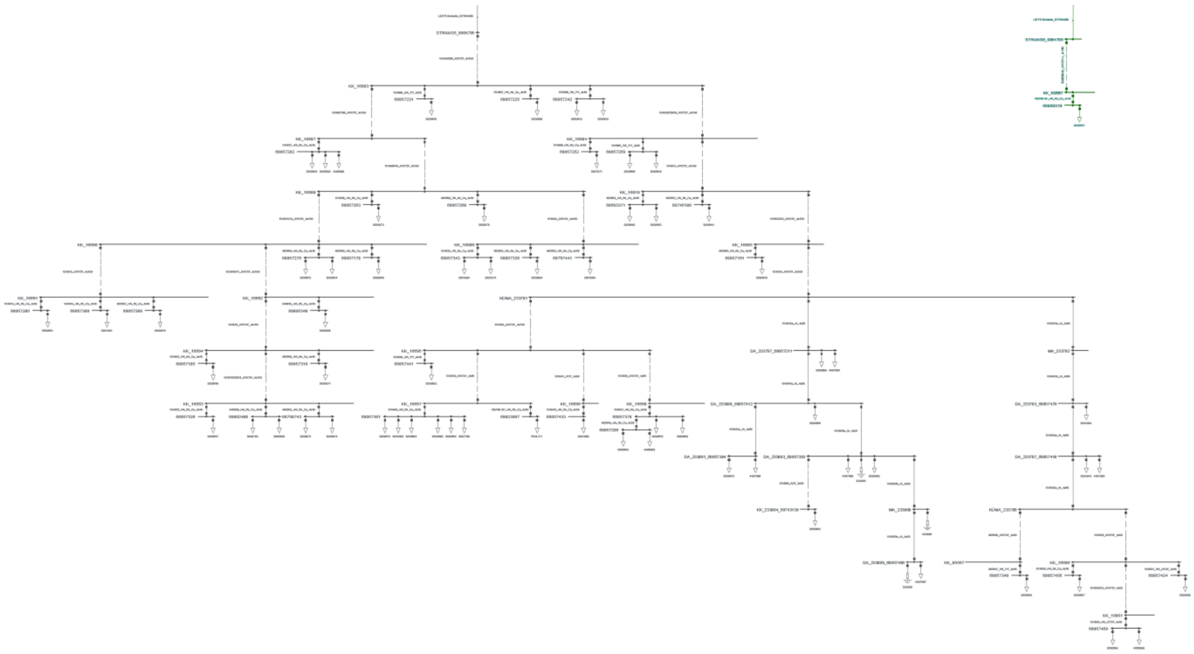
Non-electrical cluster 7 – Network 05 feeder 4



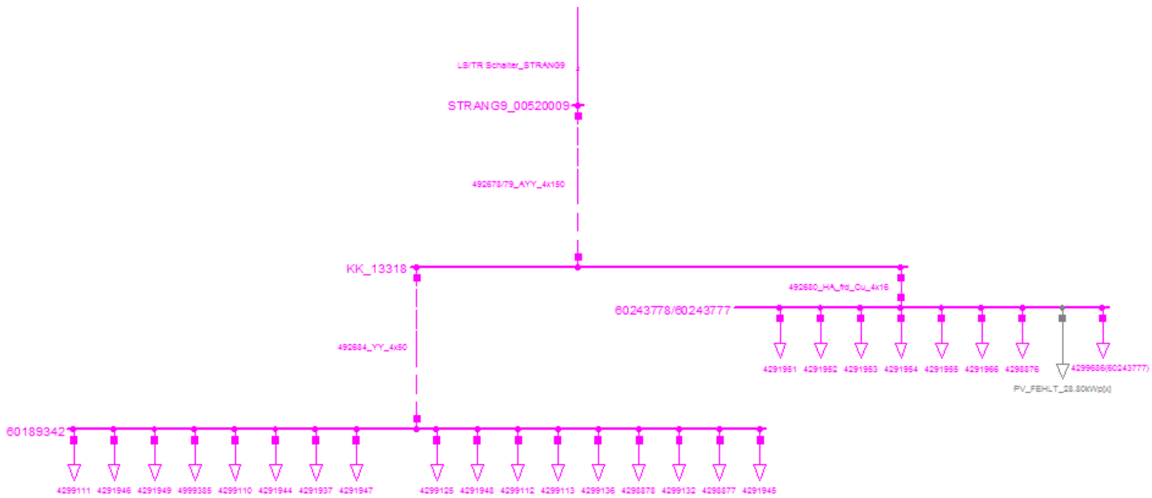
Cluster Centres

Non-electrical cluster 10 - Network 09 feeder 5 (left - electrical and non-electrical outlier)

Non-electrical cluster 8 – Network 09 feeder 6 (right - non-electrical outlier)



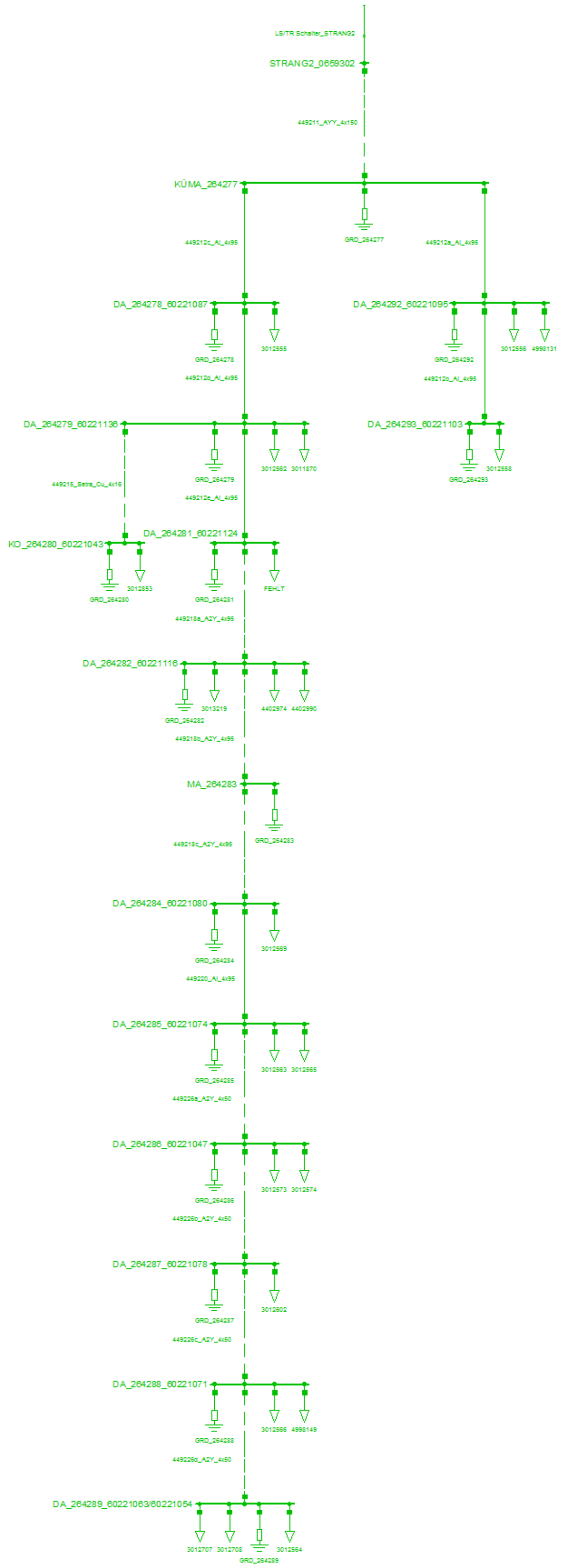
Non-electrical cluster 9 – Network 05 feeder 9



Cluster Centres

Global feeder median (electrical

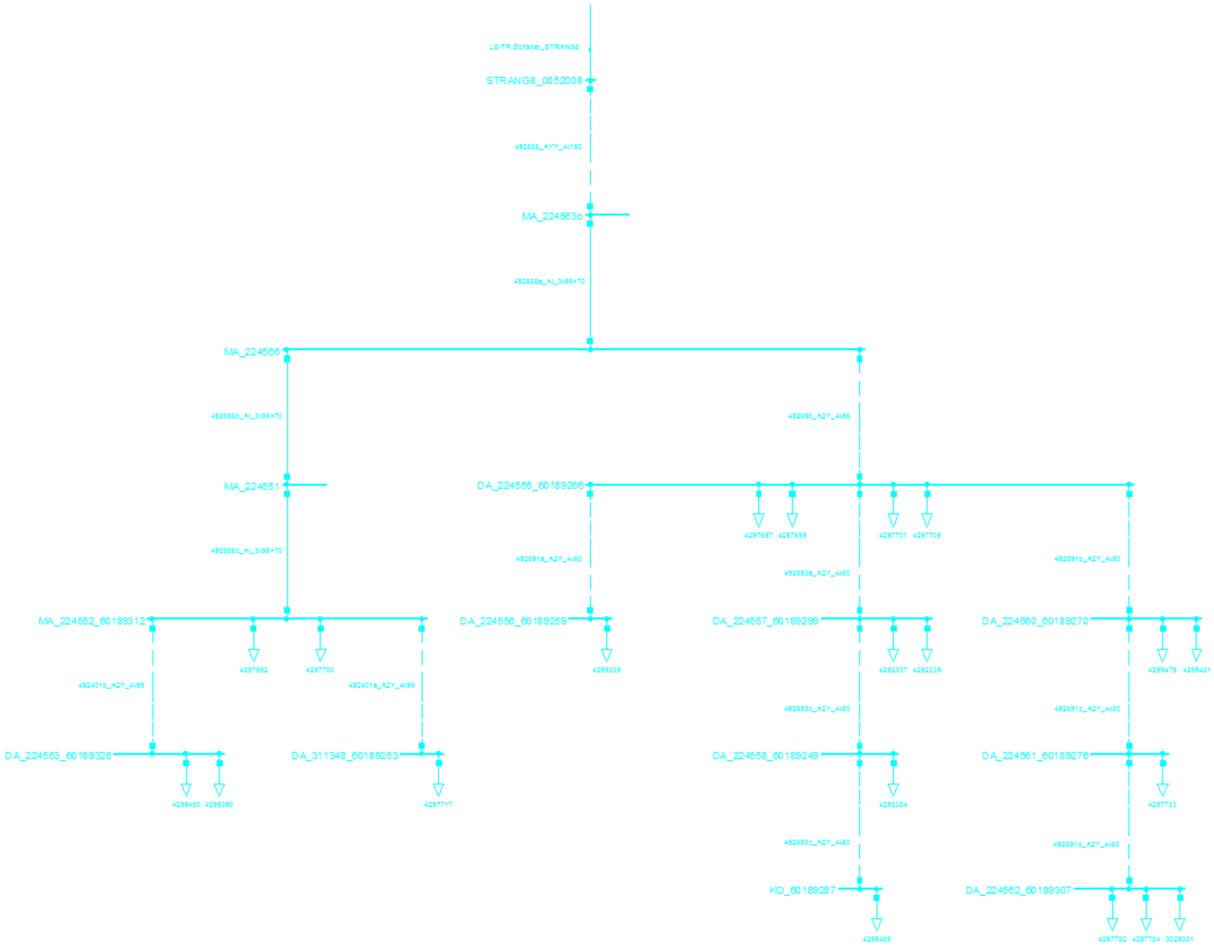
Network 30 feeder 5



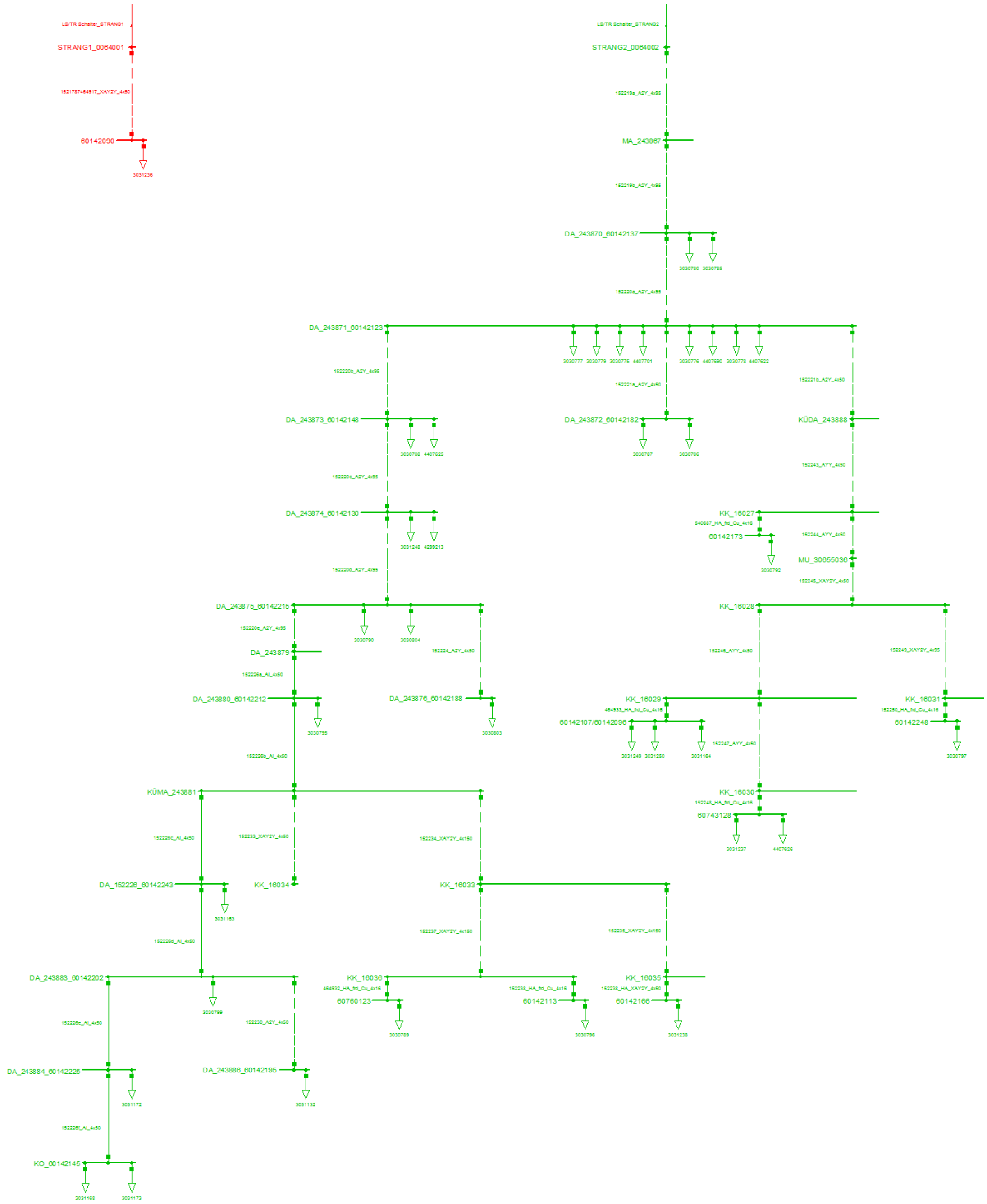
Cluster Centres

Global feeder median (non-electrical)

