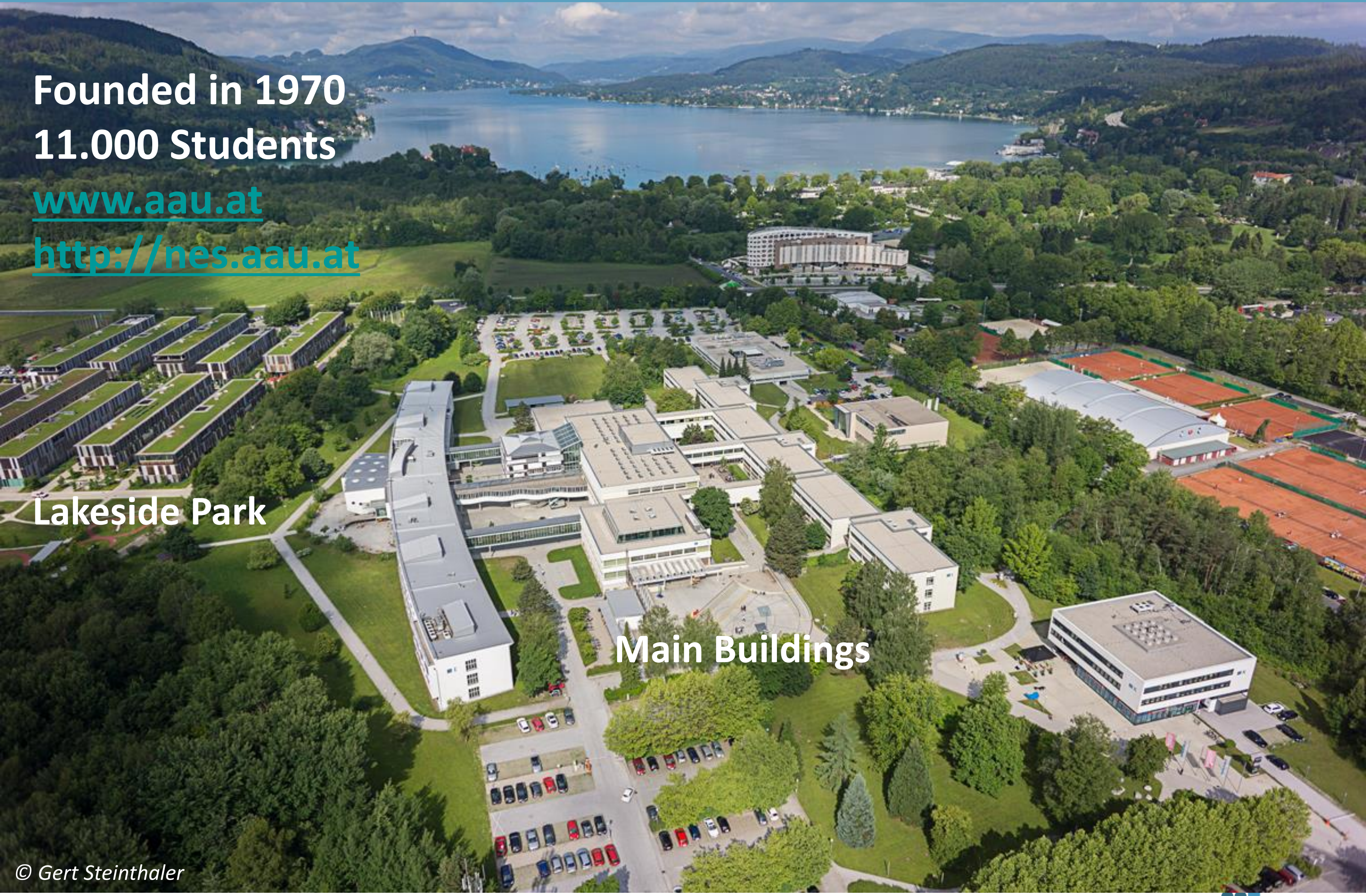
Title:	**Final project results**
Project:	**Assessment of G3-PLC technology improvements**
Contributors:	**Francesco Marcuzzi, Nunzio Letizia, Andrea Tonello**
Contact:	**Prof. Dr. Andrea Tonello** andrea.tonello@aau.at +43 463 2700 3661
Date:	**28.01.2020**
Status:	**Confidential**

Universität Klagenfurt



Chair of embedded communication systems

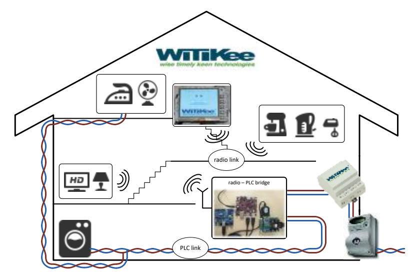
Communications
Signal processing
Electronics
Measurements
Embedded systems

Smart cars

Smart grids

IoT

Aerial robotics



Outline

- Project objectives
- Results
- Next steps

Project objectives

General aims

- Foster innovation and research collaboration in the PLC field
- Discuss research areas to advance the G3-PLC technology, its applications and use cases
- Raise awareness of said technology through dissemination activities, including eventual publications
- Identify areas of possible improvement and initiate activities to develop new or updated specifications for new use cases

Specific goals in Phase 1

- Assessment of datasets to derive performance statistics
- Study routing mechanism from the data sets and analyze dependence from network layout
- Identify potential research paths offered by data analytics
- Define deeper research actions for a second phase project (Phase 2)

Main steps and developments

❑ Project steps

- Several ideas and exploration areas have been discussed, mostly in the domain of routing and analysis of G3-PLC performance
- Collected two data sets from real deployments by ENEDIS
- **Extensive work on definition of relevant data organization and semantics**
- Decided to focus on **a methodology to extract relevant data and reconstruct topology:**
  - ➢ communication/logical topology, electrical topology, and geo-topology
- Then, we analyzed provided data sets

Overall results

Overall results

- Defined dataset semantics
- Effective extraction and visualization of electrical topology with embedded routing information
- Methodology to detect anomalies in data set
- Methodology to detect anomalies in routing
- Studied correlations between routing path and electrical topology. Some past evidence and ideas offered in [*Refs*]

Ref. A. M. Tonello, N. A. Letizia, D. Righini, F. Marcuzzi, "**Machine Learning Tips and Tricks for Power Line Communications**," IEEE ACCESS, 2019.

Ref. F. Marcuzzi, A. M. Tonello, "**Artificial-Intelligence-Based Performance Enhancement of the G3-PLC LOADng Routing Protocol for Sensor Networks**," IEEE ISPLC 2019

Ref. F. Marcuzzi, A.M. Tonello, "**Synthetic Training of a Machine Learning Approach for Routing in PLC Networks**," proc. of WSPLC 2018.

Ref. F. Marcuzzi, A.M. Tonello, "**Artifical Intelligence Based Algorithm for Routing in PLC**," IEEE ISPLC 2018.

Topology Reconstruction

Research question

Research question

- Is it possible from the datasets and through a portable algorithm to reconstruct the electrical topology of the network – including also routing information?

What topology?

- Logical topology
- Electrical topology
- Geographical topology

Data sets

- To answer the question, we have processed the following data sets:
  - **Link table:** it contains "To"/"From" pairs of logical network elements

	A	B
1	from	to
2	92050R0405	D_9205070801
3	92050R0405	D_9205070802
4	92050R0405	D_9205070803
5	92050R0405	D_9205070804
6	92050R0405	D_9205070809
7	92050R0405	D_9205070810

  - **Network table:** it contains information about physical, topology and management configuration of all meters in the network

	S	T	U	V	W	X	Y	Z	AA	AB
1	id_ligne	id_raccorc	nb_lambert_x_rac	nb_lambert_y_rac	id_troncon	nb_longueur_tron	lb_composition_cable	cd_type_troncon	id_troncon_pere	id_rang
2	92050R040(	46106807€	590453.694	2431586.545	9205070844	34	3 x 150 AL + 70 AL	S	9205070842	3
3	92050R040(	461068007	590445.002	2431602.756	9205070842	35	3 x 240 AL + 95 AL	S	9205070841	2
4	92050R040(	539332163	590444.707	2431603.187	9205070842	35	3 x 240 AL + 95 AL	S	9205070841	2
5	92050R040(	461068141	590413.639	2431611.715	9205070863	49	3 x 150 AL + 70 AL	S	9205091911	3
6	92050R040(	461068141	590413.639	2431611.715	9205070863	49	3 x 150 AL + 70 AL	S	9205091911	3
7	92050R040(	461068141	590413.639	2431611.715	9205070863	49	3 x 150 AL + 70 AL	S	9205091911	3

Data sets (cont'd)

- **Correspondence table:** correspondence between different identifiers (*prm*, *compteur* and *short mac*)

