

Topology Inference for Radial Distribution Feeder based on Power Flow

Jie Xu (s181238)

2020-12-15

Contents

1	Introduction	5
2	Radial Distribution Feeder	7
2.1	Bus and Edge	7
2.2	Two Special Concepts	7
2.3	Case with 70 Buses	8
3	Problem Formulation	11
3.1	Directed Graph	11
3.2	Remove Overlapping Edge	12
3.3	Integer Programming	12
3.4	Local Search	13
4	AC Power Flow	15
4.1	Two Matrices	15
4.2	Bus Impedance Matrix	16
4.3	Direct Impedance Method	16
5	Linearised Power Flow	19
5.1	Bus Resistance Matrix	19
5.2	Pseudo Linearised Power Flow	21
5.3	Assessment of Candidate	22
5.4	Error from Linearisation	22
6	Result and Summary	25
6.1	Result for Case-70	25
6.2	Summary	27

Chapter 1

Introduction

Problem Setting

How households are connected to distribution network is usually unknown. There are three kinds of available information that we can rely on to infer the topology in this project.

Available information for topology inference:

For example, where households are located and barriers between them. There is no literature mentioning how to handle geographical information.

- geographical information about buses
- voltage magnitudes of all the phases of all the buses
- some real power injection profiles

In the literature, a technique called association network inference is usually used. Correlations between buses are derived based on some entity attributes. Highly correlated buses are connected directly.

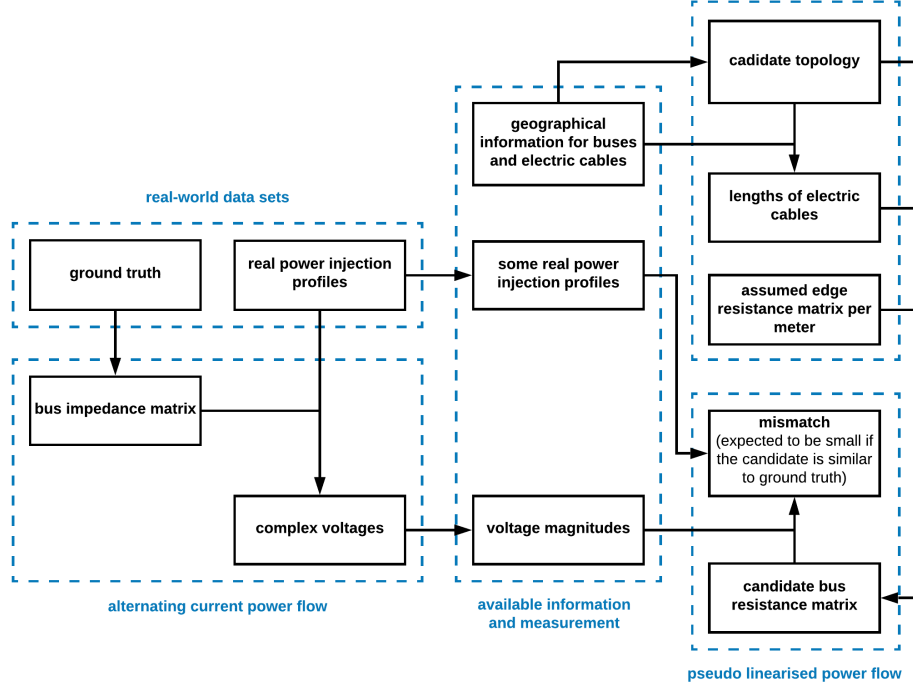
Association network inference: correlation between buses

However, there are two reasons why such technique cannot be used here.

- Spurious correlation resulted from correlated profiles.
- Missing entity attribute.

The second reason will be discussed in detail later.

Flowchart



There are three columns. In the left column, ground truth, the true topology, is known, and used to simulate available measurement using power flow program. Three boxes in the middle represents three kinds of information discussed before. Note that ground truth is not known and we want to find it. Programs in the right column are used to make inference about the topology.

Ground truth is to be found in topology inference, but used to simulate available measurement first.

Two batches of computer programs:

- power flow
- three algorithms to handle directed graphs

The second batch will not be discussed in detail.

Chapter 2

Radial Distribution Feeder

Essential concepts are discussed in following three parts.