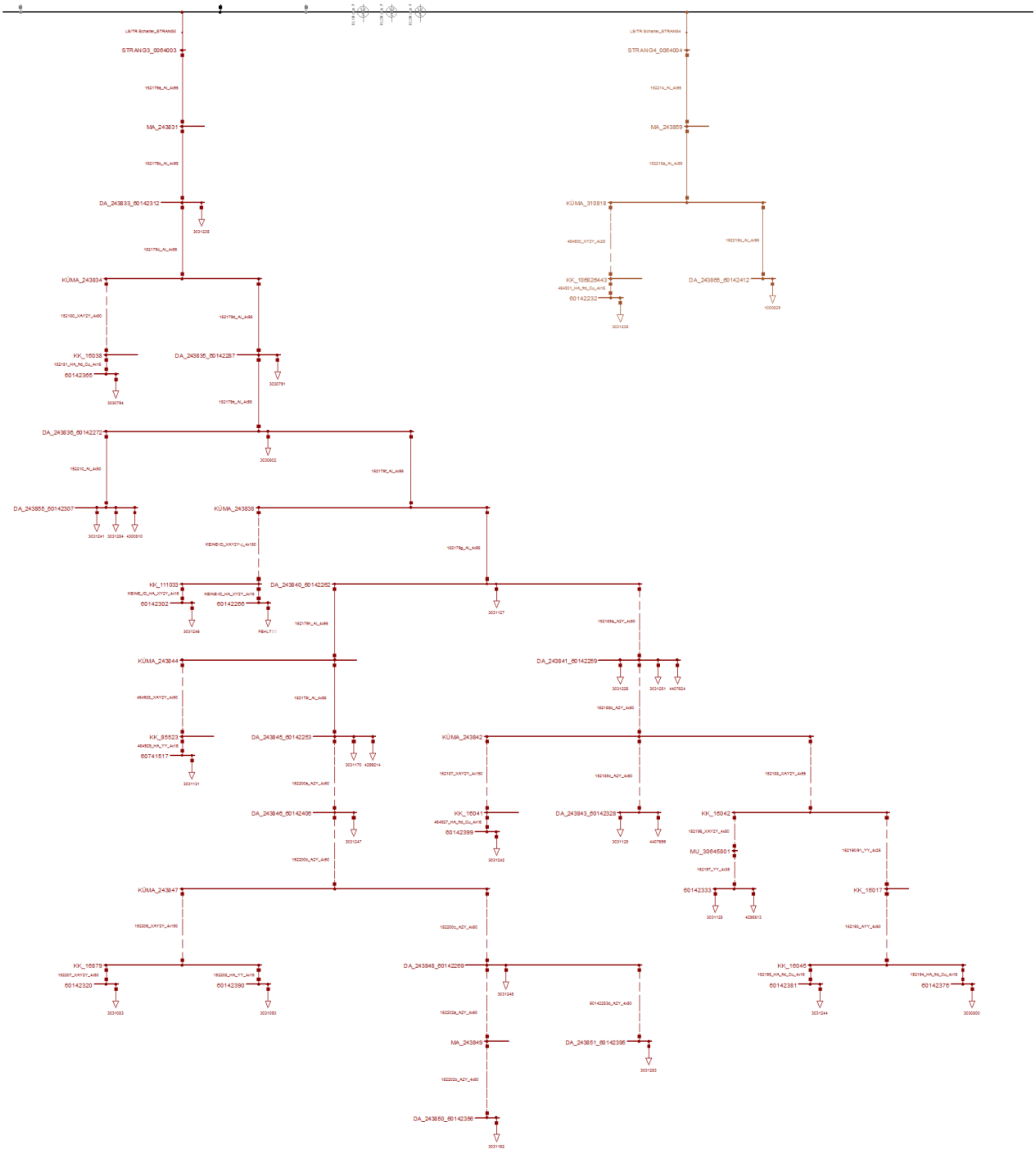
Network 05 feeder 8



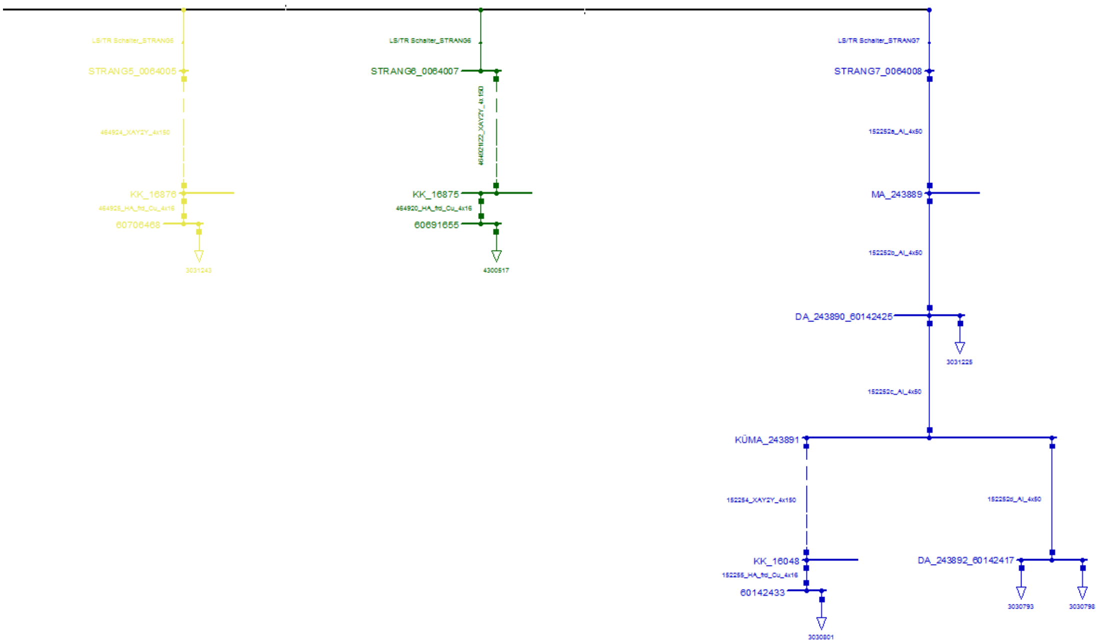
Network 08 (electrical outlier)



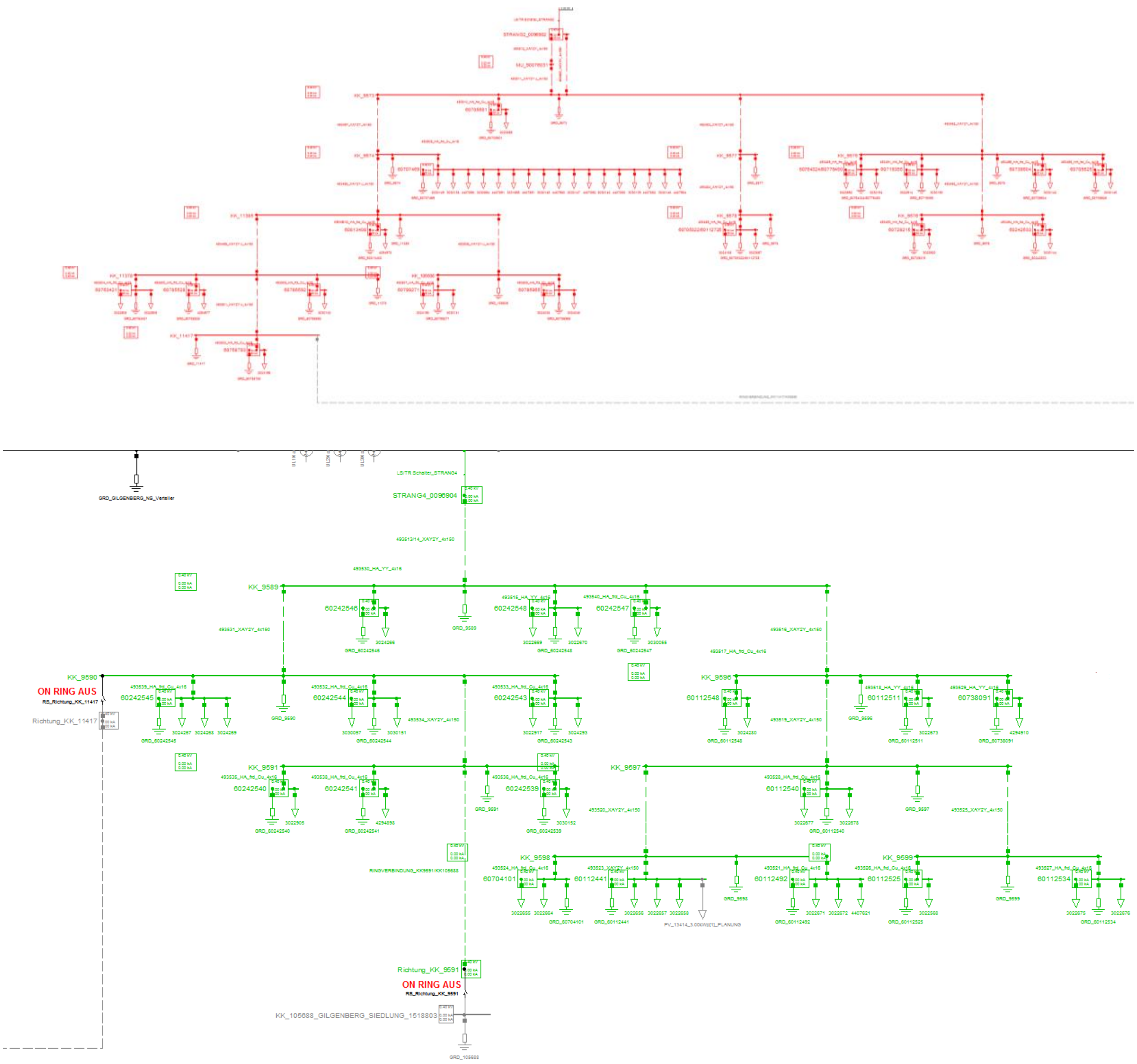
Network 08 (electrical outlier)



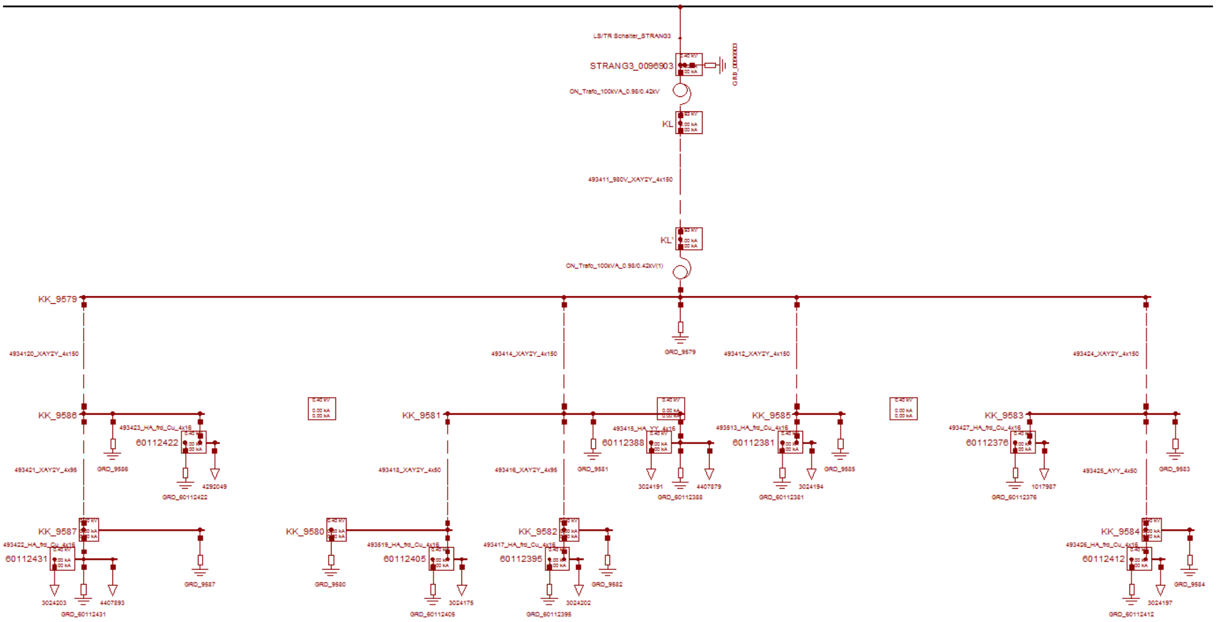
Network 08 (electrical outlier)