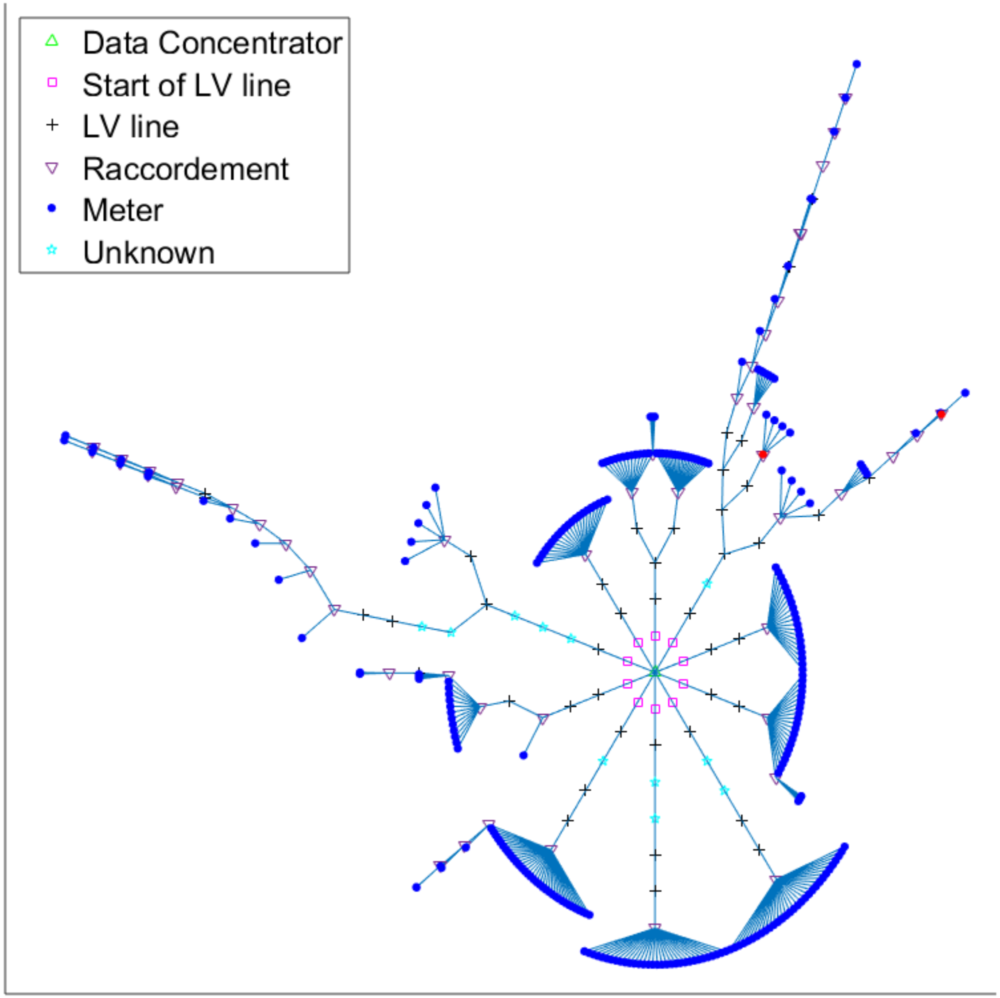
	A	B	C
1	id_prm	id_compteur	id_mac_etendue
2	21424891403336	#N/A	#N/A
3	21419247409932	031864451028	000781FFFE338C22
4	21418089667584	041864212203	000F93FFFE9967C0
5	21443849465894	021875790988	B2982BFFFEE05E80
6	21444717772630	041864211998	000F93FFFE98FA25
7	21442112852273	041864212000	000F93FFFE99C209

- **Log files (*3800 !*):** information bases (*routing set, neighbour table, LQI*) extracted from G3-PLC routers (using path request primitives)

	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	IdC	IdDem	NomOpe	HExec	NumOrd	CR	Code	Nom	HDate	destination	nextHop	coutRoute	coutHop	weakLink	validite
2	31861017238	-2.00001E+17	LectPanIdAlternatif	2019-07-04T14:03:32.595Z		OK	LU_INF007								
3	31861017238	-2.00001E+17	LectTableRoutage	2019-07-04T14:03:37.487Z	1	OK		TableRoutage	2019-07-04T14:03:37.486Z	0	0	10	1	0	1440
4	31861017238	-2.00001E+17	LectTableRoutage	2019-07-04T14:03:37.487Z	1	OK		TableRoutage	2019-07-04T14:03:37.486Z	253	253	10	1	0	526
5	31861017238	-2.00001E+17	LectTableRoutage	2019-07-04T14:03:37.487Z	1	OK		TableRoutage	2019-07-04T14:03:37.486Z	6	6	9	1	0	990
6	31861017238	-2.00001E+17	LectTableRoutage	2019-07-04T14:03:37.487Z	1	OK		TableRoutage	2019-07-04T14:03:37.486Z	4	4	9	1	0	990
7	31861017238	-2.00001E+17	LectTableRoutage	2019-07-04T14:03:37.487Z	1	OK		TableRoutage	2019-07-04T14:03:37.486Z	248	248	10	1	0	520

Link topology

Logical topology reconstruction: incorrect Link-table dataset



- PV COUTURIER: LV network in France
- 10 buses/sectors
- 320 meters (end nodes)

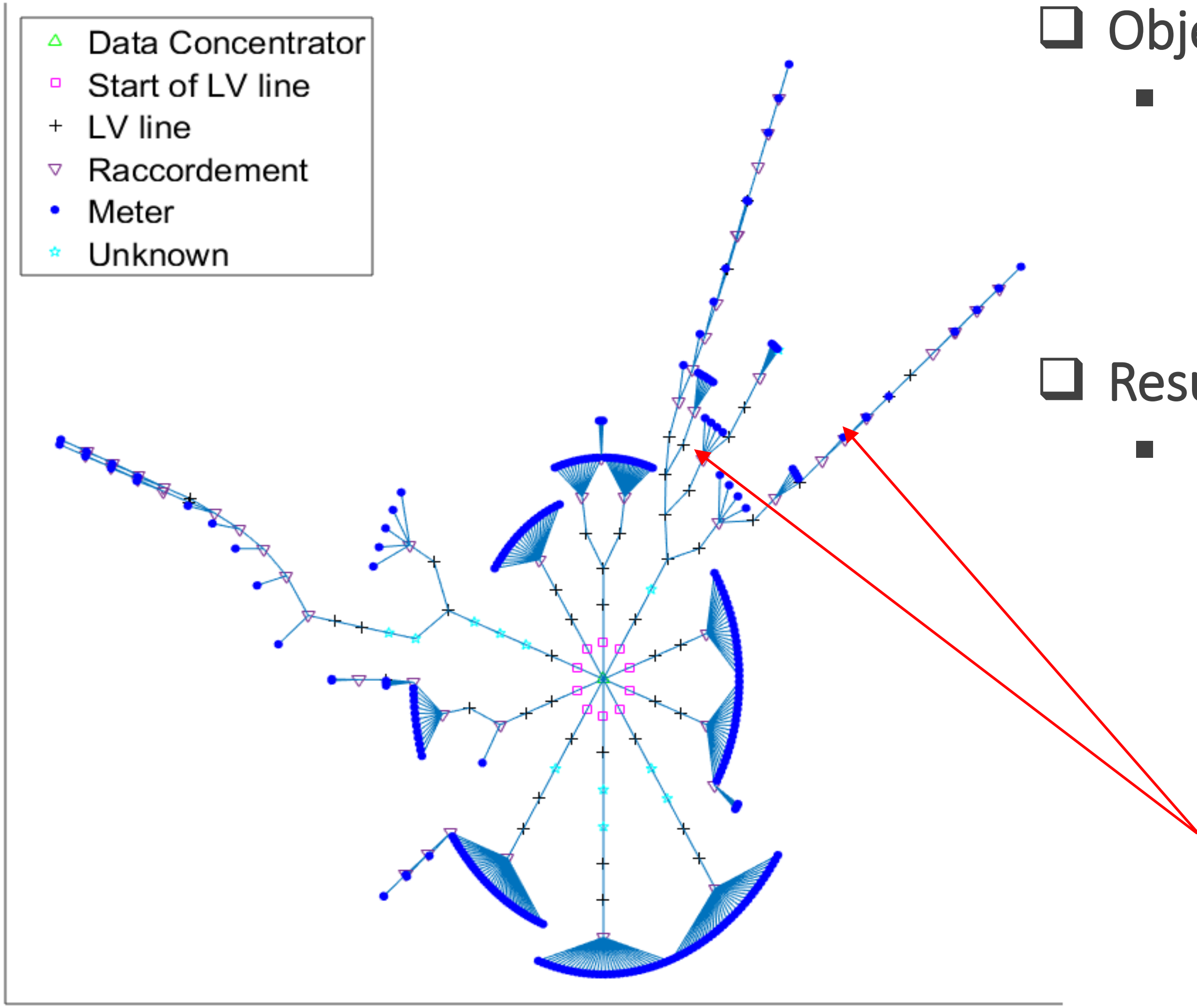
Challenge:
- Reconstruct logical topology and easily visualize it

Starting point:
- Use only the link table
- Hypothesis: radial structure

Result:
- Anomaly - not all elements from the link table fit at the beginning the reconstruction: *data set was incomplete*

UNIVERSITÄT KLAGENFURT

Logical topology reconstruction: correct Link-table dataset



❑ Objective:

- Develop a fast reconstruction algorithm and fix the data set

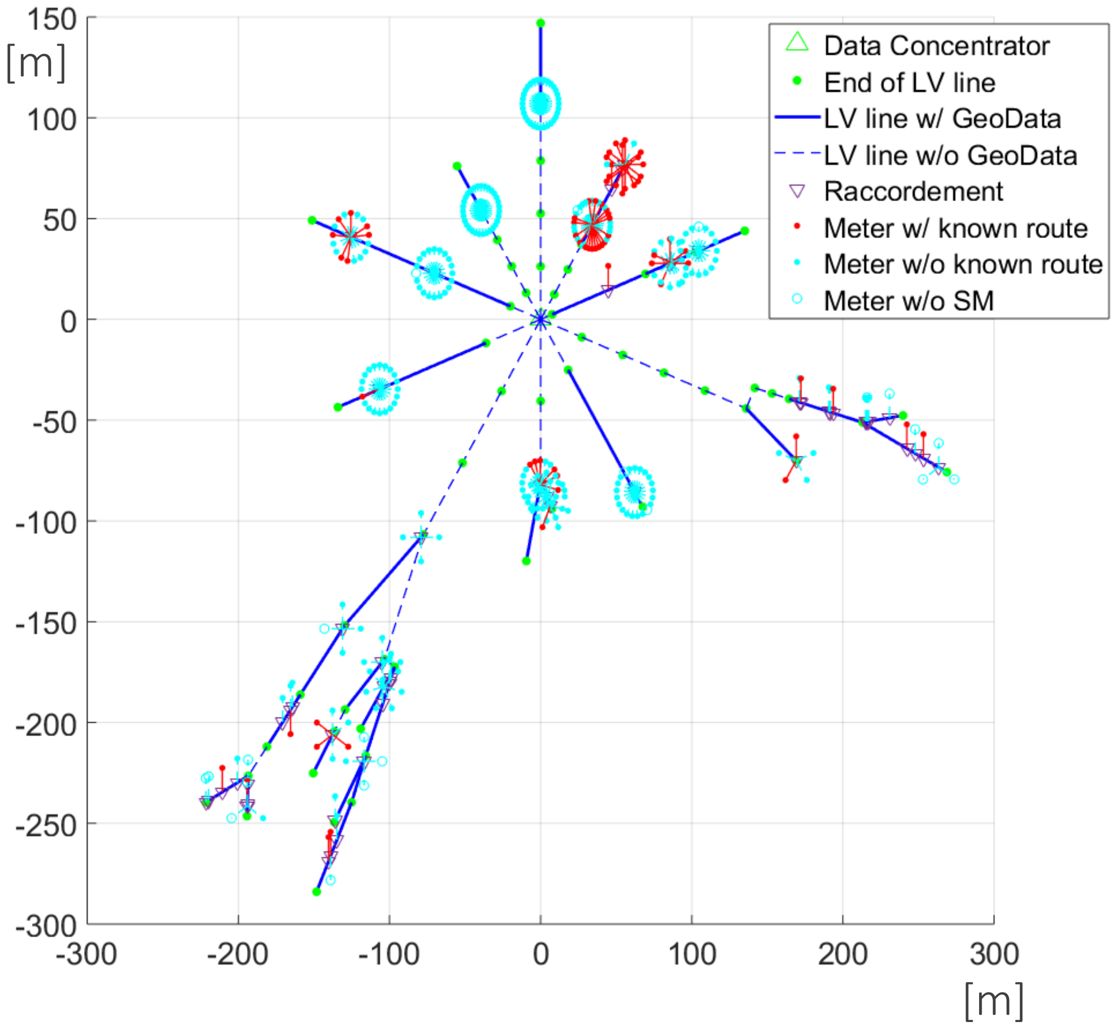
❑ Result:

- No dangling links, all elements from the link table are used for the reconstruction (no anomaly detected)

New links where the incomplete dataset expected some links but there weren't

Electrical Topology with Spatial Information (Geo-topology)

Electrical topology reconstruction



- **Challenge:**
  - Reconstruct electrical topology
- **Starting point:**
  - Geometric data introduces spatial information w.r.t. link topology
  - Distances are real
  - Use all provided datasets
  - Every red node has at least one entry in its Routing Set to the data concentrator
- **Result:**
  - Portable algorithm
  - Clear visualization and labeling of components
  - Geometric properties (cable lengths) preserved
  - Lots of incomplete info

Data Semantics

Guidelines for data semantics

- The algorithm for visualization works with databases with specific properties
- The analysis of the provided dataset, enabled the definition of data structure and semantics for easy reconstruction

Important data

Topological info (table)	node ID, position, link table as in original dataset or adjacency matrix to identify direct connections
LV lines (struct)	length, position of vertices, number and ID of „raccordements“, previous LV segment and subsequent one(s), type of cable
Linking points (struct)	referred to as „raccordements“ also, are LV lines end points reaching a private premise. Need to specify ID, LV line they pertain to, number of inside nodes
Meters/end nodes (struct)	Specify ID, related linking point, distance from the linking point, any routing information

Topological info

- Link table provided with original dataset

	A	B
1	from	to
2	92050R0405	D_9205070801
3	92050R0405	D_9205070802
4	92050R0405	D_9205070803
5	92050R0405	D_9205070804
6	92050R0405	D_9205070809
7	92050R0405	D_9205070810

or ...

- Adjacency matrix also contain all the information

	A	B	C	D	E	F
1		Node ID 1	Node ID 2	Node ID 3	...	
2	Node ID 1	0	1	0	0	
3	Node ID 2	1	0	0	1	
4	Node ID 3	0	0	0	0	
5	...	0	1	0	0	
6						

LV lines - struct

- Struct LV lines: full description of the power line infrastructure

Fields	ID	ID_prev	ID_next	ind_next	empty	raccs	rank	coords	direction	amplitude	DSFC	sector	len
1	"9205070801"	"D_9205070...	"9205070841"	11	1	[]	0	[0,-7.6085,0,-2.4721]	3.4558	0.6283	0	1	8
2	"9205070802"	"D_9205070...	"9205091755"	12	1	[]	0	[0,-9.0127,0,-12.4049]	4.0841	0.6283	0	2	15.3333
3	"9205070803"	"D_9205070...	"9205091757"	13	1	[]	0	[0,-4.8220e-15,0,-26.2500]	4.7124	0.6283	0	3	26.2500
4	"9205070804"	"D_9205070...	"9205091758"	14	1	[]	0	[0,9.5515,0,-13.1465]	5.3407	0.6283	0	4	16.2500
5	"9205070809"	"D_9205070...	"9205001260"	15	1	[]	0	[0,19.9722,0,-6.4894]	5.9690	0.6283	0	5	21
6	"9205070810"	"D_9205070...	"9205001273"	16	1	[]	0	[0,36.1401,0,11.7426]	6.5973	0.6283	0	6	38
7	"9205070811"	"D_9205070...	"9205071041"	17	1	[]	0	[0,25.8626,0,35.5967]	7.2257	0.6283	0	7	44

- ID is the unique identifier of the LV line; prev and next refer to the LV lines respectively toward the DC or further away
- ind_next: index in the same struct of the elements in ID_next
- empty: boolean flag. Is there any raccordement on this line?
- raccs: list of raccordements pertaining to this LV line
- rank: logical distance from DC
- coords: position of vertices on the graph [x_1, x_2, y_1, y_2]
- direction and amplitude: angle values for representation on the graph
- DSFC and len: distance of the start of the line from DC and length of the line
- sector: network is divided in sector, based on number of buses (see slide 13)

Linking points - struct

- Struct Linking points: full description of the branching points, where end nodes are connected to the main infrastructure

Fields	ID	LVline	NNodes	IDNodes	ShortIDNodes	DFSL	DFC	sector	pos	LVlinesChain
1	"461068002"	"9205070841"	1	"214507959...	"281"	39	47	1	[-44.6997,-1...	[1,11]
2	"461081493"	"9205001260"	20	*20x1 string*	*1x20 string*	53	74	5	[70.3782,-22...	[5,15]
3	"461081643"	"9205001273"	20	*20x1 string*	*1x20 string*	74	112	6	[106.5183,3...	[6,16]
4	"461082073"	"9205001274"	21	*21x1 string*	*1x23 string*	75	106	9	[-62.3052,85...	[9,19]
5	"461081574"	"9205001260"	17	*17x1 string*	*1x17 string*	111	132	5	[125.5395,-4...	[5,15]
6	"461081910"	"9205001266"	21	*21x1 string*	*1x21 string*	1	82	8	[0.2359,81.9...	[8,18,25]
7	"461081995"	"9205001275"	12	*12x1 string*	*1x13 string*	8	89	8	[-3.9475,87....	[8,18,27]

- ID is the unique identifier of the Linking point
- It also refers to the LV line (unique) to which the Linking point is attached
- NNodes and IDNodes are respectively the number of nodes connected to this Linking point, and their ID
- ShortIDNodes contains the short MAC address used by the routing protocol
- DFSL and DFC are distances from, respectively, the start of the LV line and the DC
- sector: as before
- pos: coordinates [x,y] of the Linking point on the graph
- LVlinesChain: ordered set of LV lines (indices in the LV lines struct) to reach the current Linking Point