- bus and edge, -> 2.1
- two special concepts for power flow, -> 2.2
- case with 70 buses, -> 2.3

2.1 Bus and Edge

Power grids can be roughly described here by following three concepts.

type	definition	examples
edge	transport power from one place to another	cable, transformer, capacitor
conversion element	convert power from or to another form	solar panel, battery
bus	where two edges joint or end of an edge	slack bus, PQ bus, PV bus

Only one type of edge, cable, is considered. The method applies when there are transformers, capacitors, and other devices. Usually, there is one slack bus in distribution network and is referred to as root here.

- Cable.
- One slack bus -> **root**.
- Ignore conversion elements. Not necessary in power flow calculation.

2.2 Two Special Concepts

Such two concepts cannot be found in the literature, but are essential for power flow when taking multiple phases into account.

Essential for power flow calculation.

Channel

- **channel**: refer to one phase in some bus
- **active channel**: connect to household
- **observed active channel**: power is measured

It is assumed that all inactive channels are observed.

That is, it is known that there is no power injection at those channels.

Snapshot

Snapshot: include power injections and voltages at one time index

- duration: 1 s

Zero-load snapshot : when power injections at all the channels are zero

Such two symbols will be used later.

- \bar{V}_{zero} : voltages in zero-load snapshot
- V_{rate} : rated voltage magnitude, 230 V

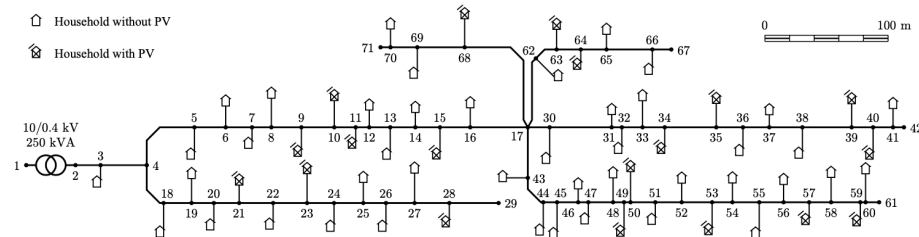
2.3 Case with 70 Buses

To make the project manageable:

Assumptions about feeders:

- spanning arborescence (SA)
- one step-down transformer
- rated voltage, 230 V
- three-phase four-wire cable
- one phase star connection

A case with 70 buses is primarily used here:



- located in Belgium

- bus 1 is omitted
- 70 buses
- 62 households -> 62 active channels

Chapter 3

Problem Formulation

- information in a directed graph -> 3.1
- integer programming formulation -> 3.3
- local search heuristic algorithm -> 3.4

There is an issue here that I will discuss after a summary in the end.

- remove overlapping edge -> 3.2

3.1 Directed Graph

All the information can be stored in a complete directed graph.

complete (directed) graph for a set of buses

All edges are **potential edges**, and they must have weights. In association network inference, it is impossible when some entity attributes are missing.

- all edges are **potential edges**
- select a set of edges -> SA -> candidate
- **impossible potential edge**
- 2-D Euclidean distance as weight

feasible region

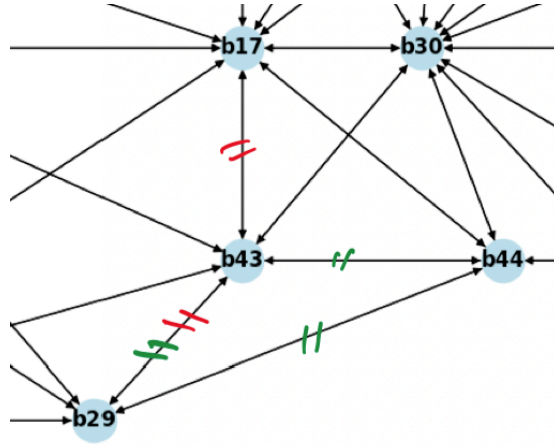
- All the candidates (SAs).
- Number of SAs is finite, making it a combinatorial optimisation problem.
- Count number of SAs.

3.2 Remove Overlapping Edge

A potential edge can overlap some bus on its way from source to target.

For example, in case-70:

shortest path	<	direct edge	\times threshold	-> remove direct edge
"b17-b43-b29"	<	"b17-b29"	$\times 1.1$	-> remove "b17-b29"
"b44-b43-b29"	>	"b44-b29"	$\times 1.1$	-> keep "b44-b29"



However:

After all overlapping edges are removed (or saying detected as impossible to exist), there are still 446 possible potential edges and over 10^{45} SAs. This is the only technique used in this project. When more information is available, more advanced techniques can be applied.