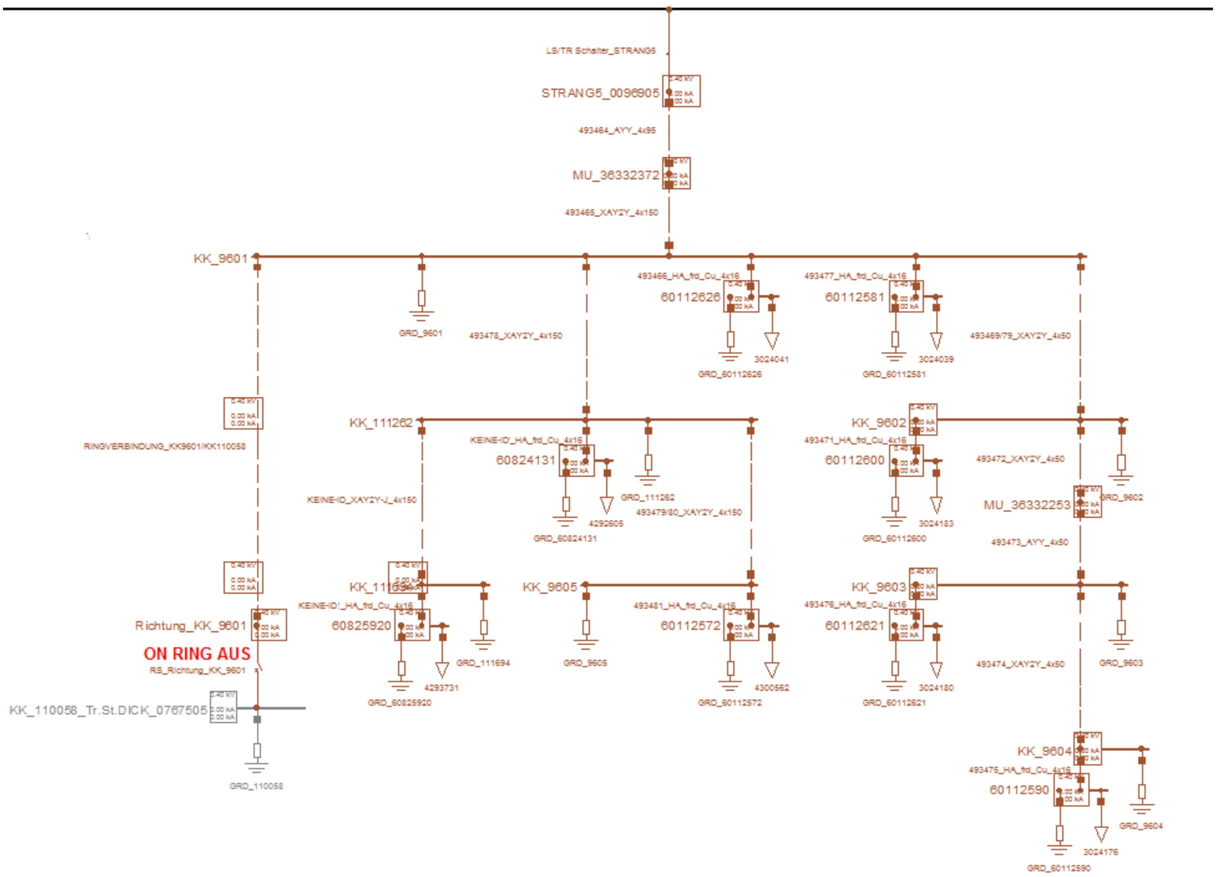


Network 18 (electrical and non-electrical outlier)

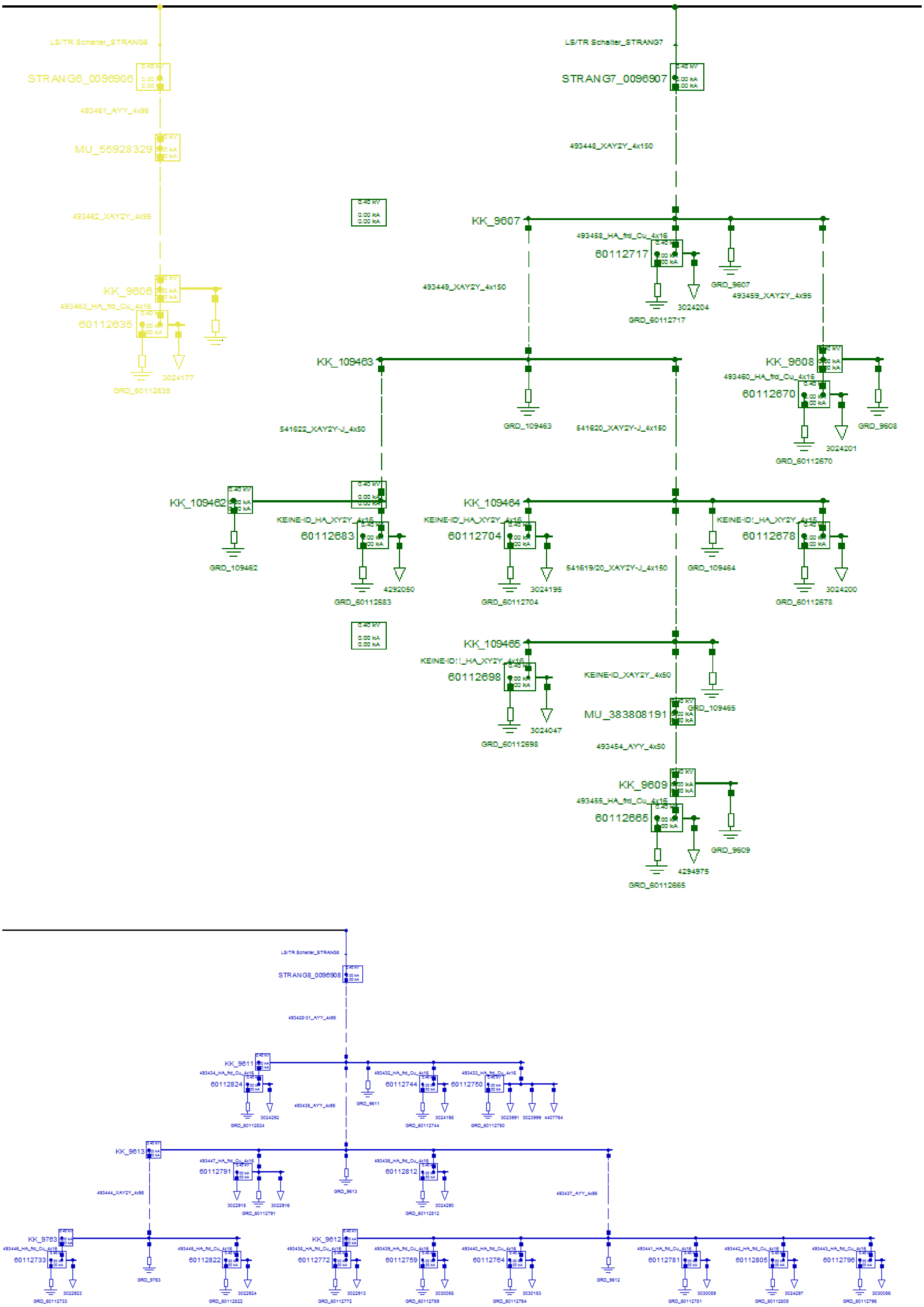


Network 18 (electrical and non-electrical outlier)

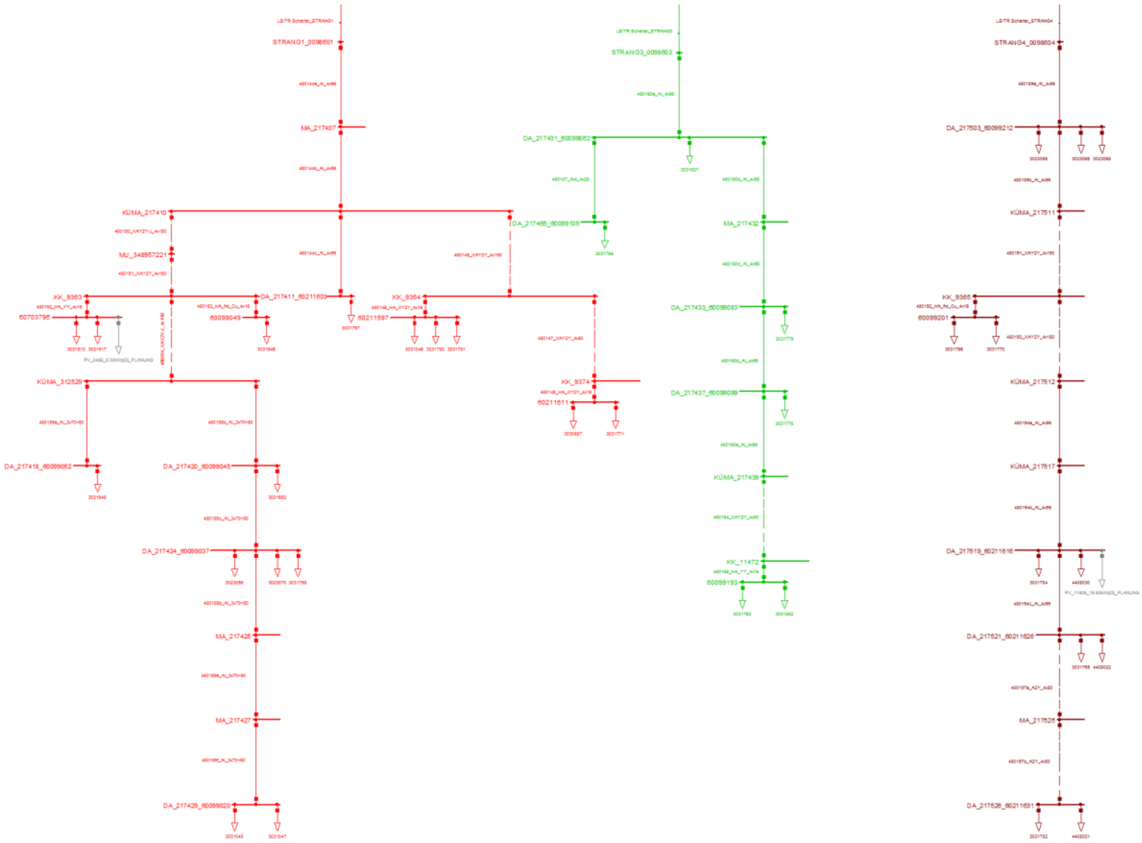




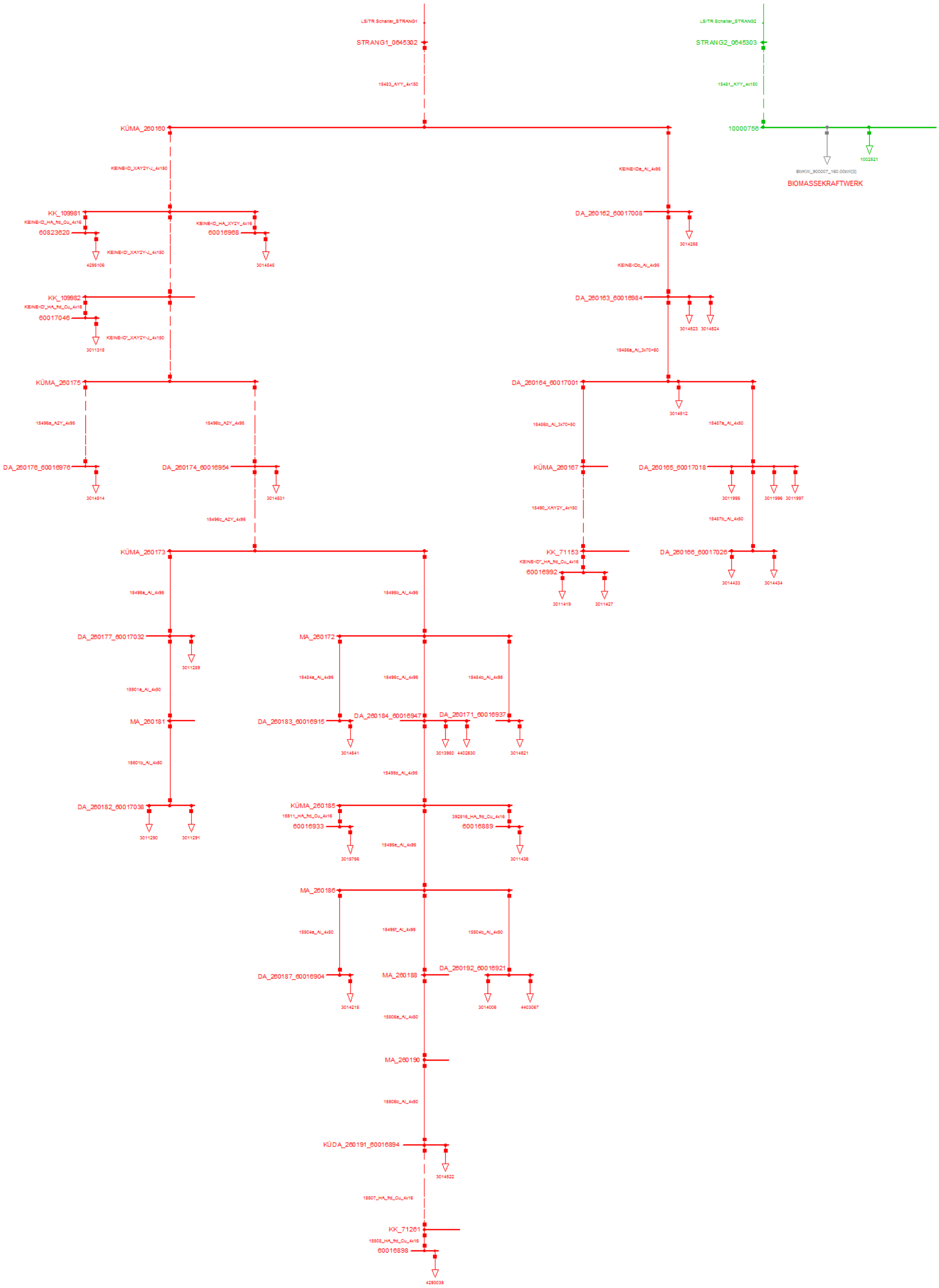
Network 18 (electrical and non-electrical outlier)