Meters/End nodes - struct

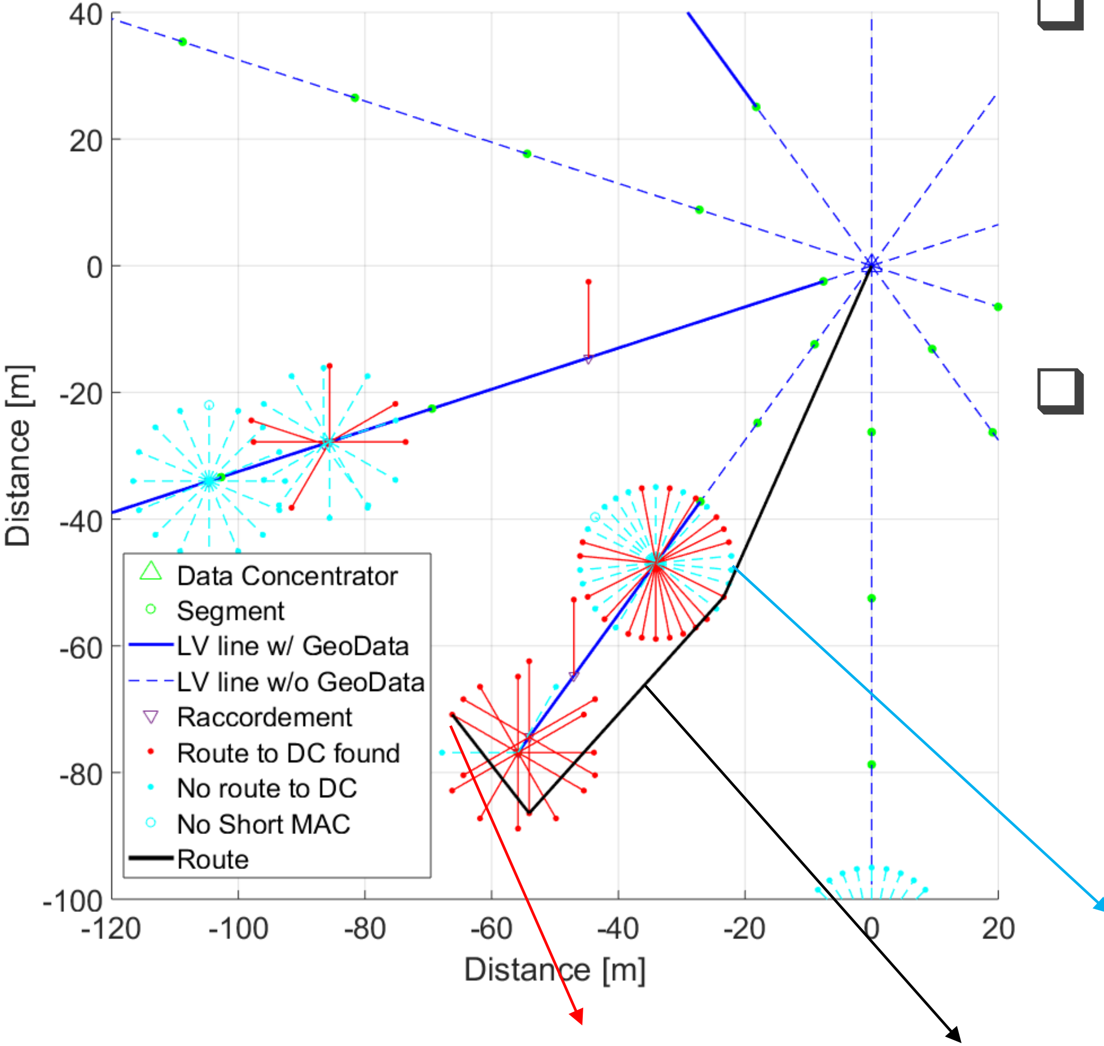
- Struct End nodes: full description of the end nodes, which also act as routers

Fields	ID	SM	pos	routeSM	racc_ind	racc_ID	sector	layer	met
133	"21447901564276"	"193"	[-46.0404,-4...	*1x2 string*	9	"461068141"	2	2	[23,12]
134	"21442402287867"	"212"	[-46.0404,-4...	[]	9	"461068141"	2	-1	[]
135	"21439218496229"	"294"	[-45.6335,-5...	[]	9	"461068141"	2	-1	[]
136	"21439507931837"	"218"	[-44.8335,-5...	*1x2 string*	9	"461068141"	2	2	[20,10]
137	"21445586079467"	"208"	[-43.6678,-5...	[]	9	"461068141"	2	-1	[]
138	"21446454386203"	"204"	[-42.1759,-5...	*1x2 string*	9	"461068141"	2	2	[21,11]
139	"21442981159084"	"83"	[-40.4087,-5...	[]	9	"461068141"	2	-1	[]

- ID and SM are both identifying sequences, where the latter is the Short MAC used by the routing protocol to map the nodes in the network
- pos: coordinates [x,y] of the node in the graph
- routeSM: if routing information is present, this vector contains the Short MACs of the nodes to go from the current one to the DC
- racc_ind: index in the Linking points struct where this node is found
- racc_ID: ID of the Linking point connected to this meter
- sector: as above
- layer and met: number of hops to the DC and metric from each node along the route to the DC

Route Tracing
+
Discovery of Routing Anomalies

Route tracing



Node with route to DC

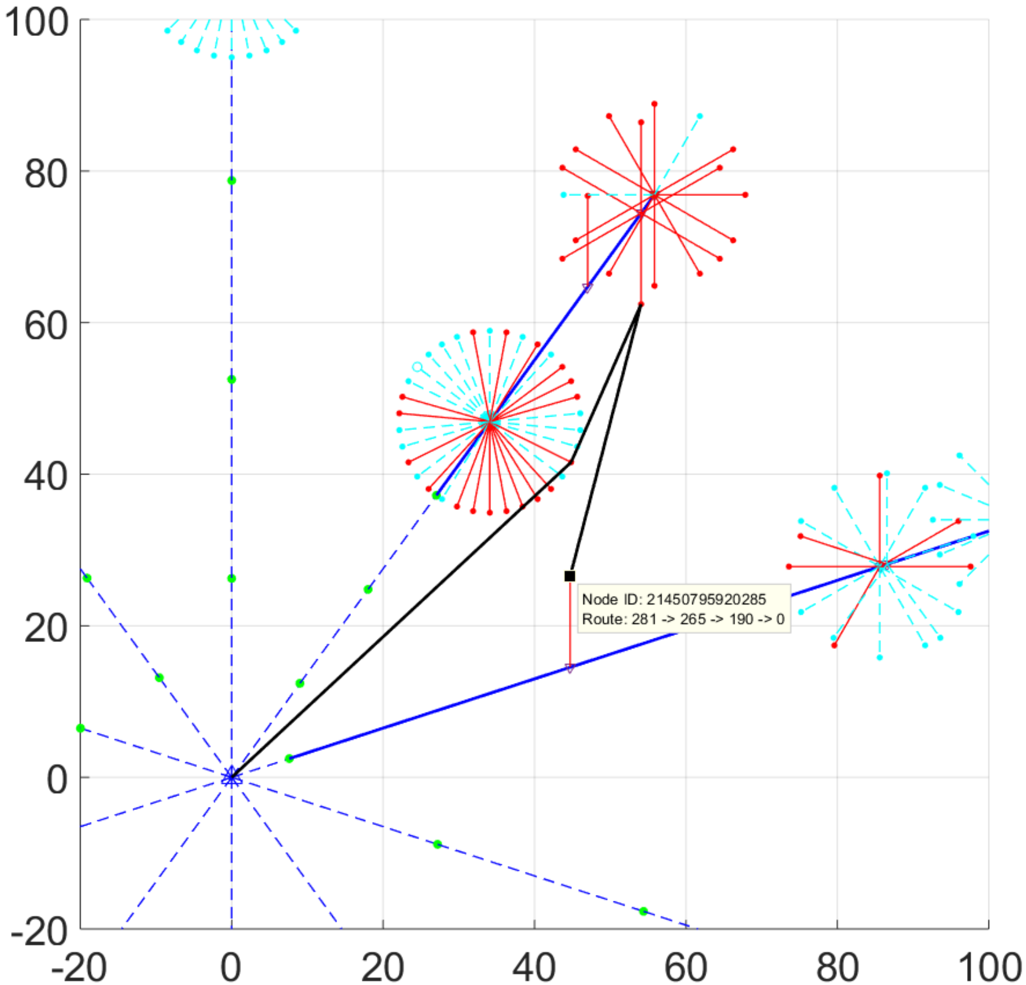
Routing path

Node with no route to DC

- **Challenge:**
  - Trace and display routes
  - Reformat data from router information base for visual interpretation
- **Result:**
  - Can visualize routes if information is found in log files