- 446 possible potential edges
- over 10^{45} SAs

-> summary

3.3 Integer Programming

Sets:

symbol	definition
\mathcal{E}	all the potential edges
\mathcal{C}	available measurements
$\mathcal{E}_{\text{impossible}}$	impossible potential edges

Variables:

symbol	definition	type	set
x_{ij}	whether to choose edge from i to j	binary	\mathcal{E}

Constants:

symbol	definition	set
$d_{i,j}$	Euclidean distance	\mathcal{E}

$$\begin{aligned}
\min_{x_{ij} \forall (i,j) \in \mathcal{E}} \quad & (1 - \alpha) \sum_{(i,j) \in \mathcal{E}} d_{ij} x_{ij} + \alpha \mathcal{H}(\{x_{ij} \forall (i,j) \in \mathcal{E}\}, \mathcal{C}) \\
\text{s.t.} \quad & \sum_{(i,j) \in \delta^-(j)} x_{ij} = 1 \quad \forall j \in V' \quad (\text{a directed forest}) \\
& \sum_{(i,j) \in \delta^-(S)} x_{ij} \geq 1 \quad \forall S \subseteq V', |S| \geq 2 \quad (\text{a connected graph}) \\
& x_{ij} = 0 \quad \forall (i,j) \in \mathcal{E}_{\text{impossible}} \quad (\text{remove impossible potential edges})
\end{aligned}$$

Two terms in the objective function:

term	definition	coefficient
$(1 - \alpha) \sum_{(i,j) \in \mathcal{E}} d_{ij} x_{ij}$	total weight of candidate	$1 - \alpha$
$\alpha \mathcal{H}(\{x_{ij} \forall (i,j) \in \mathcal{E}\}, \mathcal{C})$	assessment of candidate	α

Three sets of constraints:

Last one makes sure that impossible potential edges do not appear in any candidate.

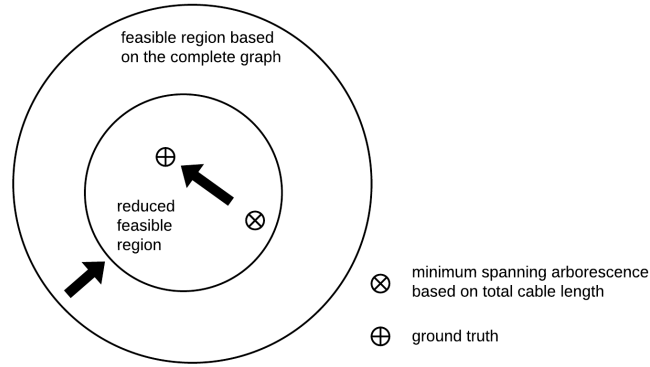
- First two sets ensure arborescence. (Fischetti and Vigo, 1997)
- Last set removes impossible potential edges.

3.4 Local Search

At least two possible values for α :

value	term lefted	to find	disadvantage
1	$\mathcal{H}(\{x_{ij} \forall (i,j) \in \mathcal{E}\}, \mathcal{C})$	ground truth	NP-hard and non-linear
0	$\sum_{(i,j) \in \mathcal{E}} d_{ij} x_{ij}$	topology with min total cable length	cannot find ground truth

Such two situations can be visualised:



For this combinatorial optimisation problem, a **local search heuristic algorithm** is proposed to move from \otimes to \oplus . (Michiels et al., 2007)

All the algorithms in this category have two parts, an objective function to assess candidates and a neighbourhood function to generate candidates systematically. Here, pseudo linearised power flow will be used as the objective function and resulted mean squared error is to be minimised. It will be discussed later. An algorithm to rank spanning arborescences according to their total cable lengths is implemented.

function	what it does	in this project
objective	assess candidate	pseudo linearised power flow
neighbourhood	generate candidate	rank spanning arborescence

Because power grids are to be built with less cost, the total cable length of ground truth is kept as short as possible, so we can find it before long.

- Ground truth should be found before long.
- Not in parallel.