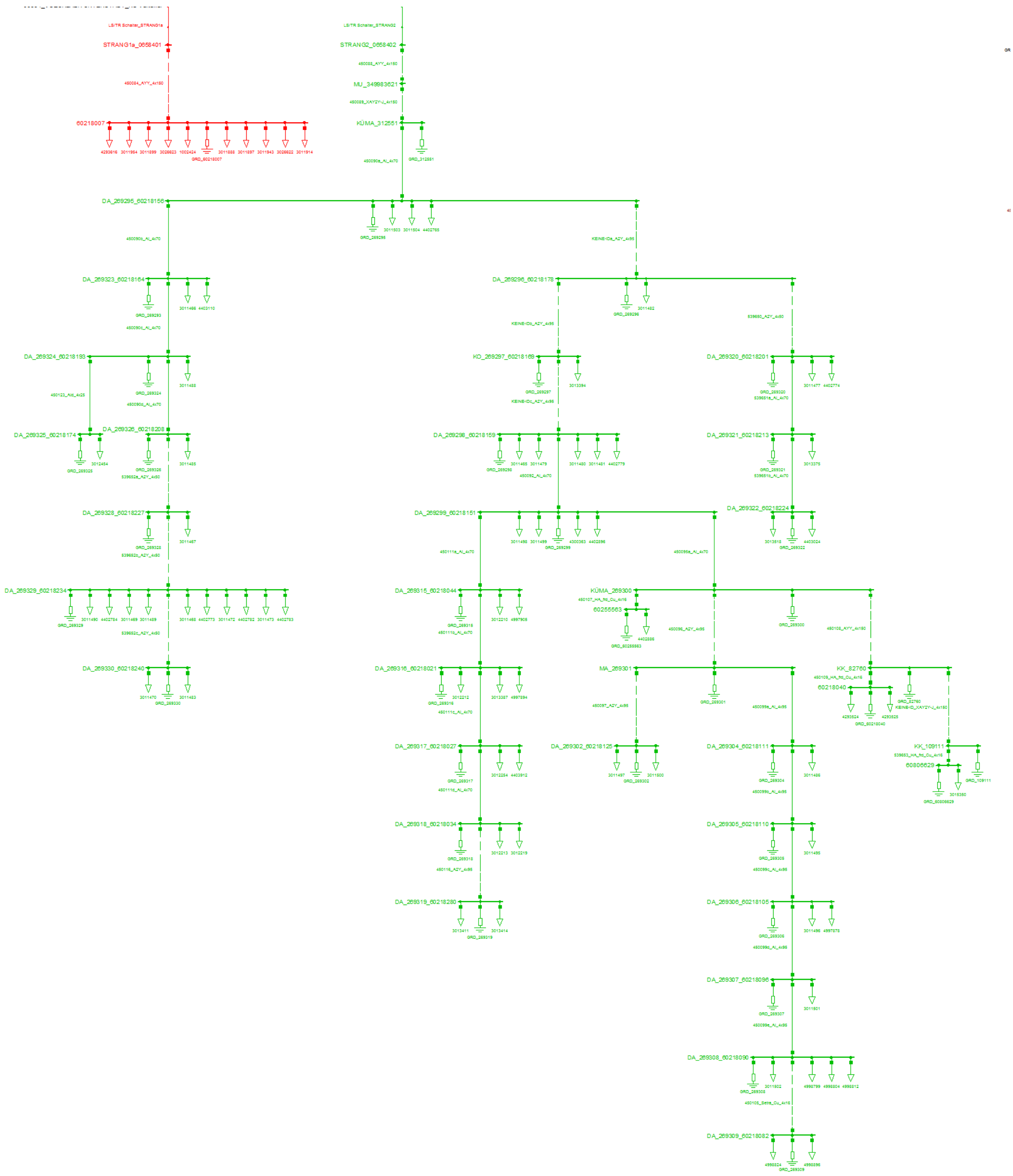
Network 20 (electrical outlier)



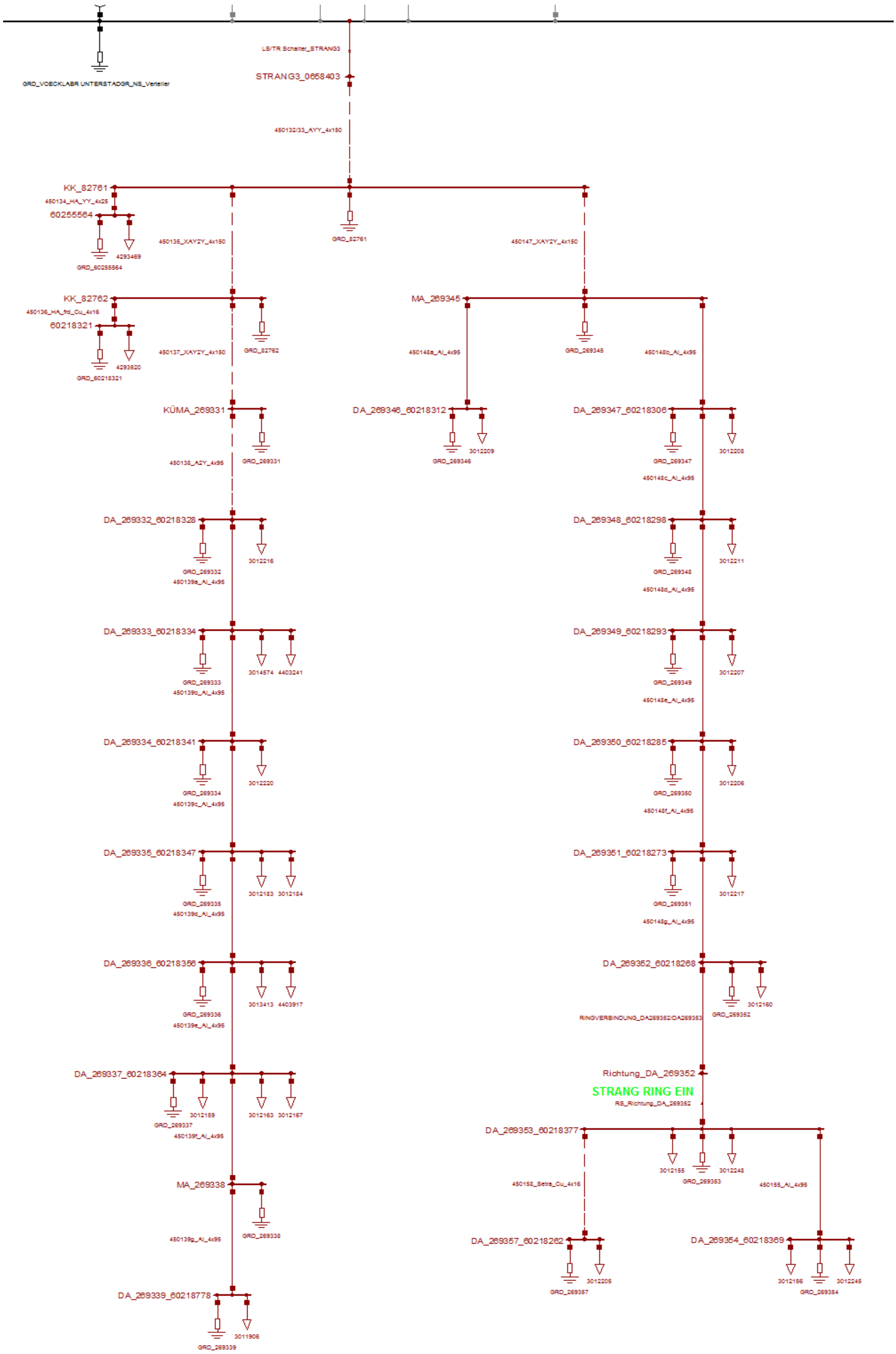
Network 27 (non-electrical outlier)



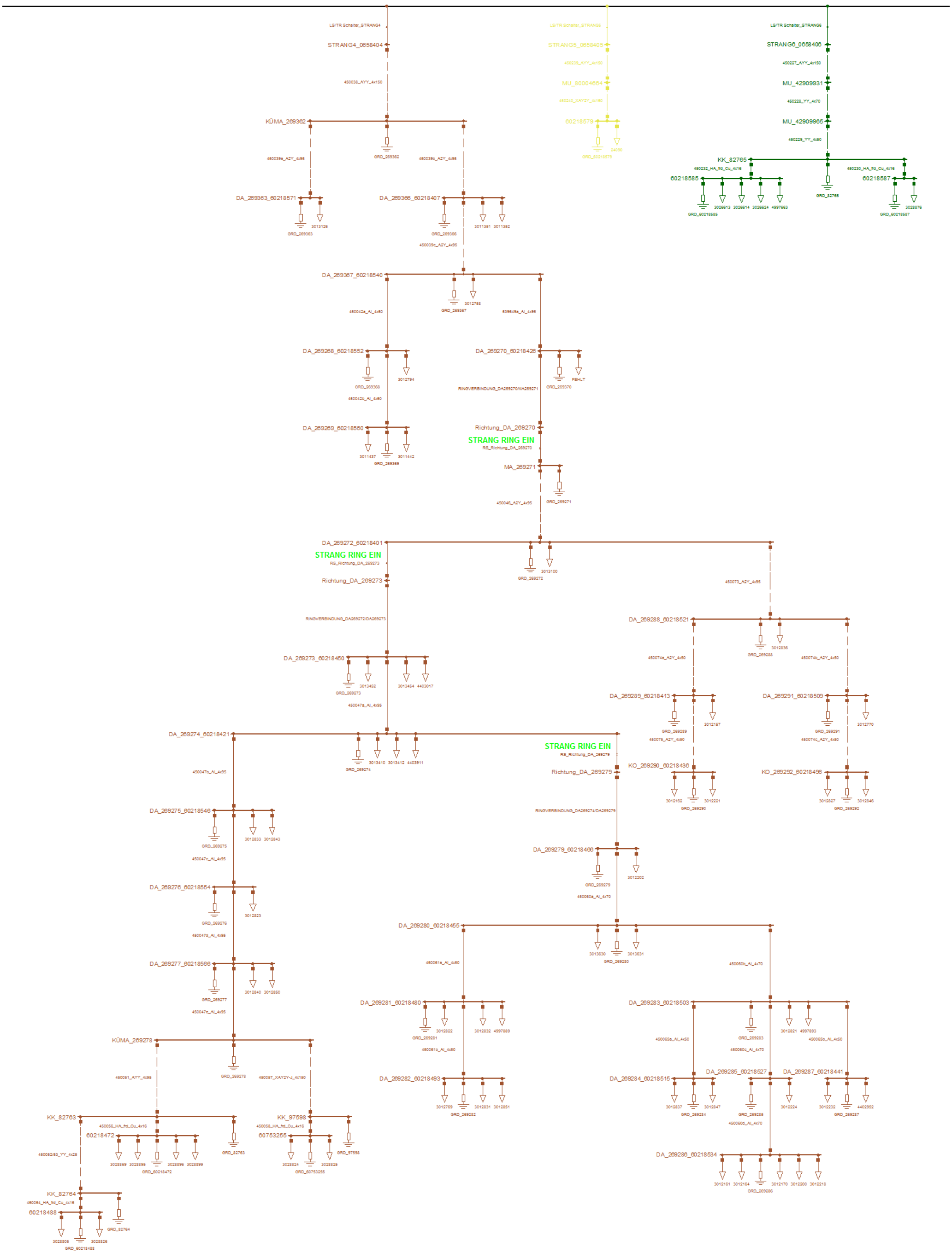
Network 29 (non-electrical outlier)



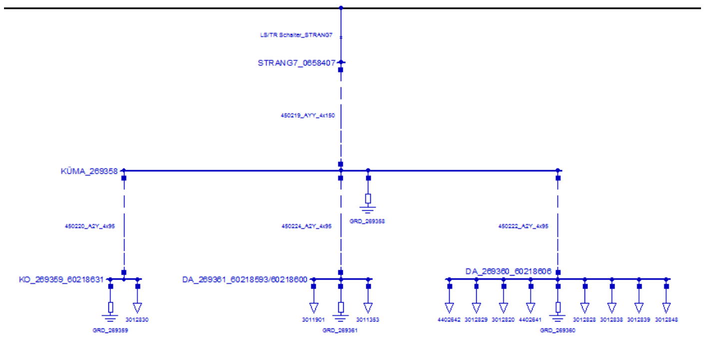
Network 29 (non-electrical outlier)

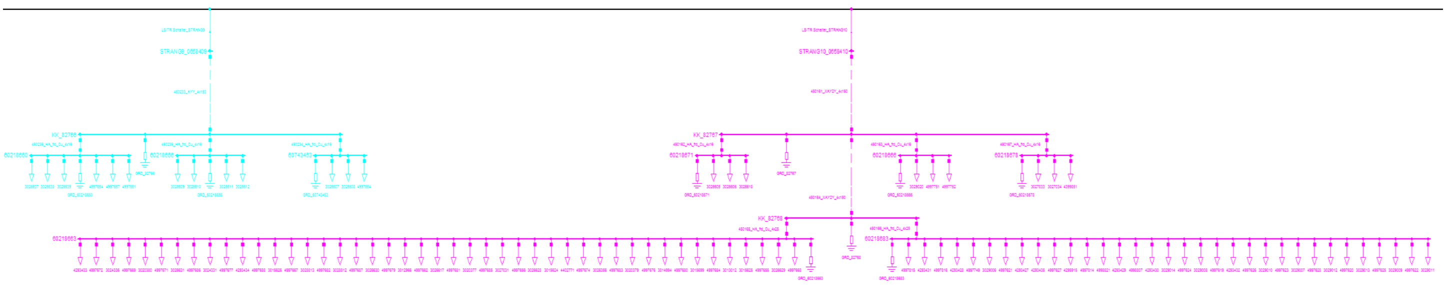


Network 29 (non-electrical outlier)