UNIVERSITÄT KLAGENFURT

Routing anomalies

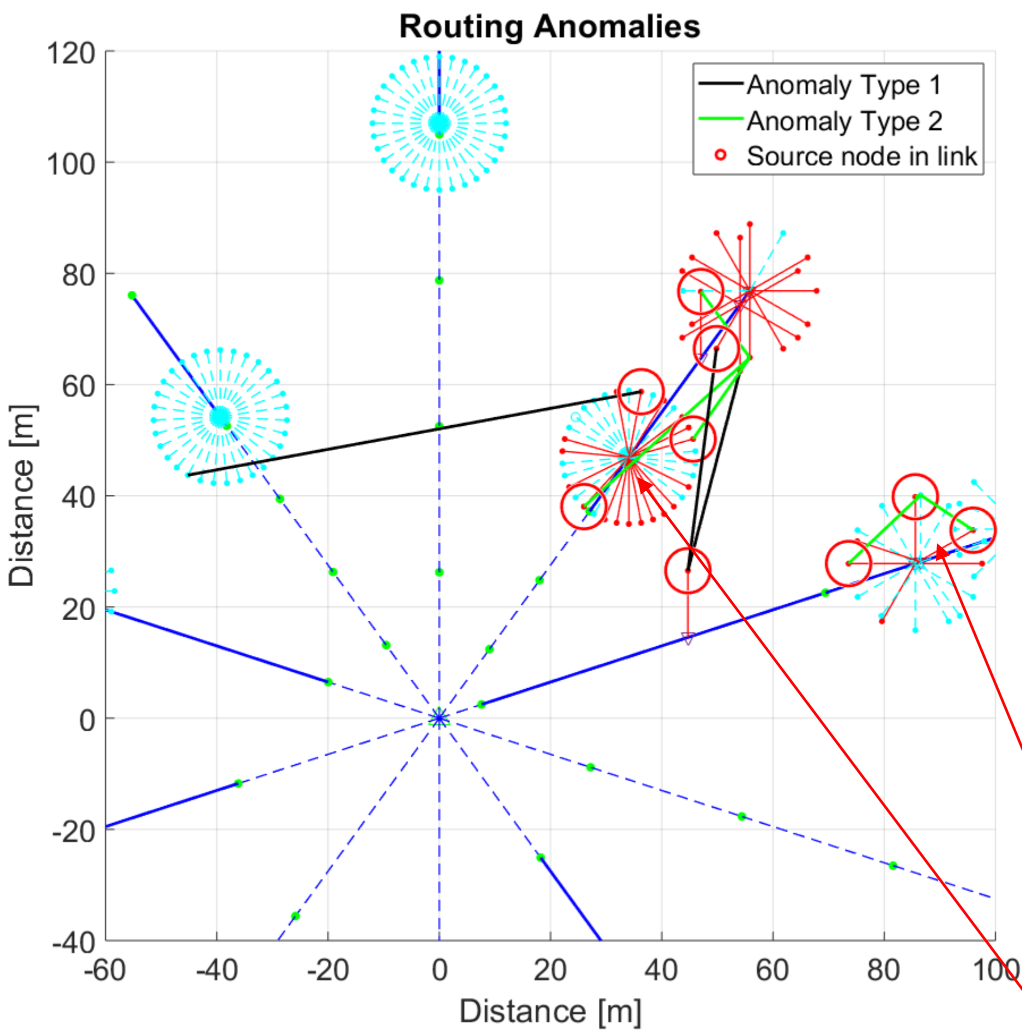


- Challenge:
  - Trace and display routes
  - Algorithm to detect anomalies
- Result:
  - Can visualize routes if information is found in log files, and…
  - …Anomalies found!
- Possible way forward:
  - Anomaly categorization

- Some routes from the log files display unexpected behavior
- Routes would be expected to move towards DC in the same sector

UNIVERSITÄT KLAGENFURT

Automatic anomaly detection



❑ **Challenge:**

- Automatic anomaly detection

❑ **Result:**

- Anomaly categorization:
- **Type 1,** if route connects nodes in different sectors
- **Type 2,** if the next hop along the route to the DC is further away from DC than the source node

Nodes are very close, false positive

Real anomaly

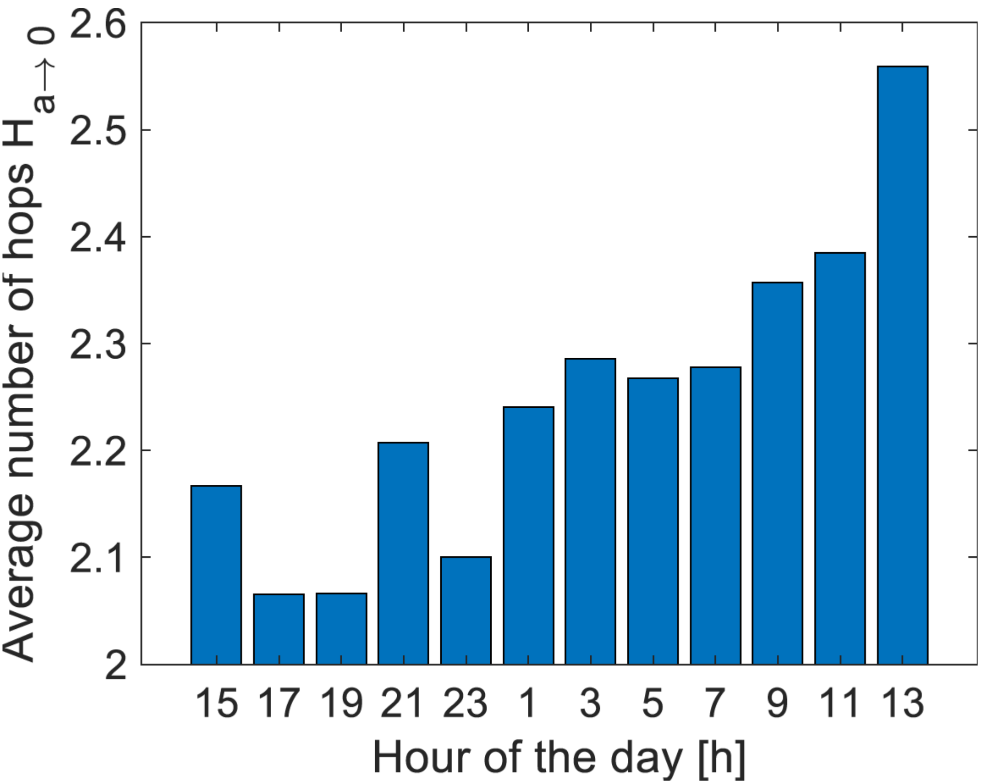
UNIVERSITÄT KLAGENFURT

Possible reasons for anomalies

- **Propagation/physical effects:**
  - Cross talk because cables are actually close
  - Very different channel attenuation in different paths
  - Time variability of noise
- **Successive Route Repairs** issued after link failures might create longer routes
- **Firmware bug** at layer 2-3
- **Network table** with data inconsistency

Routing time dependency

Time dependency



- **Time dependency**
  - Investigated how routing path changes over time
- **Anomalies (Type 3)**
  - Heuristic approach by defining

1. $I^{(1)}_{a\to b} = \max_t H_{a\to b}(t) - \min_t H_{a\to b}(t)$
2. $I^{(2)}_{a\to b} = \max_t \left|\frac{dH_{a\to b}(t)}{dt}\right|$

With $H_{a\to b}(t)$ number of hops between a and b observed in the time window t

- We spot an anomaly

If $I^{(1)}_{a\to b} > 3$

or

If $I^{(2)}_{a\to b} > 2$

Observations:

- Greater number of hops in the morning
- Number of hops should stay more or less constant
- When it grows too much, the route becomes more complex if Route Repair is used. Better procedure might be found.

Possible reasons for anomaly Type 3

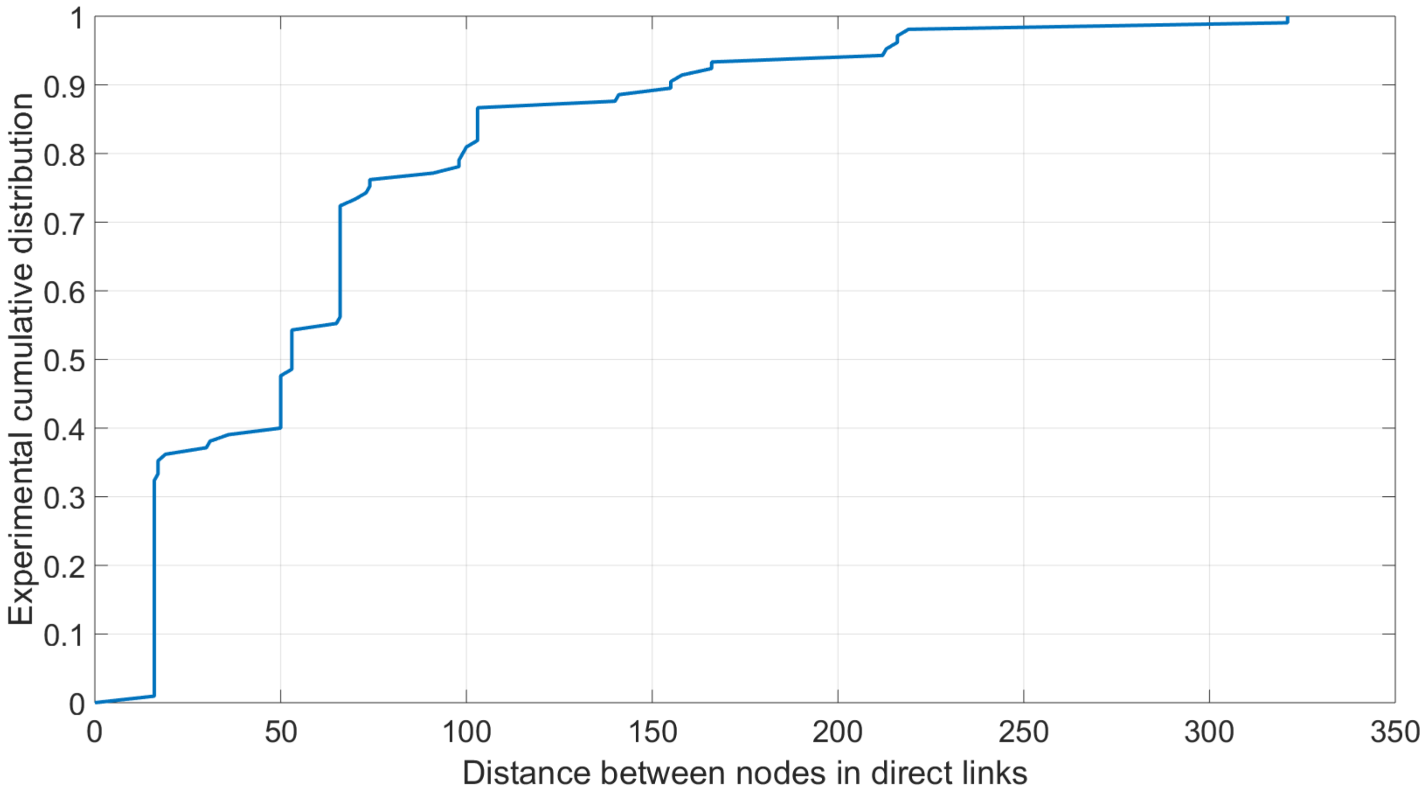
- **Bad links cause Route Repair generation:**
  - **RREQs** are reissued to reconstruct route
  - Path information is not found locally. This may increase latency.