Chapter 4

AC Power Flow

- two essential matrices -> 4.1
- bus impedance matrix -> 4.2
- direct impedance method for power flow calculation -> 4.2

Can be generalised for multi-phase model. (Hsieh et al., 2017)

4.1 Two Matrices

current injection -> current flow:

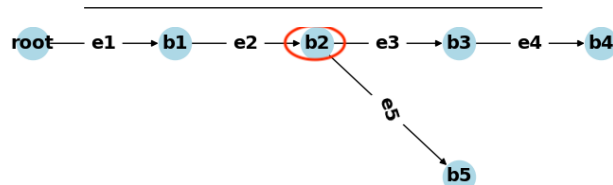
$$\bar{I}_{\text{edge}} = -K\bar{I}$$

where **edge path incidence matrix (EPI)**, K .

voltage drop -> voltage:

$$\bar{V} = \bar{V}_{\text{zero}} - K^{\top} \bar{Z}_{\text{edge}} \bar{I}_{\text{edge}}$$

where **edge impedance diagonal block matrix (EIDB)**, \bar{Z}_{edge} .



Each row indicates whether power injection from some bus will pass the edge.

$$\begin{bmatrix} \bar{I}_{\text{edge},1} \\ \bar{I}_{\text{edge},2} \\ \bar{I}_{\text{edge},3} \\ \bar{I}_{\text{edge},4} \\ \bar{I}_{\text{edge},5} \end{bmatrix} = - \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{I}_1 \\ \bar{I}_2 \\ \bar{I}_3 \\ \bar{I}_4 \\ \bar{I}_5 \end{bmatrix}$$

$$\begin{bmatrix} \bar{V}_1 \\ \bar{V}_2 \\ \bar{V}_3 \\ \bar{V}_4 \\ \bar{V}_5 \end{bmatrix} - \begin{bmatrix} \bar{V}_{\text{rate}} \\ \bar{V}_{\text{rate}} \\ \bar{V}_{\text{rate}} \\ \bar{V}_{\text{rate}} \\ \bar{V}_{\text{rate}} \end{bmatrix} = - \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^\top \begin{bmatrix} Z_{\text{edge},1} & 0 & 0 & 0 & 0 \\ 0 & Z_{\text{edge},2} & 0 & 0 & 0 \\ 0 & 0 & Z_{\text{edge},3} & 0 & 0 \\ 0 & 0 & 0 & Z_{\text{edge},4} & 0 \\ 0 & 0 & 0 & 0 & Z_{\text{edge},5} \end{bmatrix} \begin{bmatrix} \bar{I}_{\text{edge}} \\ \bar{I}_{\text{edge}} \\ \bar{I}_{\text{edge}} \\ \bar{I}_{\text{edge}} \\ \bar{I}_{\text{edge}} \end{bmatrix}$$

Alternating current power flow: (Conti et al., 2006)

$$\bar{V} = \bar{V}_{\text{zero}} + (K^\top \bar{Z}_{\text{edge}} K) \bar{I}$$

4.2 Bus Impedance Matrix

Alternating current power flow:

$$\bar{V} = \bar{V}_{\text{zero}} + (K^\top \bar{Z}_{\text{edge}} K) \bar{I}$$

Bus impedance matrix (BIM), \bar{Z} , is defined as:

$$\begin{aligned} \bar{Z} &= K^\top \bar{Z}_{\text{edge}} K \\ &= R + jX \end{aligned}$$

where **bus resistance matrix (BRM)**, R : real part of entries in BIM.

4.3 Direct Impedance Method

The whole process is called direct impedance method. BIM is built by following five steps and used to calculate power flow.

Five steps to build BIM:

1. Define a unit impedance matrix.
2. Calculate edge impedance matrices for cables.
3. Build EIDB.
4. Obtain EPI based on topology.
5. Calculate BIM using EIDB and EPI.

Fixed Point Method

The following procedure is repeated:

$$\begin{aligned}\bar{I} &= \underline{P} \otimes \underline{V}_{\text{previous}} \\ \bar{V} &= \bar{Z}\bar{I} + \bar{V}_{\text{zero}} \\ \epsilon &= (\bar{V} - \bar{V})^\top (\bar{V} - \bar{V})\end{aligned}$$

until ϵ is smaller than a pre-defined threshold.