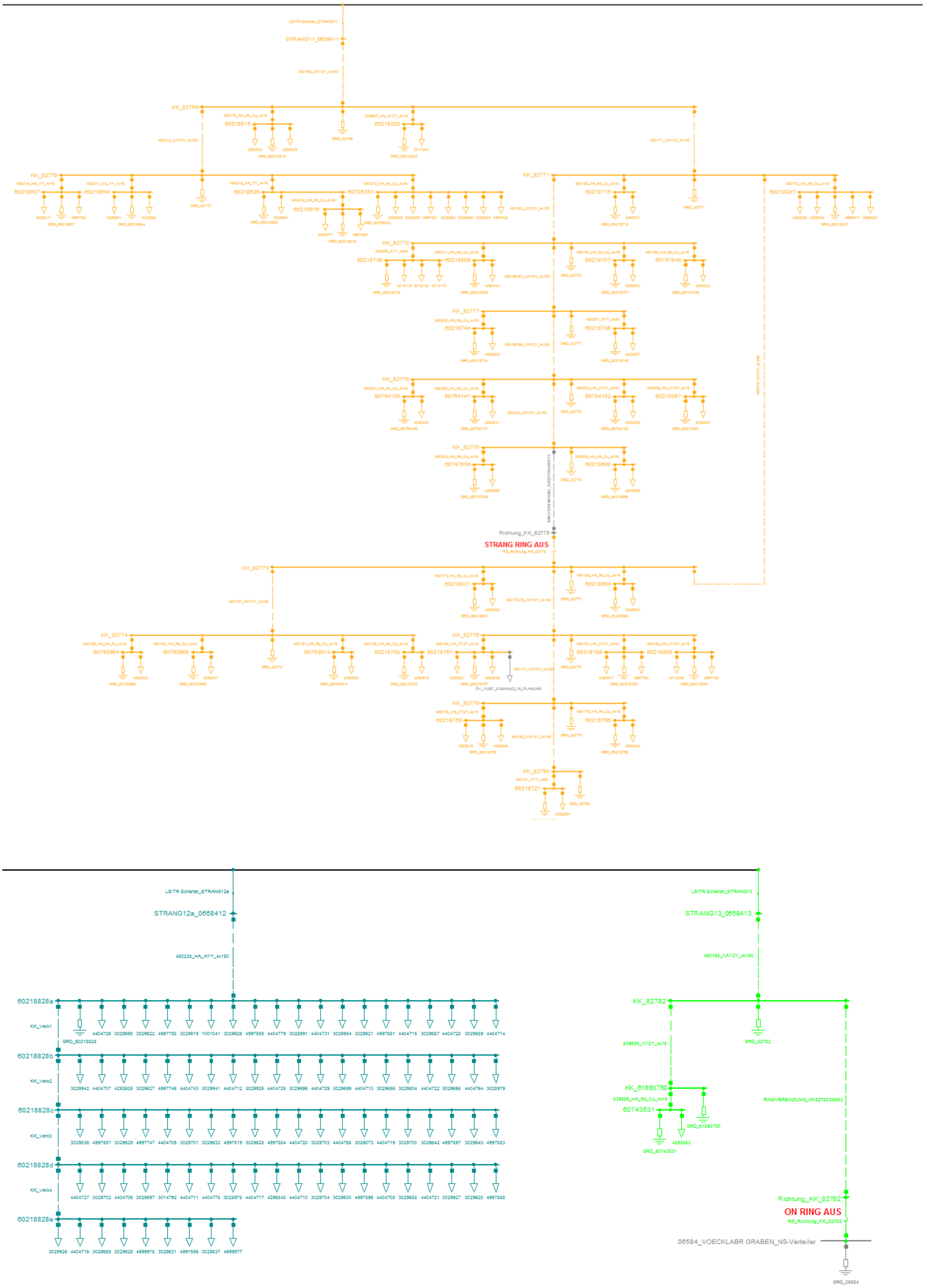


Network 29 (non-electrical outlier)

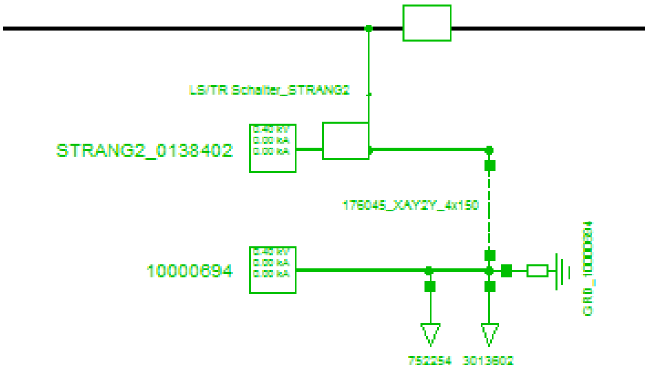




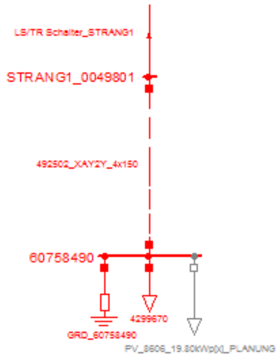
Network 29 (non-electrical outlier)



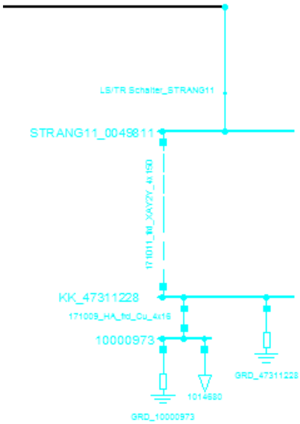
Network 01 feeder 2 (electrical outlier)



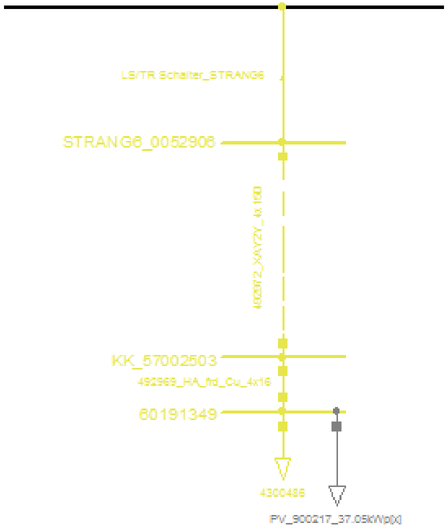
Network 04 feeder 1 (non-electrical outlier)



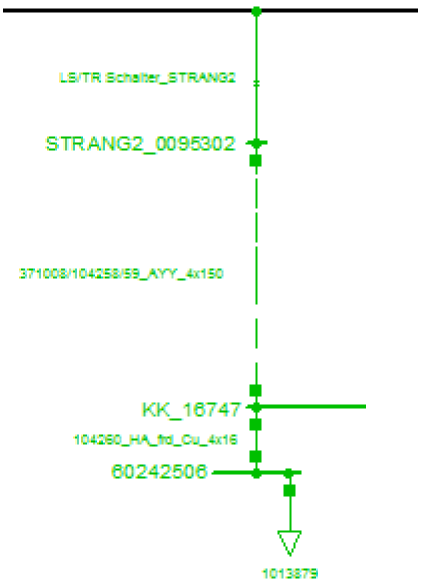
Network 04 feeder 8 (non-electrical outlier)



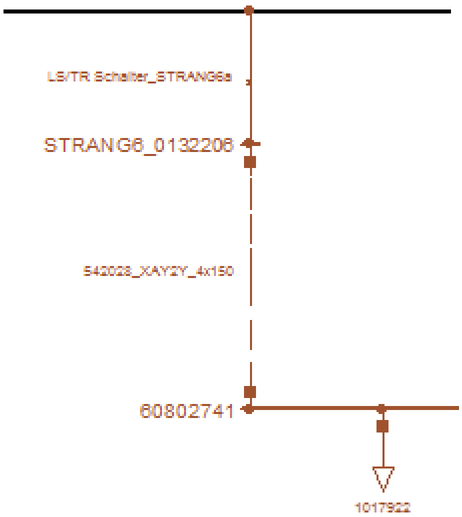
Network 06 feeder 5 (non-electrical outlier)



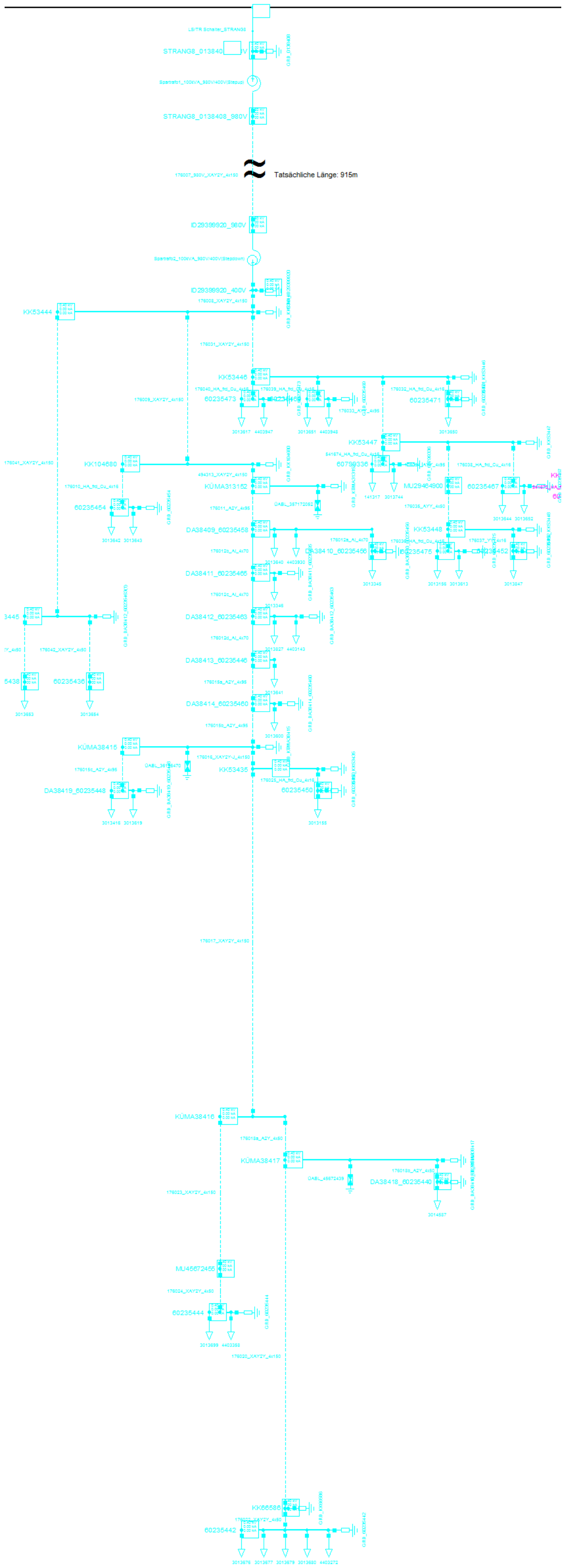
Network 14 feeder 2 (electrical and non-electrical outlier)



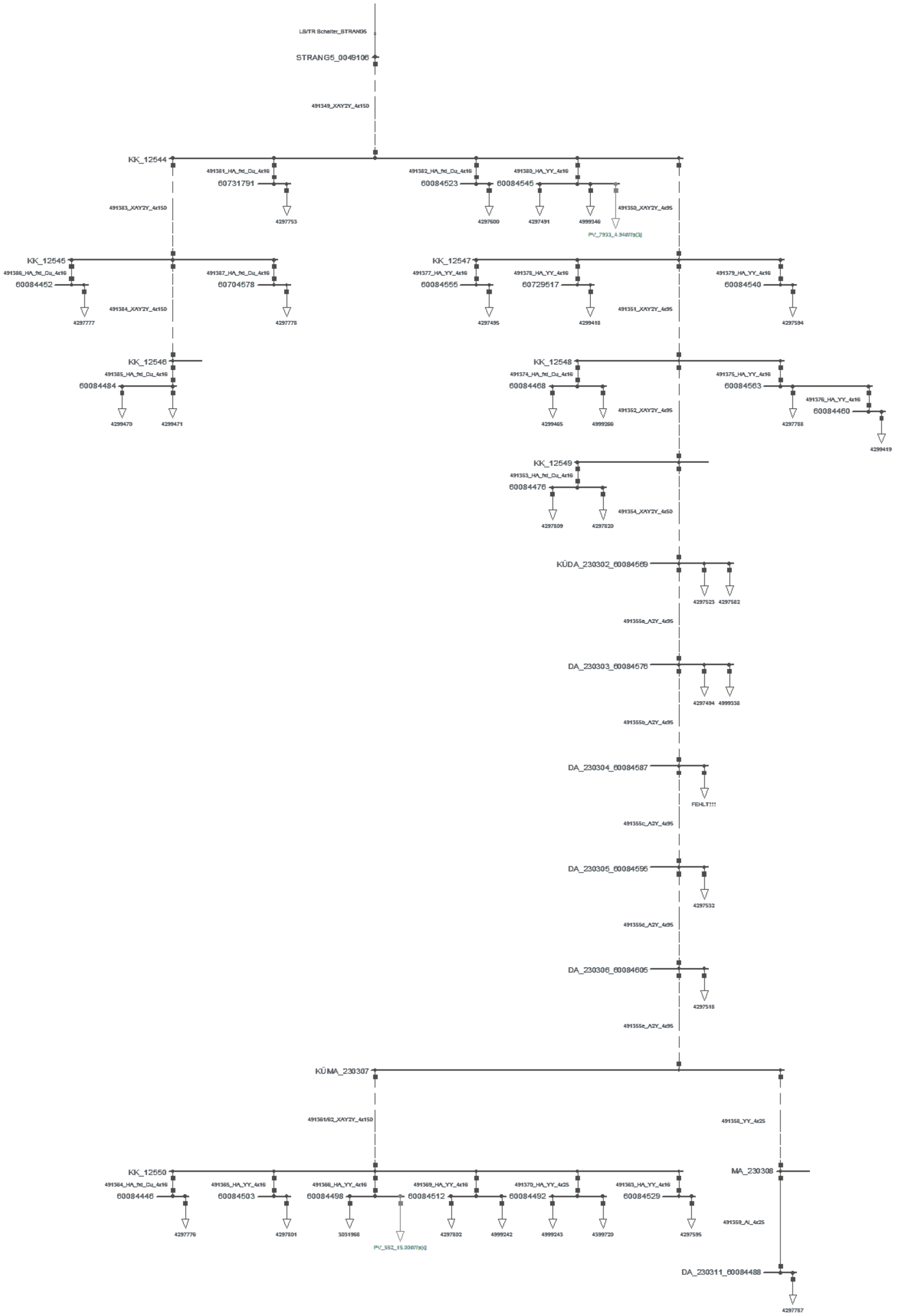
Network 23 feeder 4 (electrical outlier)



Network 01 feeder 8 (electrical outlier)