- **Detection of anomalies can highlight anomalies in the data set**
  - We found that a packet can be rerouted through nodes with MAC addresses not found in the current network. This is something apparently wrong.

Analysis of Topologic and Geometric Data

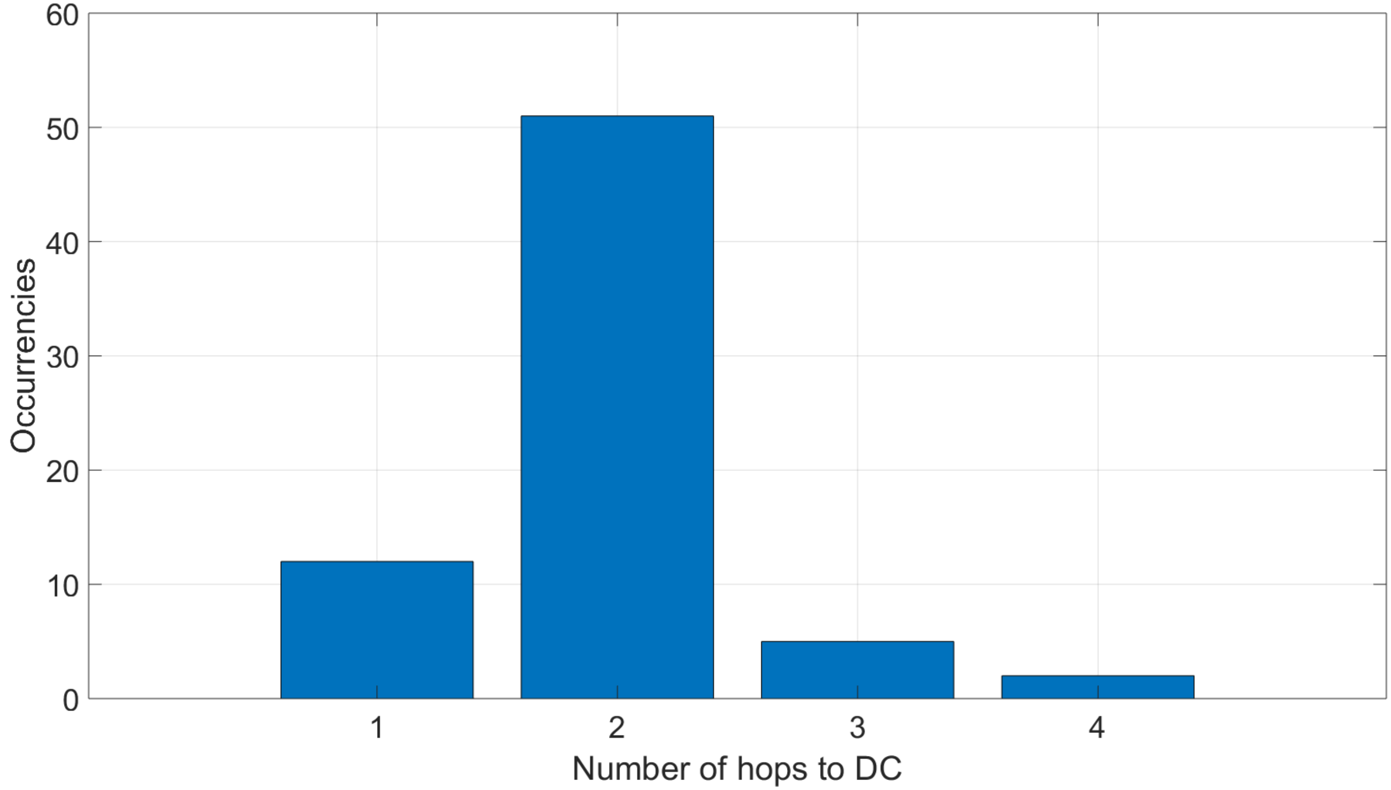
Electrical distance in direct route links

CDF



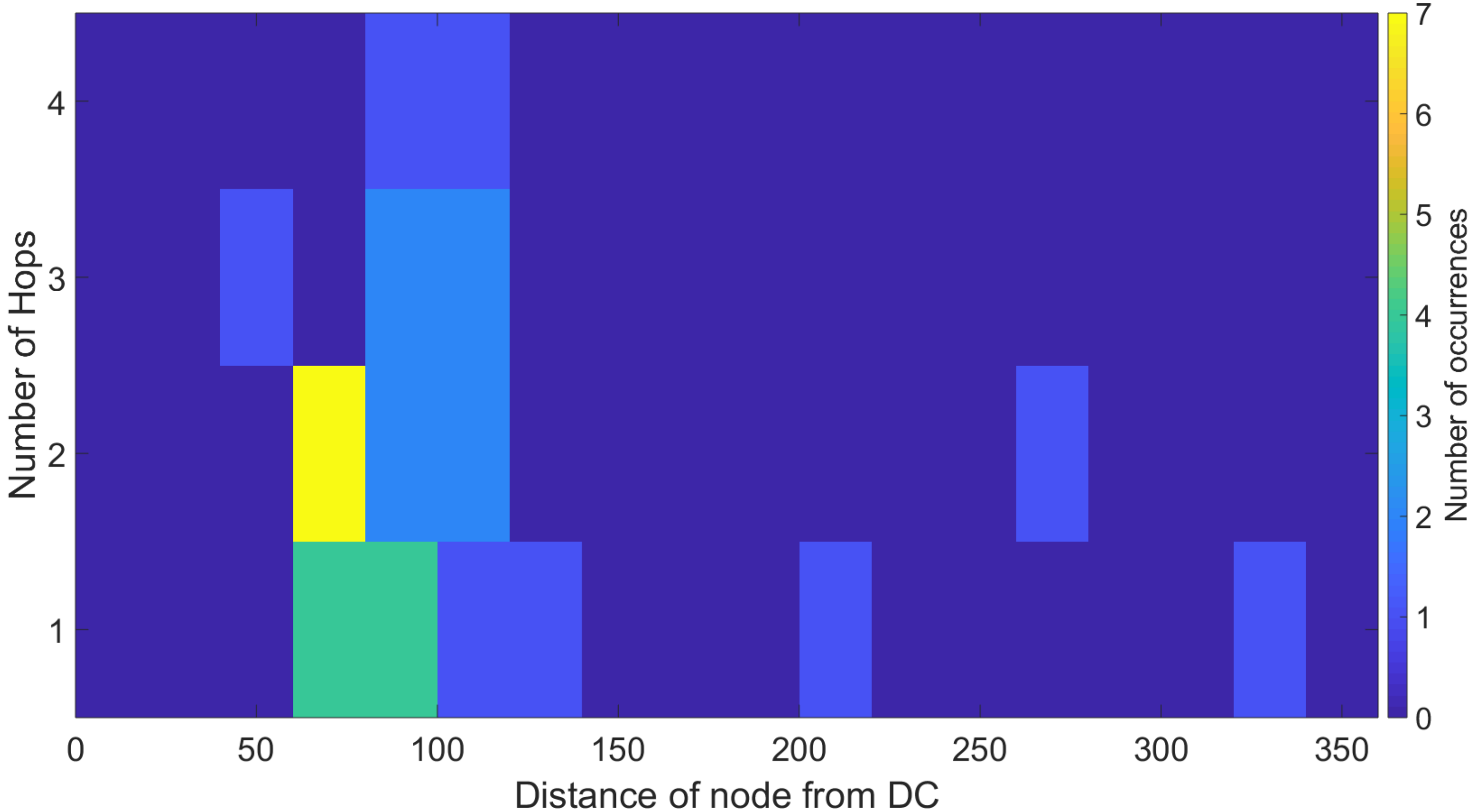
- **Direct Route Link:** when a couple of nodes appears at a 1-hop distance in any route
- **Observations:**
  - 33% of direct links have length lower than 16 m – this is because the supposed distance of a meter from the "raccordement" to which it is connected to is 8 m
  - 80% of direct links have length 100 m or shorter
  - Direct links cannot be longer than 320 m – that is also roughly the distance of the furthest node from the DC

How many hops to get to the DC?