Chapter 5

Linearised Power Flow

- assessment of candidate -> 5.3

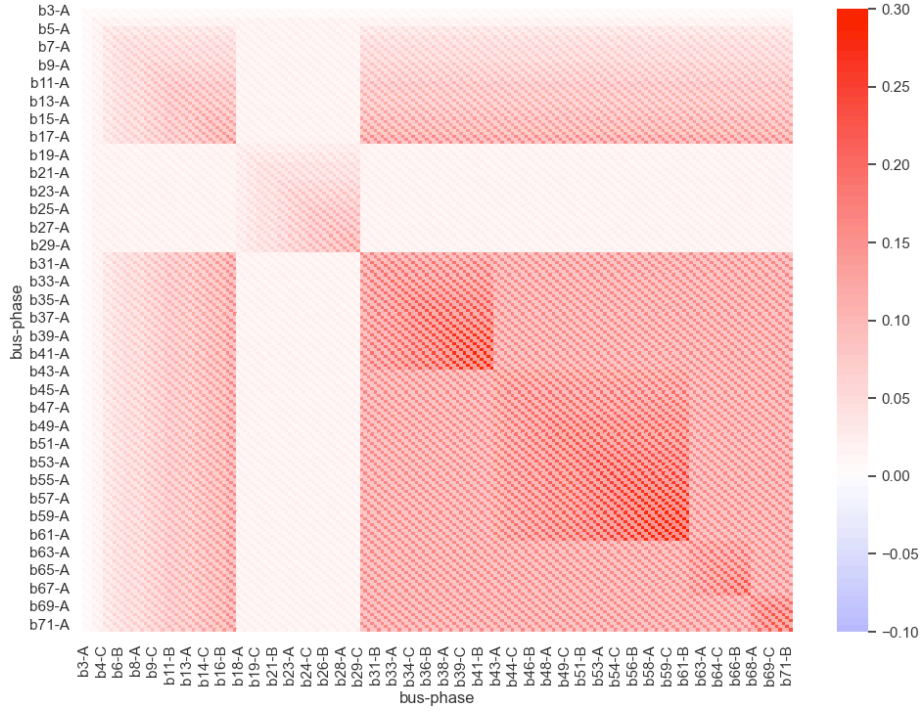
Besides, there are three parts that I will discussed in detail after the summary.

- bus resistance matrix -> 5.1
- inversed bus resistance matrix -> 5.2
- error from linearisation -> 5.4

5.1 Bus Resistance Matrix

BRM of case-70:

- bus 2 -> root
- 69 PQ buses
- 207 channels
- 207 rows and 207 columns



There is some pattern that can be explained by the analytical expression.

Lowest Common Ancestor Problem

Entry (i, j) \rightarrow sum of edge resistances in the path from root to their lowest common ancestor (LCA):

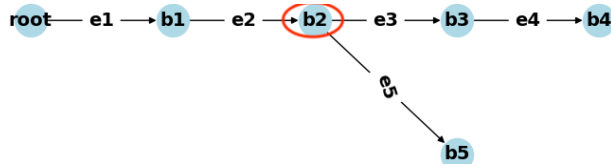
$$R_{i,j} = \sum_{k \in U_i \cap U_j} R_{\text{edge},k}$$

where U_i is set of edges on the path from root to bus i .

BRM can be calculated efficiently using LCA for all pairs of buses. and the pattern can be used in future work.

- Calculated efficiently using LCA for all pairs.
- Useful pattern.

For example,



pair of buses	entry in BRM
b3-b5	$R_{e1} + R_{e2}$
b4-b5	$R_{e1} + R_{e2}$

-> summary

5.2 Pseudo Linearised Power Flow

Based on linearised power flow, $V = V_{\text{zero}} + \frac{1}{V_{\text{rate}}} RP$:

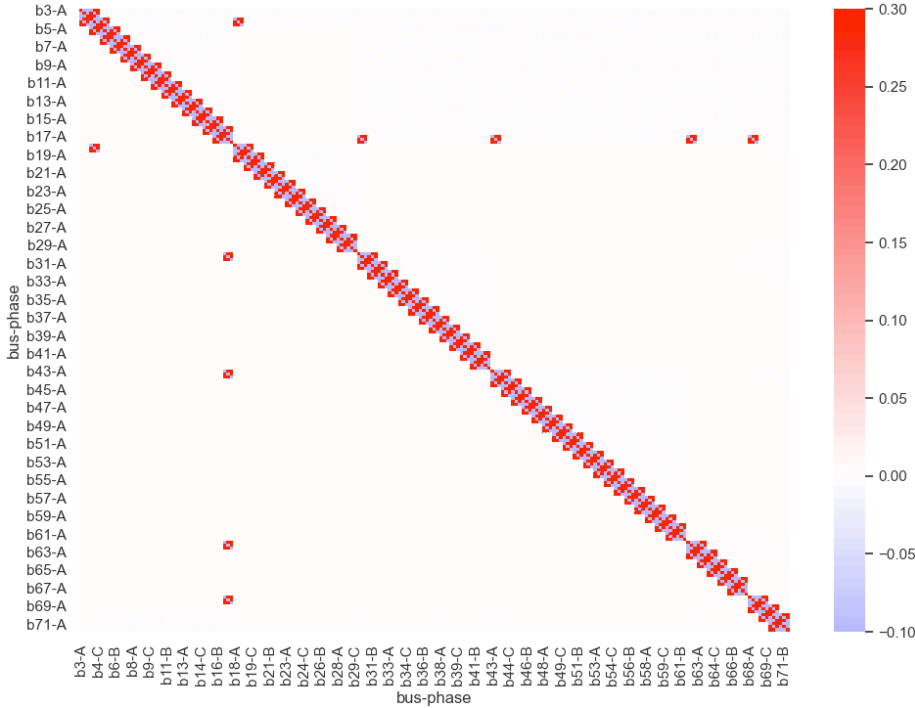
$$P_{\text{assess}} = V_{\text{rate}} R^{\top} (V - V_{\text{zero}})$$

-> **pseudo linearised power flow.**

which is referred to as pseudo linearised power flow.

Inversed BRM for case-70:

Inversed BRM for the case looks like:



- Sparse.
- Full rank.
- Voltage magnitude at any channel can have a huge impact.
- Useful pattern.

5.3 Assessment of Candidate

Linearised power flow:

$$\begin{aligned} V &= V_{\text{zero}} + \frac{1}{V_{\text{rate}}} (K^\top R_{\text{edge}} K) P \\ &= V_{\text{zero}} + \frac{1}{V_{\text{rate}}} RP \end{aligned}$$

- Calculate P_{assess} using voltage magnitudes.
- Compare with available measurements.

Mean squared error (MSE):

$$\mathcal{H}(R) = [(P_{\text{assess}} - P_{\text{measure}}) \otimes O]^\top \cdot [(P_{\text{assess}} - P_{\text{measure}}) \otimes O] / |\mathcal{O}|$$

where:

It is the second term in the objective function. Entries for unobserved active channels are ignored.

- \mathcal{O} : set of observed active channels and inactive channels.
- O : binary vector indicating observed active channels.

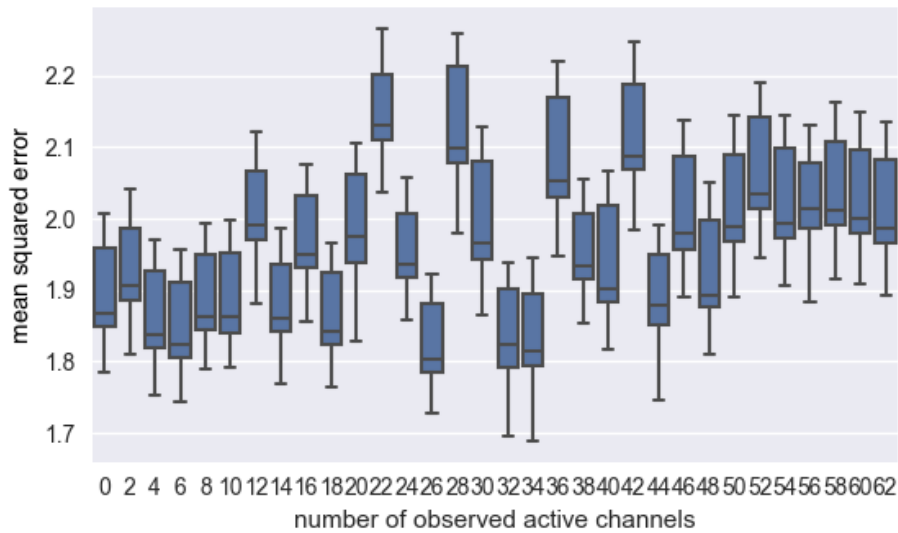
5.4 Error from Linearisation

This is a box plot for assessment of ground truth. Each box contains 50 MSE values for a given number of observed active channels.

Box plot:

- with respect to different number of observed active channels
- based on ground truth and 50 snapshots¹

¹during 00:00:00 and 00:00:50 on Dec 2, 2020 from Sonnen data set.



Error is not 0, which can also be seen in the result before. The effect of number of observed active channels is insignificant. Rated voltage magnitudes will increase the error dramatically, so full observability of voltage magnitudes are still required for now.