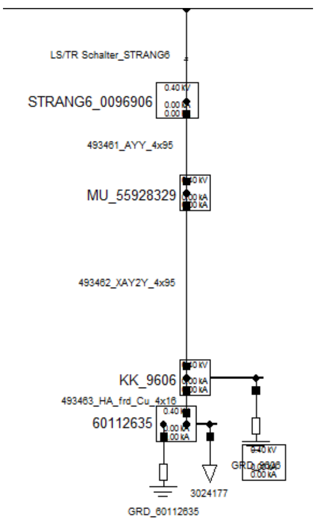
Network 03 feeder 5 (electrical outlier)



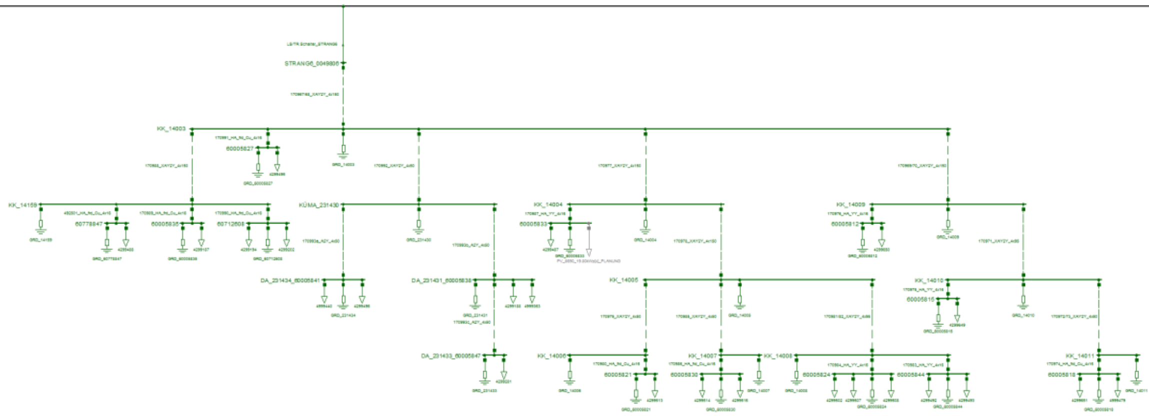
LS/TR Schalter_STRANG5
STRANG5_0049106
491349_XAY2Y_4x150
KK_12544
491381_HA_fhd_Cu_4x16
60731791
4297753
491382_HA_fhd_Cu_4x16
60084523
4297600
491380_HA_YY_4x16
60084545
4297491
4999346
491350_XAY2Y_4x95
PV_7933_4.94kWp[3]
491383_XAY2Y_4x150
KK_12545
491386_HA_fhd_Cu_4x16
60084452
4297777
491387_HA_fhd_Cu_4x16
60704578
4297778
491384_XAY2Y_4x150
KK_12547
491377_HA_YY_4x16
60084555
4297495
491378_HA_YY_4x16
60729517
4299418
491351_XAY2Y_4x95
491379_HA_YY_4x16
60084540
4297594
KK_12546
491385_HA_fhd_Cu_4x16
60084484
4299470
4299471
KK_12548
491374_HA_fhd_Cu_4x16
60084468
4299465
4999266
491352_XAY2Y_4x95
491375_HA_YY_4x16
60084563
4297788
491376_HA_YY_4x16
60084460
4299419
KK_12549
491353_HA_fhd_Cu_4x16
60084476
4297809
4297820
491354_XAY2Y_4x50
KÜDA_230302_60084569
4297523
4297582
491355a_A2Y_4x95
DA_230303_60084576
4297494
4999338
491355b_A2Y_4x95
DA_230304_60084587
FEHLT!!!
491355c_A2Y_4x95
DA_230305_60084595
4297532
491355d_A2Y_4x95
DA_230306_60084605
4297518
491355e_A2Y_4x95
KÜMA_230307
491361/62_XAY2Y_4x150
491358_YY_4x25
KK_12550
491364_HA_fhd_Cu_4x16
60084446
4297776
491365_HA_YY_4x16
60084503
4297801
491366_HA_YY_4x16
60084498
3031968
PV_552_15.00kWp[x]
491369_HA_YY_4x16
60084512
4297802
4999242
491370_HA_YY_4x25
60084492
4999243
4399720
491363_HA_YY_4x16
60084529
4297595
MA_230308
491359_Al_4x25
DA_230311_60084488
4297787

Network 02 feeder 1 (non-electrical outlier)

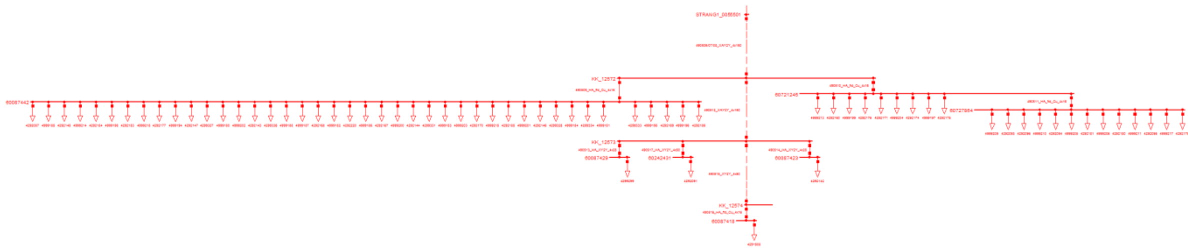
Network 18 feeder 5 (electrical outlier)



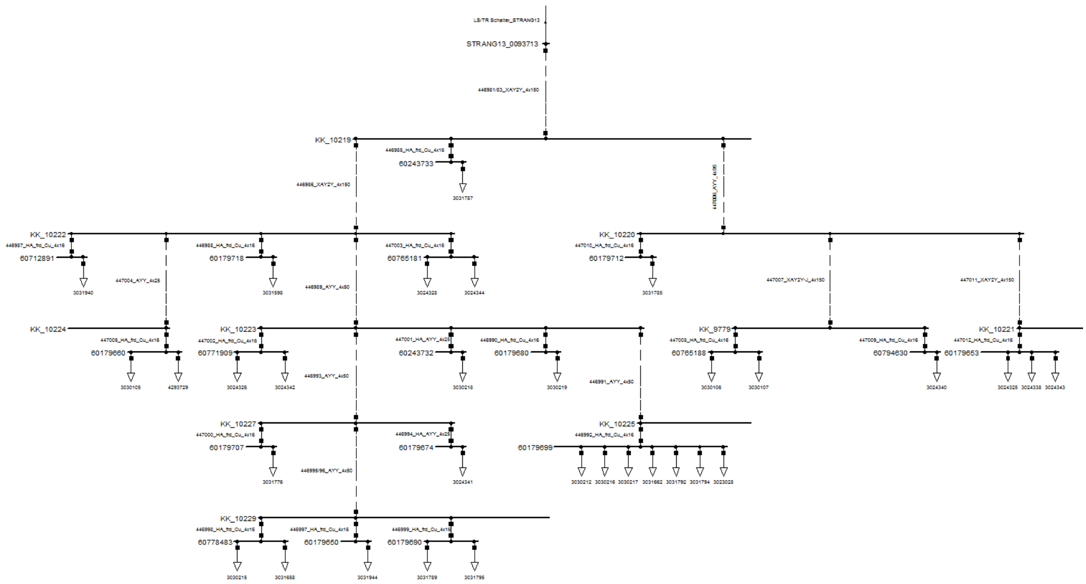
Network 04 feeder 6 (electrical outlier)



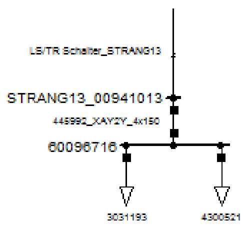
Network 07 feeder 1 (electrical and non-electrical outlier)



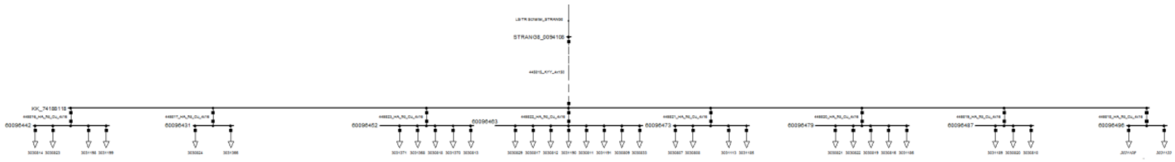
Network 11 feeder 8 (electrical outlier)



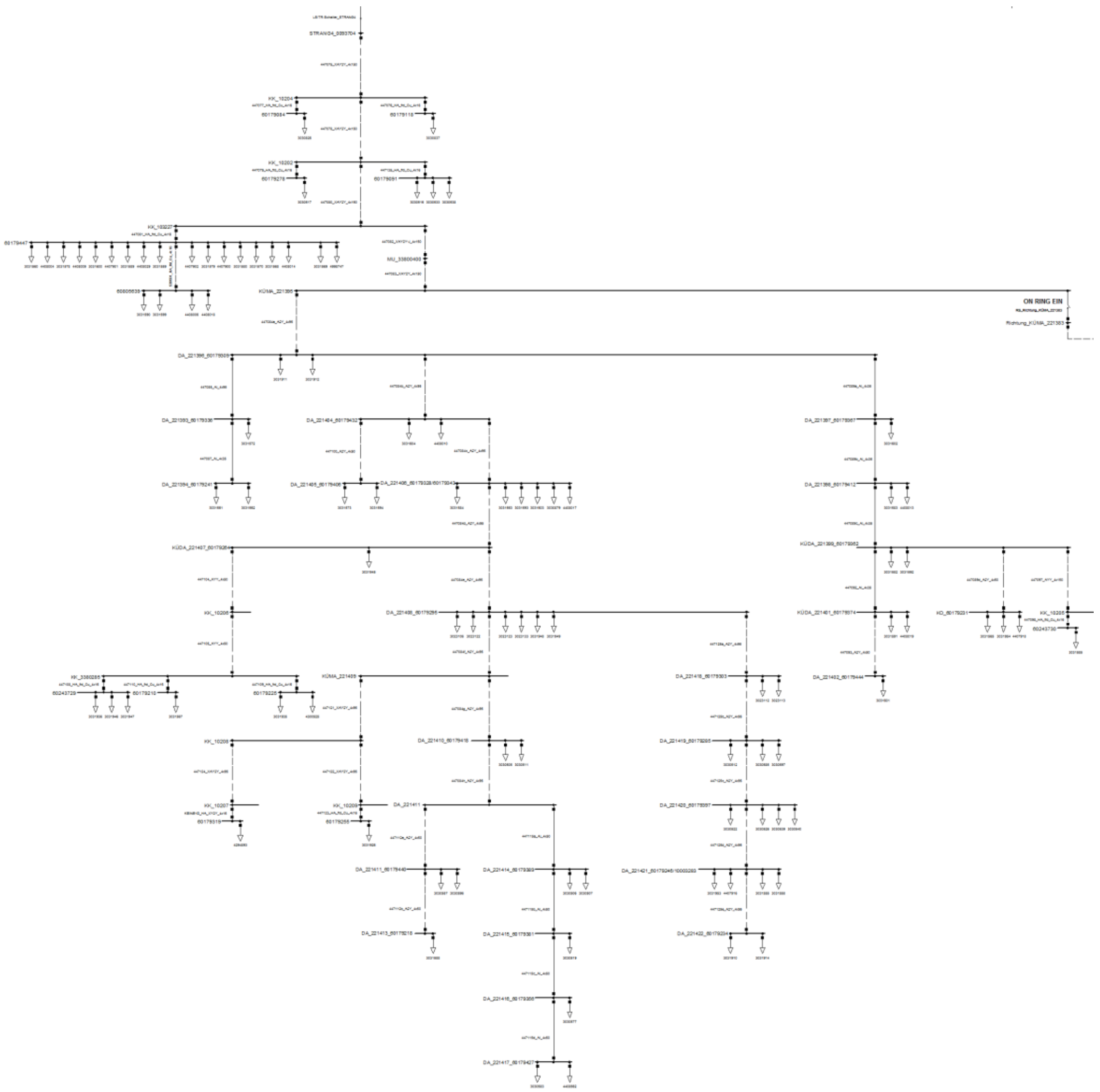
Network 12 feeder 2 (non-electrical outlier)



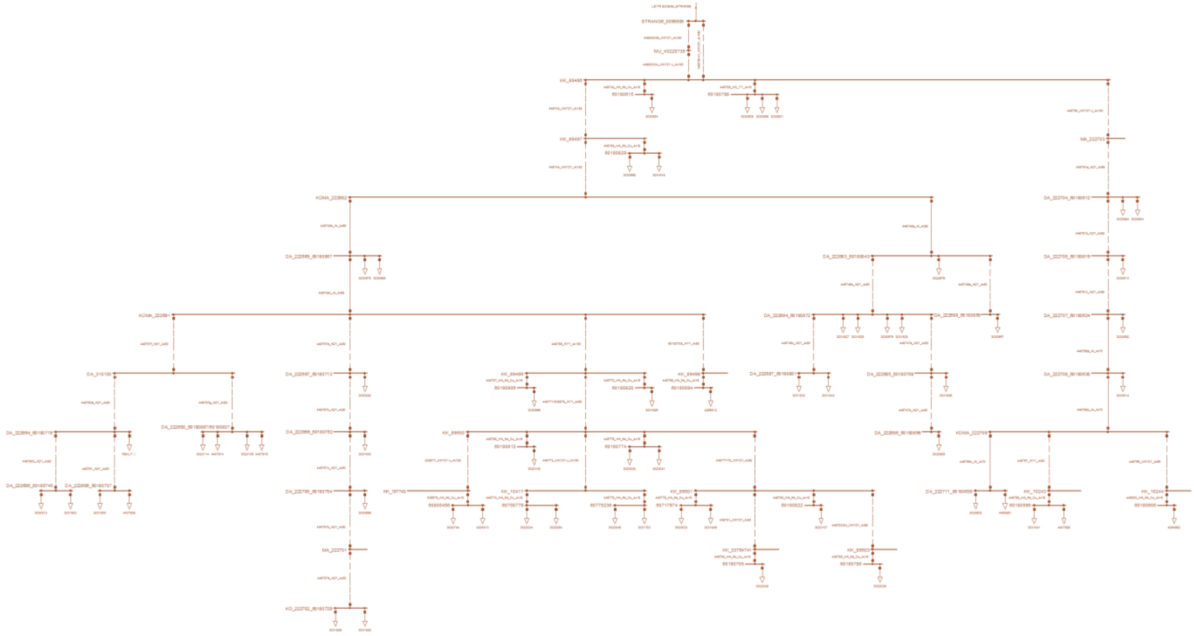
Network 12 feeder 7 (electrical outlier)