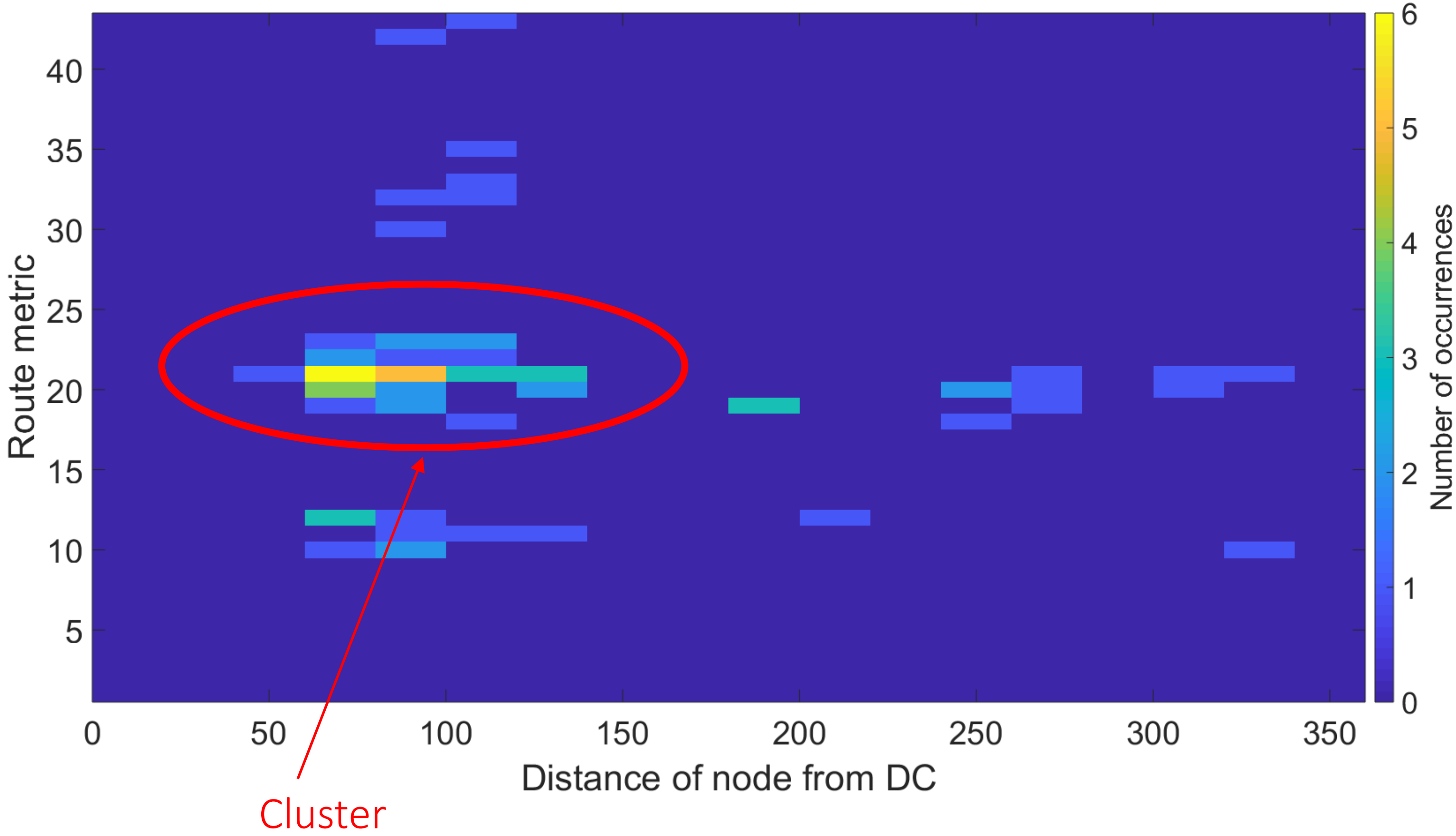
- **Observations:**
  - The DC log files report most nodes being able to connect to the data concentrator (DC) with two hops, i.e., 50 nodes reach the DC in 2 hops
  - Does this relate to a node's distance from the DC? See next slides ...

Number of hops to DC vs distance from DC



- **Observations:**
  - The distance considered here is the electrical distance of a node from the DC
  - To obtain statistical value, we need to process more data

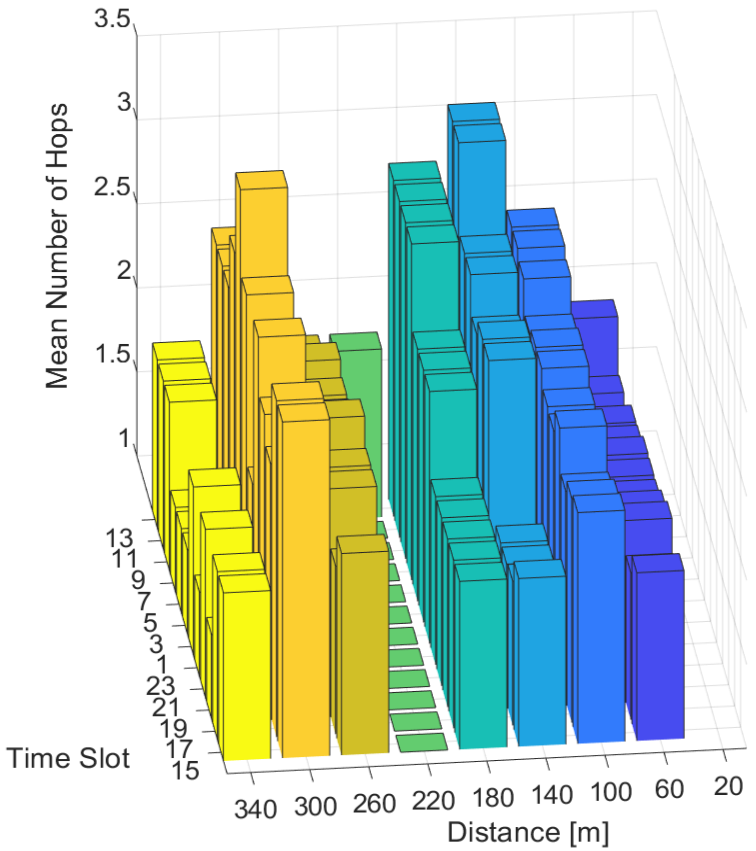
Route metric (Cost) to DC vs distance from DC

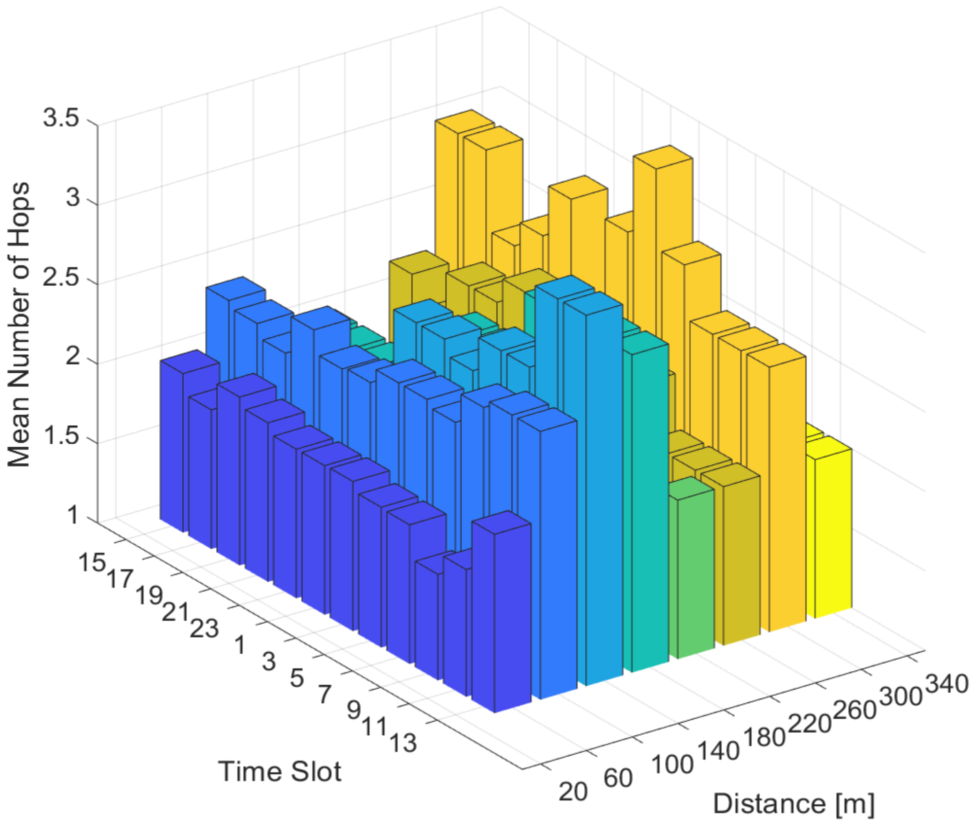


❑ Observations:

- Rout metric in line with hops: each hop cost is in the range [4,14] as per standard.
- Cluster: in line with the high number of occurrences of 2-hop links around the same distances.