- ~~Number of observed active channels.~~
- Rated voltage magnitudes will increase the error dramatically.

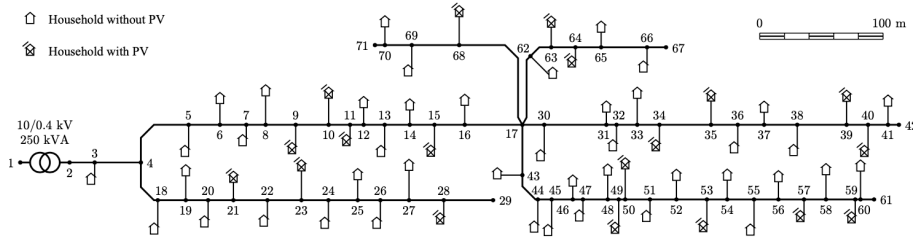
-> summary

Chapter 6

Result and Summary

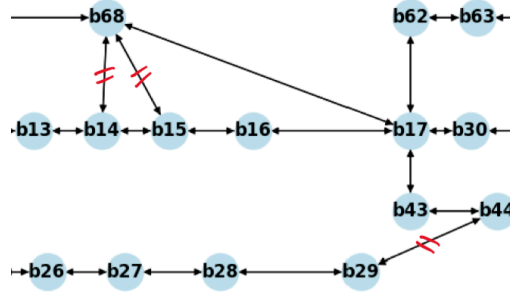
- result for case-70 -> 6.1
- summary -> 6.2

6.1 Result for Case-70



Here we only show how result might look like. Because there is an issue that will be discussed later. Instead, 75 possible potential edges are manually added besides 69 edges in ground truth here. So there are 144 edges in total and 288 SAs rooted at bus 2. There are three new pairs, which are highlighted. Any connected pair has two edges with opposite directions. Full observability over power injections is assumed, and we will talk about it later.

- three new pairs, “b29-b44”, “b68-b14”, and “b68-b15”
- 144 potential edges in total
- 288 SAs rooted at bus 2
- full observability



Rank candidates according to total cable lengths:



It increases all the time.

Assessment based on 50 snapshots¹:

¹during 00:00:00 and 00:00:50 on Dec 2, 2020 from Sonnen data set.



The red cross point, with minimum mean squared error, corresponds to ground truth.

6.2 Summary

The project is summarised here. We turn topology inference into a combinatorial optimisation problem. A new framework is proposed, and its core is a local search heuristic algorithm.

- Topology inference -> combinatorial optimisation problem.
- A new framework is proposed.
- Core: local search heuristic algorithm.

Four steps:

1. Shrink feasible region (reduce number of SAs).
2. Measure the size of feasible region.
3. Get candidates sequentially according to total cable lengths.
4. Assess candidates based on available measurements.

Advantages:

There are two advantages. It is robust to partial observability over power injections, and is capable to integrate all kinds of information in weights and directions.

- Robust to partial observability.
- Integrate all kinds of information in weights and directions.

Issues

There are three issues that I want to discuss in detail. Firstly, there are too many candidates, so we have to reduce the number, for example, by removing overlapping edges.

1. Too many candidates. (remove overlapping edges)
2. Full observability over voltage magnitudes. (matrices with full rank)
3. Error in linearised power flow calculation. (error from linearisation)

Future Work

- How to detect more impossible potential edges. (for issue 1)
- How to assess candidates based on a fraction. (for issue 2)
- How to use voltage sensitivity matrix in linearised power flow. (for issue 3)

Bibliography

- Conti, S., Greco, A., and Raiti, S. (2006). Voltage sensitivity analysis in mv distribution networks. In *Proceedings of the 6th WSEAS/IASME International Conference on Electric Power Systems, High Voltages, Electric Machines, Tenerife, Spain*, pages 16–18.
- Fischetti, M. and Vigo, D. (1997). A branch-and-cut algorithm for the resource-constrained minimum-weight arborescence problem. *Networks: An International Journal*, 29(1):55–67.
- Hsieh, T.-Y., Chen, T.-H., and Yang, N.-C. (2017). Matrix decompositions-based approach to z-bus matrix building process for radial distribution systems. *International Journal of Electrical Power & Energy Systems*, 89:62–68.
- Michiels, W., Aarts, E., and Korst, J. (2007). *Theoretical aspects of local search*. Springer Science & Business Media.