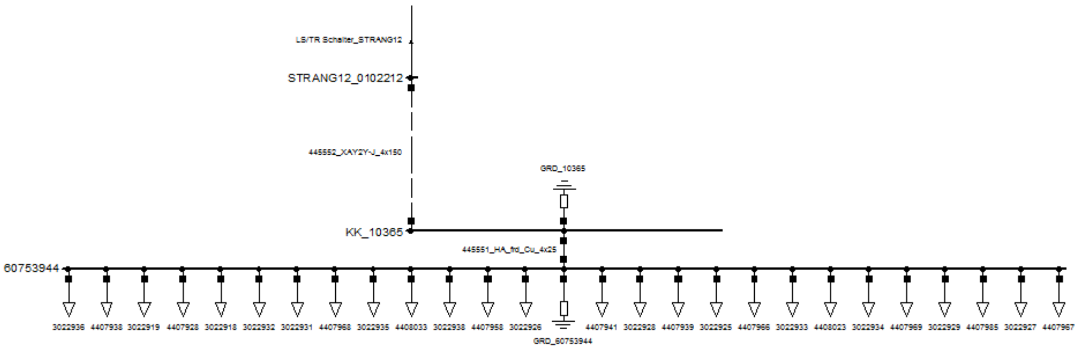
Network 11 feeder 2 (non-electrical outlier)



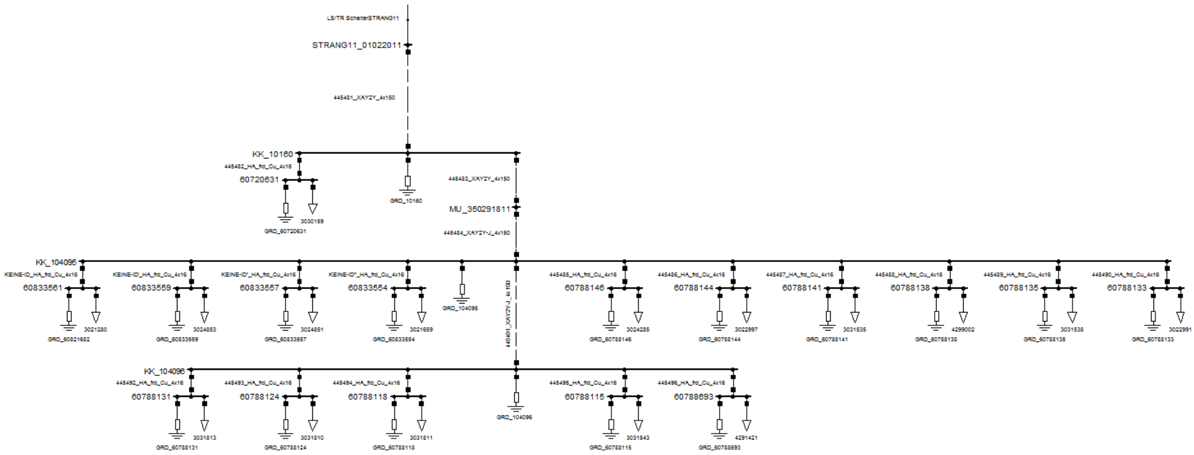
Network 15 feeder 4 (non-electrical outlier)



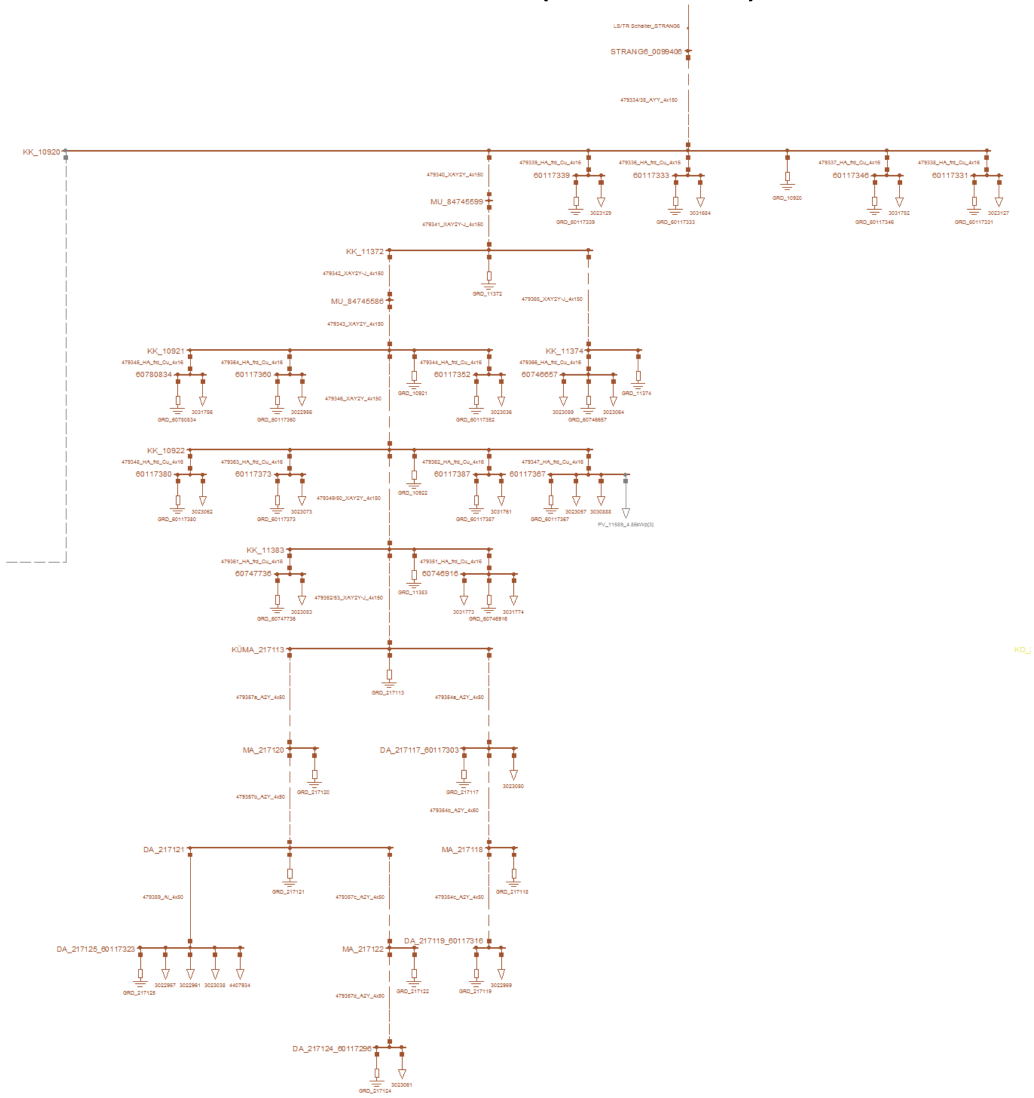
Network 22 feeder 2 (non-electrical outlier)



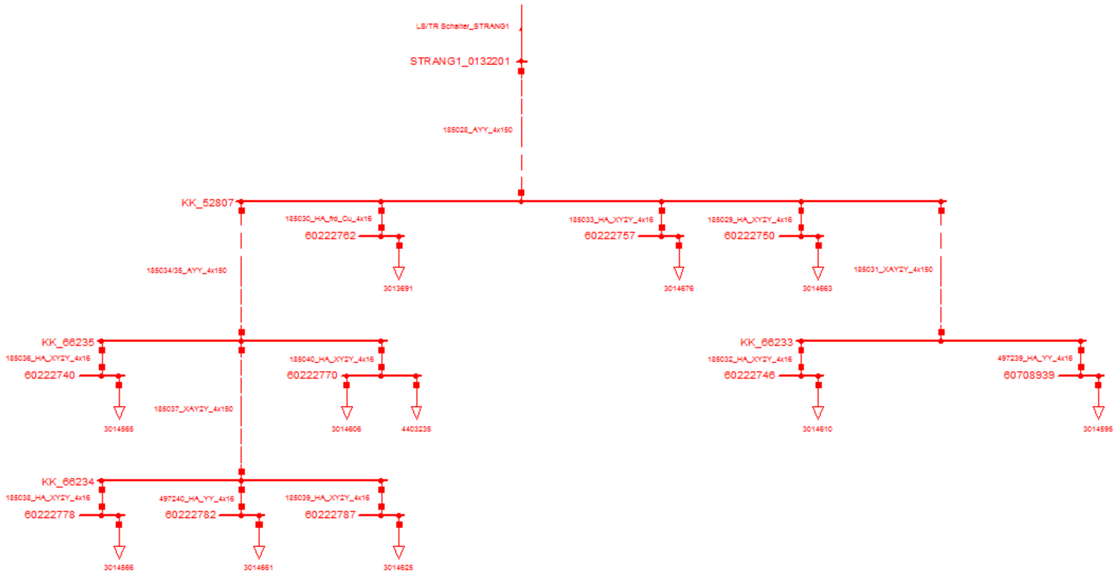
Network 22 feeder 4 (non-electrical outlier)



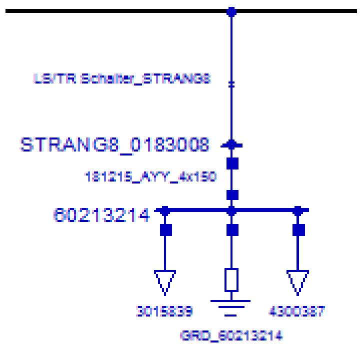
Network 21 feeder 4 (electrical outlier)



Network 23 feeder 1 (electrical outlier)

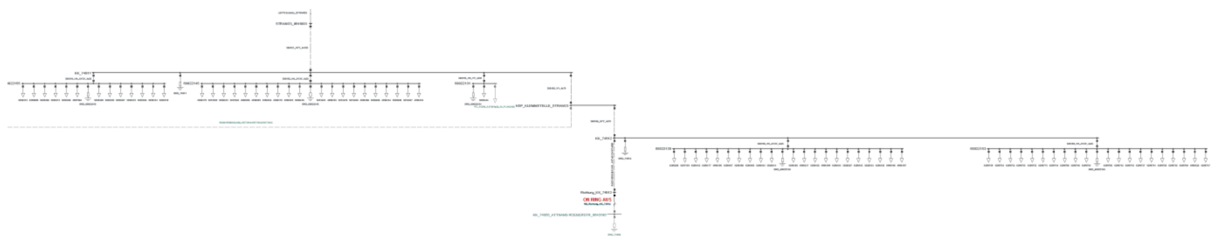


Network 24 feeder 9 (non-electrical outlier)

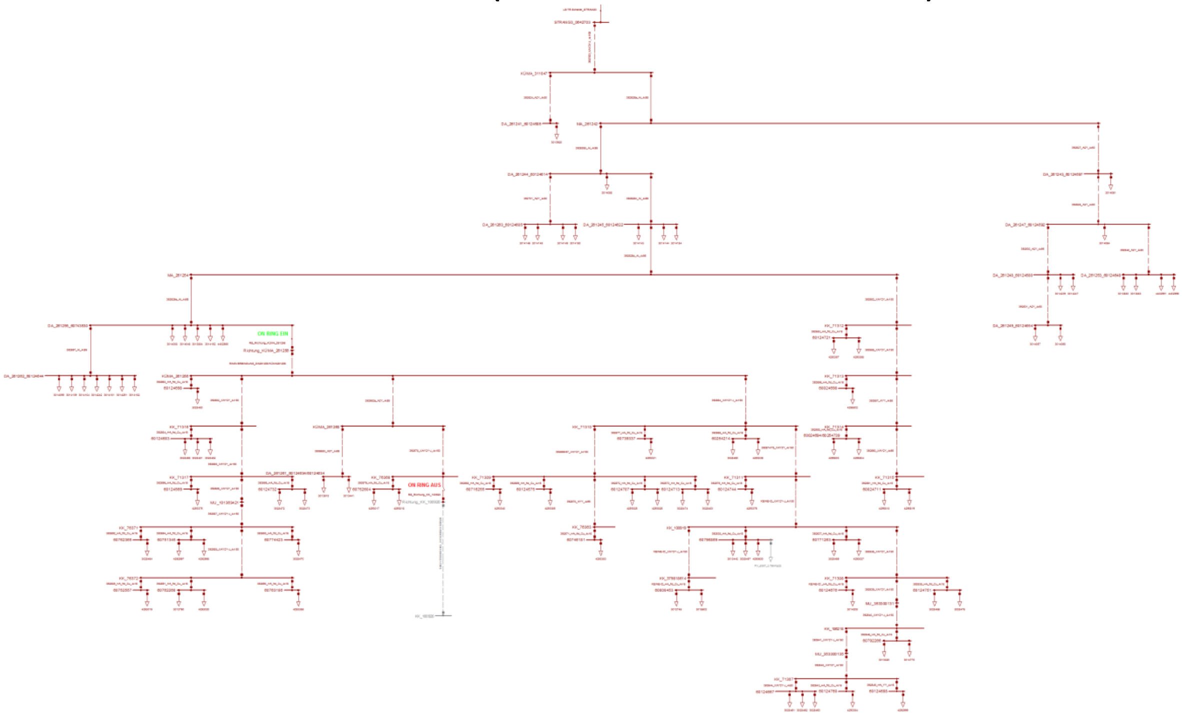


Network 26 feeder 5 (non-electrical outlier)