Number of hops to DC vs distance from DC and time

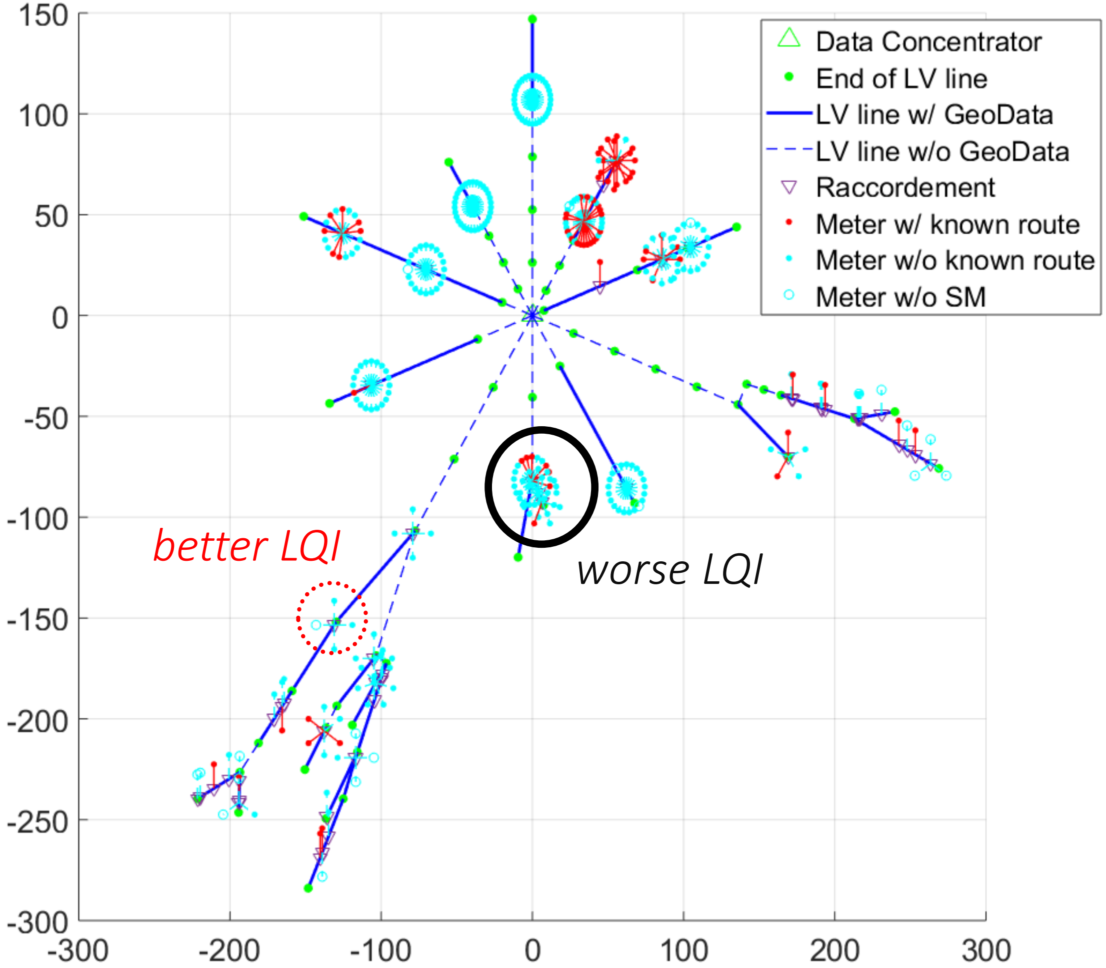




- ❑ **Time slot:** sampling every 2 hours
- ❑ **Observation**
  - The channel is time variant. The link quality is influenced not only by distance but also by environment, namely, human activity

UNIVERSITÄT KLAGENFURT

Channel response vs topology



- **Channel response:**
  - Link quality to the DC might be better for the nodes in the red/dashed circle than for the ones in the black/thick one despite longer distance
  - Density of loads is influential. More loads = more noise, more signal reflection, more interference, smaller SNR
  - How to characterize density for these networks?

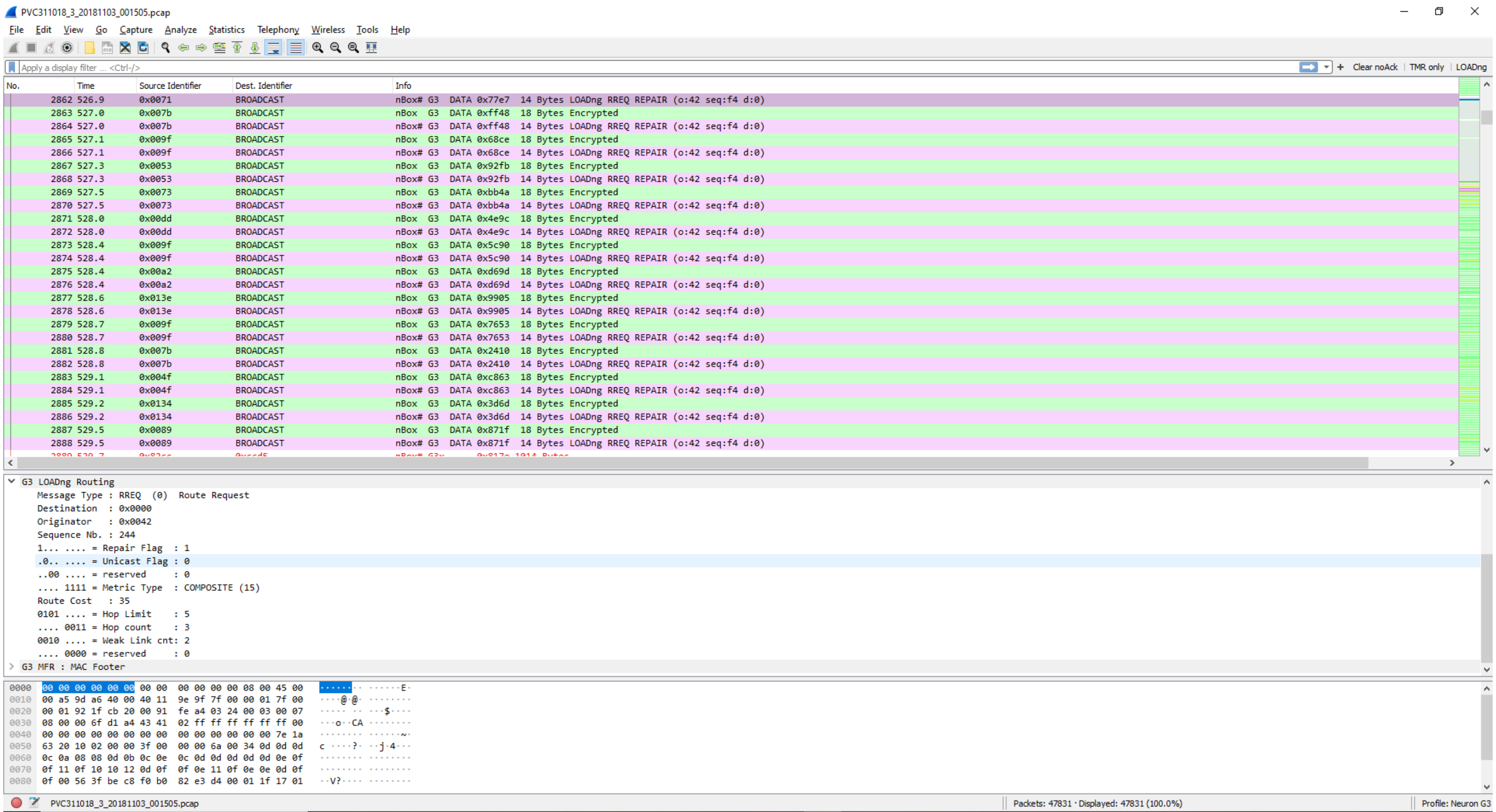
- **Observation:**
  - There exists a relation between routing path and electrical topology but distance is not all

UNIVERSITÄT KLAGENFURT

On Wireshark traces

Wireshark traces

- Activity was started to analyze WireShark traces (Routing packets)



Wireshark traces cont.ed

- Wireshark provides data for Layers 1-3
- Most traffic in the traces is generated by the Sniffer (called *noise measurement*), probably to verify that it works
- Most routing packets in the traces were Route Repairs
  - Route Repairs are issued after a link failure
- Detailed traffic reconstruction
  - It is possible to label links in the recostructed electrical topology with Wireshark data
  - However, not all nodes are able to reach the Sniffer. So info is incomplete
- For analysis of routing the Log files are more immediate than the Wireshark traces

Conclusions

Conclusions

- Quite challenging tasks with the convoluted data sets, nevertheless encouraging results
- Identified most relevant semantics in the data sets
- Developed automatic procedure for network visualization
  - Able to define proper data structure for visualization
  - Visual tool can be used to discover anomalies in data sets (missing/erroneous data) or in routing
- Started activity to infer relations between routing and topology
- Only initial look at Wirshark traces
- Submitted a paper to ISPLC

Future work (Phase 2)

Future developments

1) Data analytics

- We are in the position to:

✓ Define primitives that recursively call Path Requests to recreate complete routes at the DC

✓ Process larger data sets and discover other properties/anomalies

✓ Develop procedures for analysis and anomaly detection also exploting machine learning

✓ Propose improvements to routing

→ Data formats and primitives might be part of new specs

→ Help to identify implementation bugs

Future developments

1) Improvement of routing protocol

- More efficient waiting time strategies/jitter in the nodes for RREQ/RREP depending on past knowledge of the network
- RREQs are currently broadcasted. *Gossiping* is a technique of performing broadcast transmission with a likelihood < 1. This can be implemented with local network parameters yielding faster route establishment
- Implementation of *Collection Tree Protocol*, as specified in [1]
- Exploitation of Reinforcement Learning in nodes for adaptation of route metric calculation

[1] J. Yi, Jiazi, T. H. Clausen. **Collection Tree Extension of Reactive Routing Protocol for Low-Power and Lossy Networks**. Hindawi journal of distributed sensor networks. 2014. 10.1155/2014/352421.

Future developments (cont'd)

3) Other research paths

- Grid diagnostics and maintenance
- Improvements at PHY (adapatation and resilience). Routing is not all !
- Look at the broadband spectrum for possible evolution of G3-PLC
- Hybrid deployments: PLC/LoRa/Zigbee/WiFi
- Exploit Machine Learning approaches in all stacks

4) Convergence with other standards ?

5) Definition of new (or less exploited) applications

Ref. A. M. Tonello, N. A. Letizia, D. Righini, F. Marcuzzi, "Machine Learning Tips and Tricks for Power Line Communications," IEEE ACCESS, 2019.
Ref. . Passerini, A. M. Tonello, "Smart Grid Monitoring Using Power Line Modems: Anomaly Detection and Localization," IEEE Trans. on Smart Grids, 2019
Ref. A. M. Tonello, A. Pittolo, "Considerations on Narrowband and Broadband Power Line Communication for Smart Grids," Proc. of IEEE SmartGridComm 2015.

IEEE ComSoc Technical Committee on Smart Grid Communications

Consider becoming member:

https://comsoc-listserv.ieee.org/cgi-bin/wa?SUBED1=asgcc&A=1