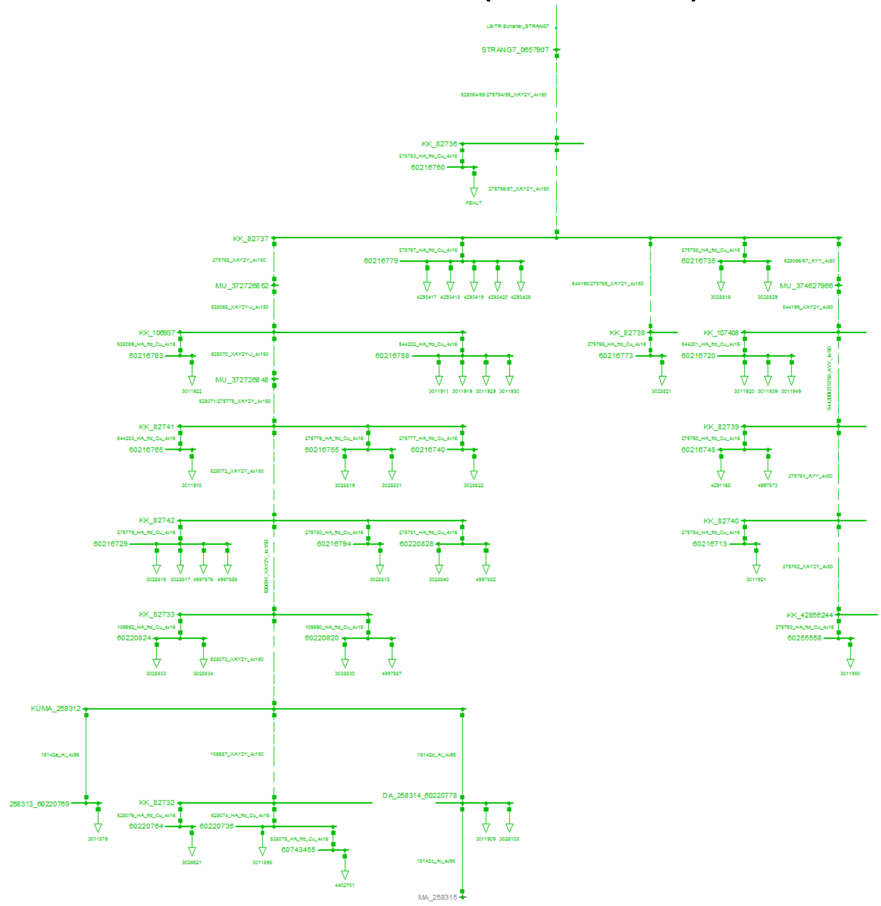
Network 25 feeder 3 (non-electrical outlier)



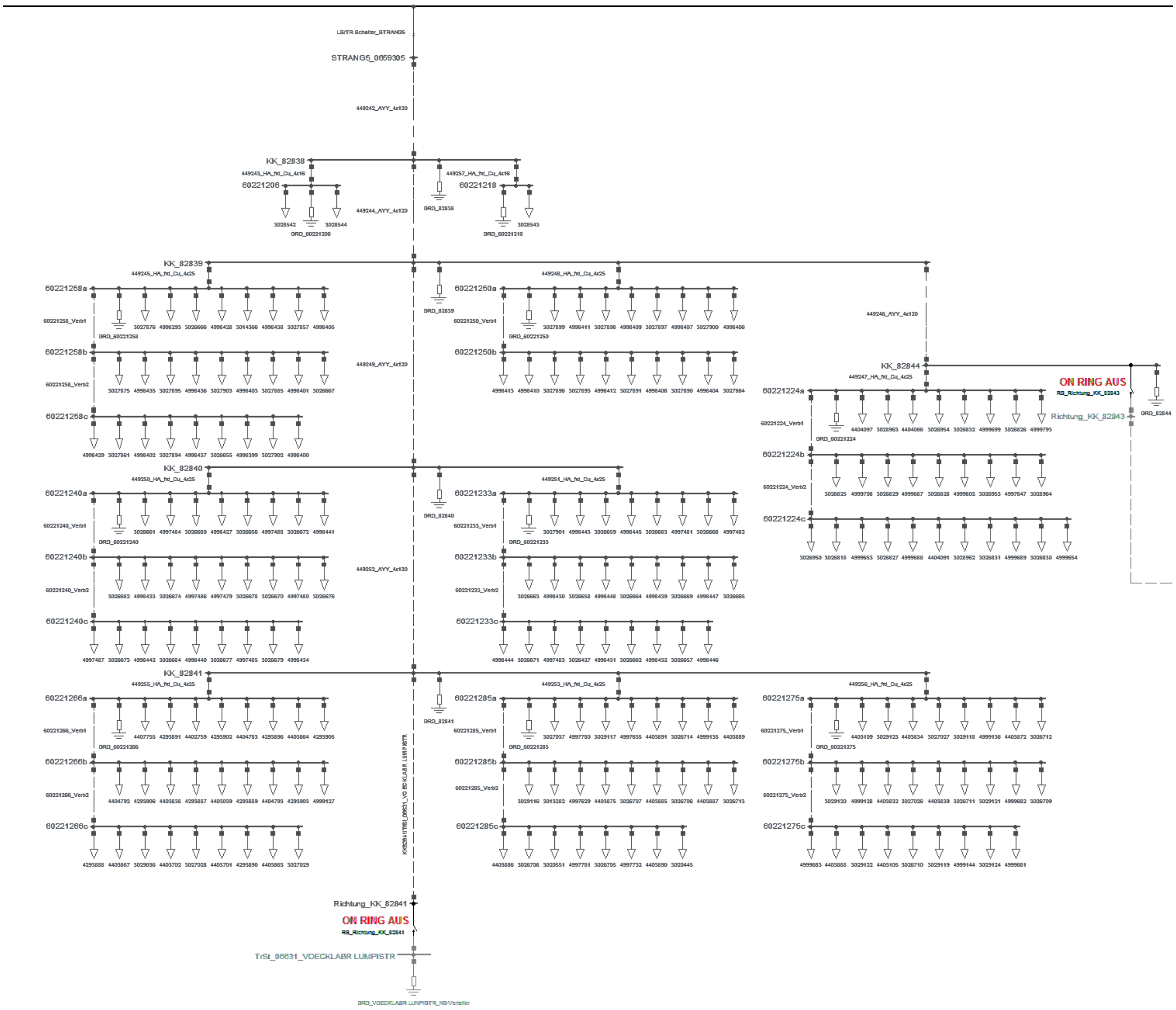
Network 26 feeder 3 (electrical and non-electrical outlier)



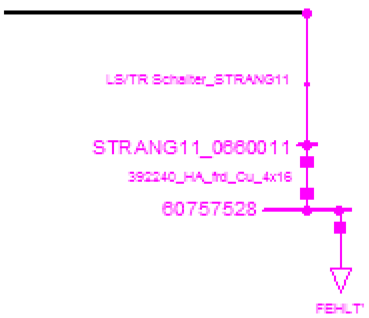
Network 28 feeder 1 (electrical outlier)



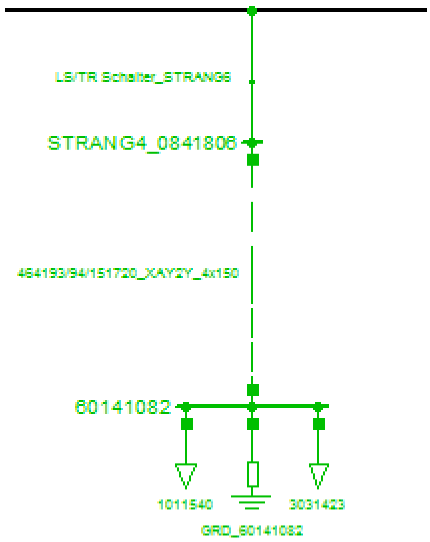
Network 30 feeder 7 (non-electrical outlier)



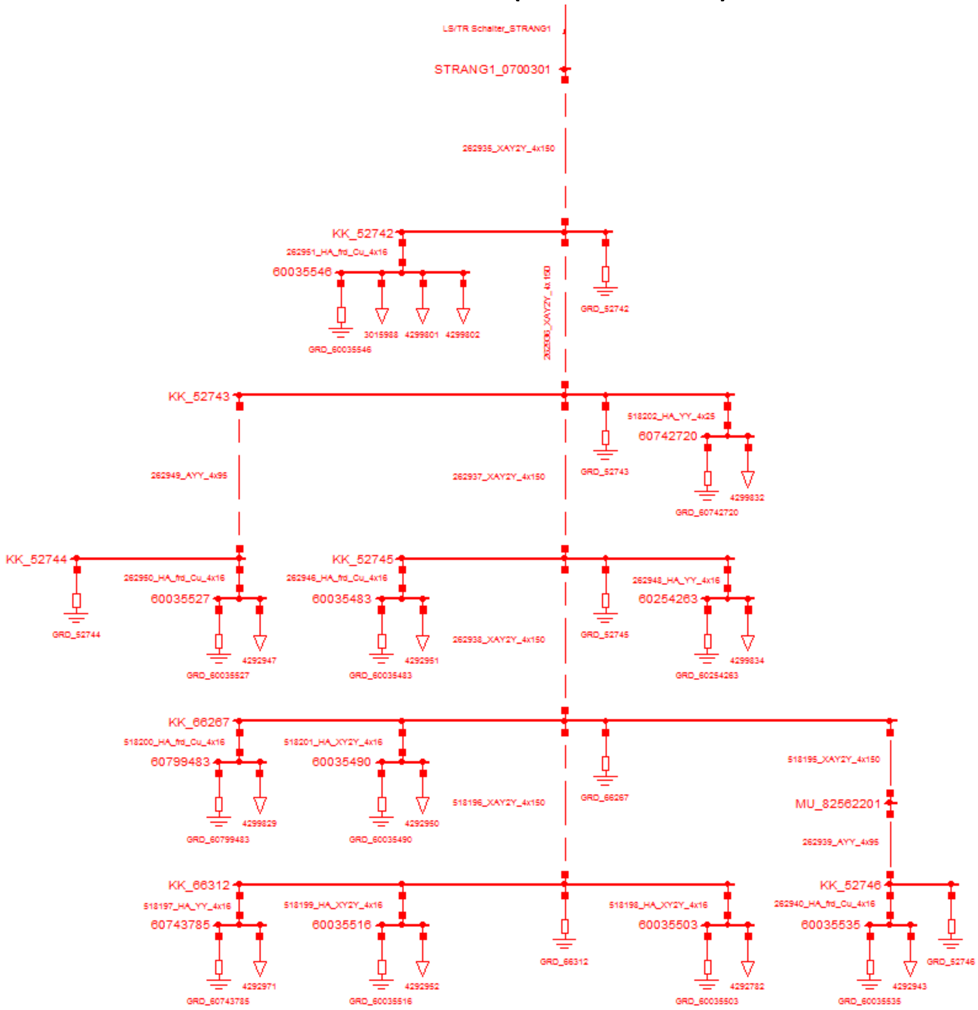
Network 31 feeder 9 (electrical outlier)



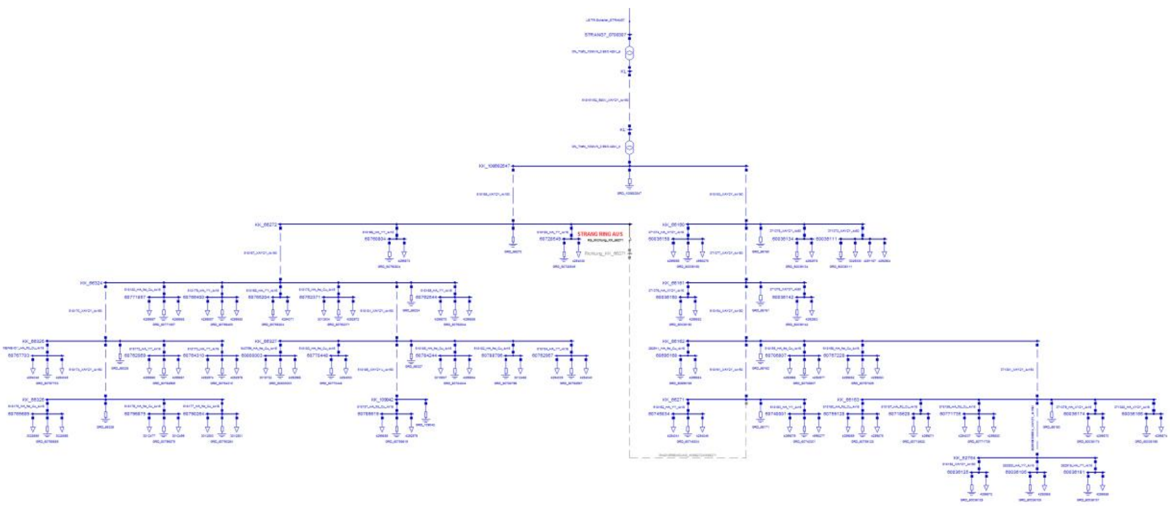
Network 34 feeder 2 (non-electrical outlier)



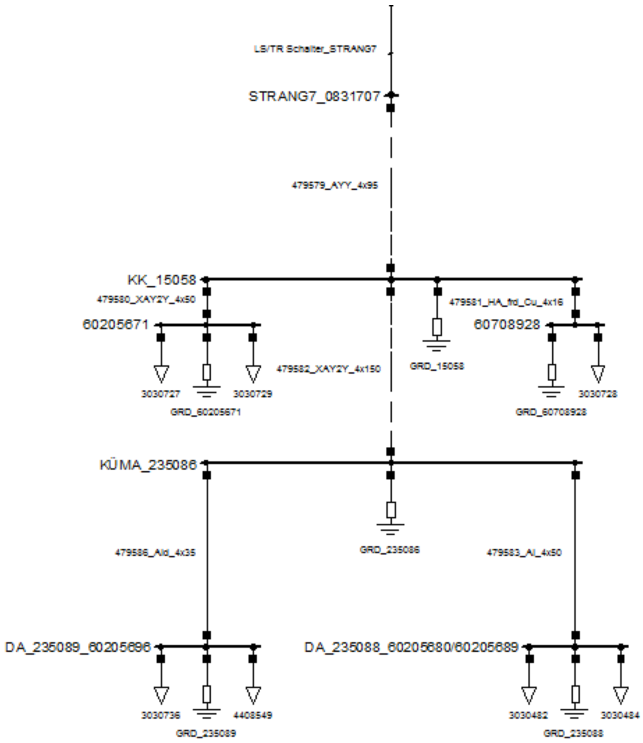
Network 32 feeder 1 (electrical outlier)



Network 32 feeder 7 (electrical and non-electrical outlier)



Network 33 feeder 8 (electrical outlier)



Network 33 feeder 9 (electrical outlier)