UNIVERSITY OF DENVER

Data-Driven Scalable Phase Identification in Modern Distribution Grids

- Behrouz Sohrabi
- Amin Khodaei, Ph.D.

November 9th, 2022

Electrical and Computer Engineering Department

University of Denver

Outline

- Introduction to Phase Identification
- Phase Identification Solutions
- Proposed Model
- Analysis Results
- Conclusion

Phase Identification

Photo: Anish Regmi

Why is this necessary?

- Power Grid Operation and Maintenance Procedures
- Distribution Network Load Balancing
- Component connectivity is often altered and not documented, especially after outages
- Grid topology reconstruction begins with identifying the phase connectivity of loads

Phase Identification

Hardware-Driven Solutions:

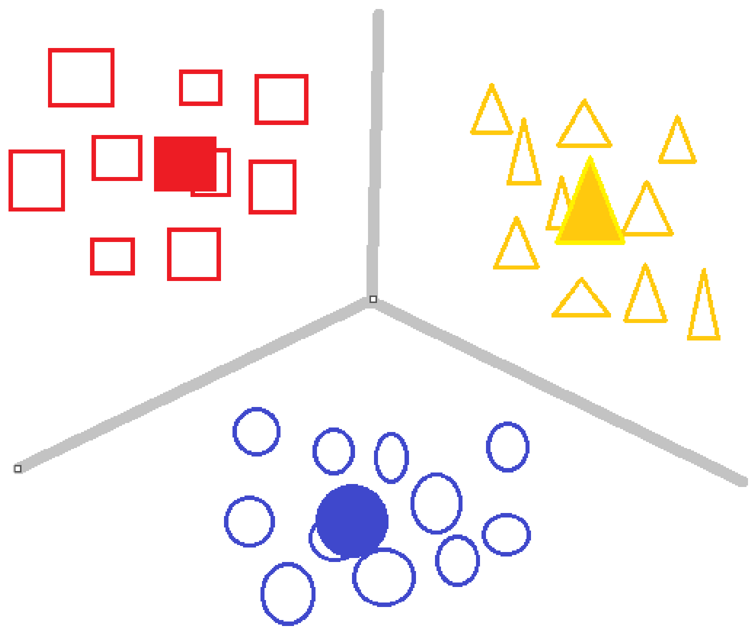
- Micro PMUs
- Voltage Signal Injection

A phasor measurement units

Phase Identification

Data-Driven Solutions:

- Voltage Clustering
- Voltage Correlation
- Consumptions Correlation
- Consumption Aggregation

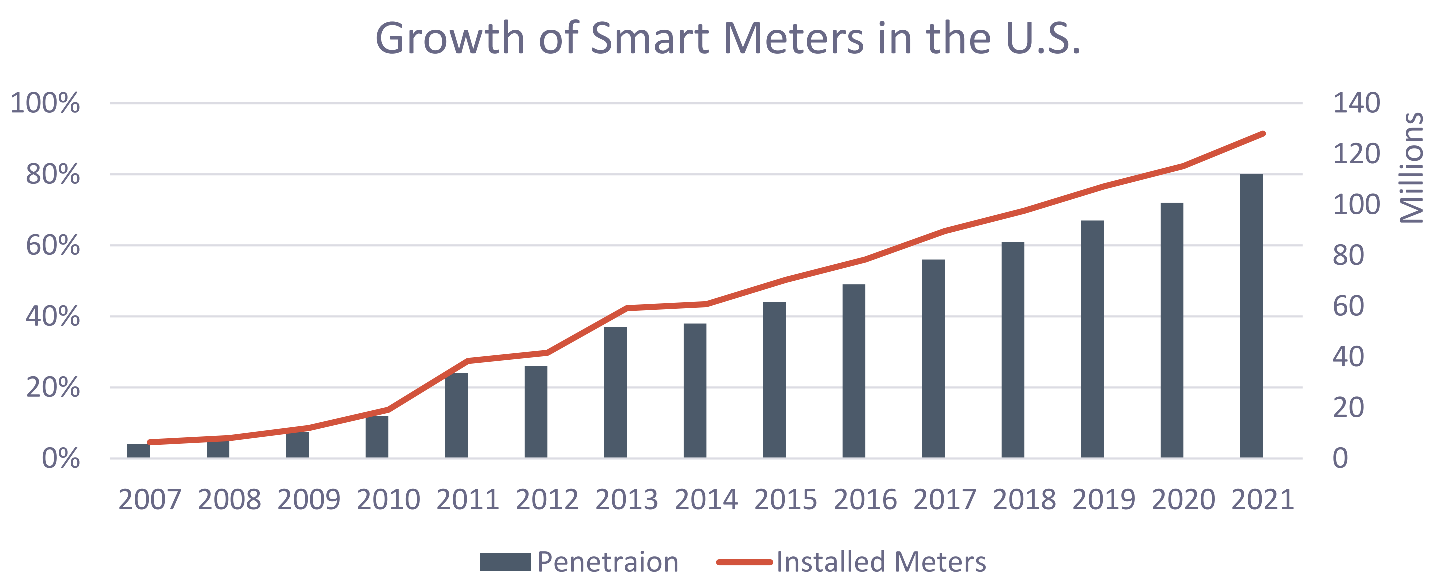


Basics of Clustering by Justin Osborne

Modern Distribution Grid

Using Smart Meters Data:

- Smart meters’ growth rate increases the observability of the network
- The data obtained from smart meters can be used for phase identification



Smart Meters growth in the United States (Source: The Federal Energy Regulatory Commission)

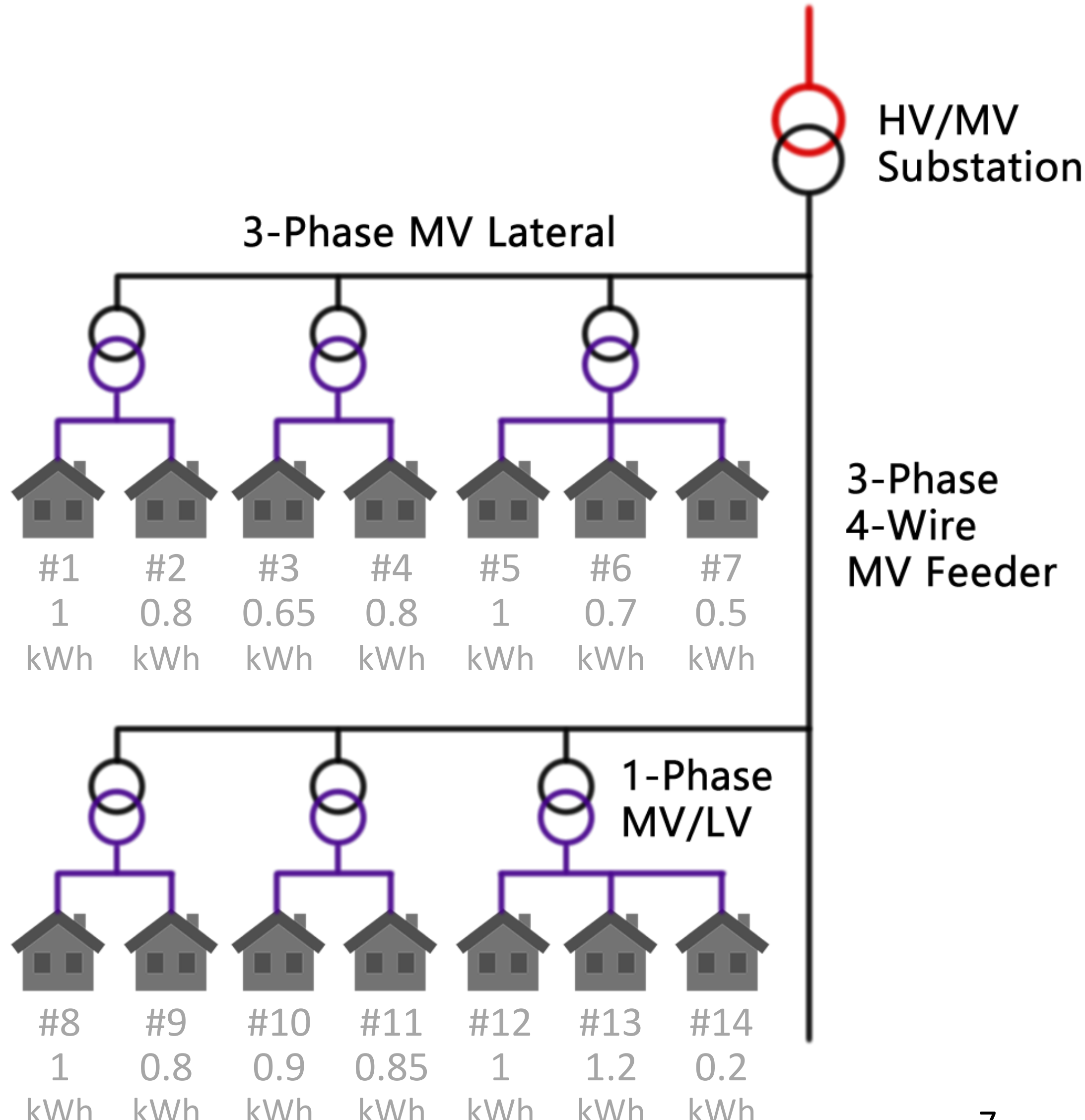
Solution

A simplified scenario:

Having consumed power measurements from smart meters and the distributed load from the substation for a given time interval:

Substation Measurement:

- Phase A = 3.5 kWh
- Phase B = 4.0 kWh
- Phase C = 4.5 kWh



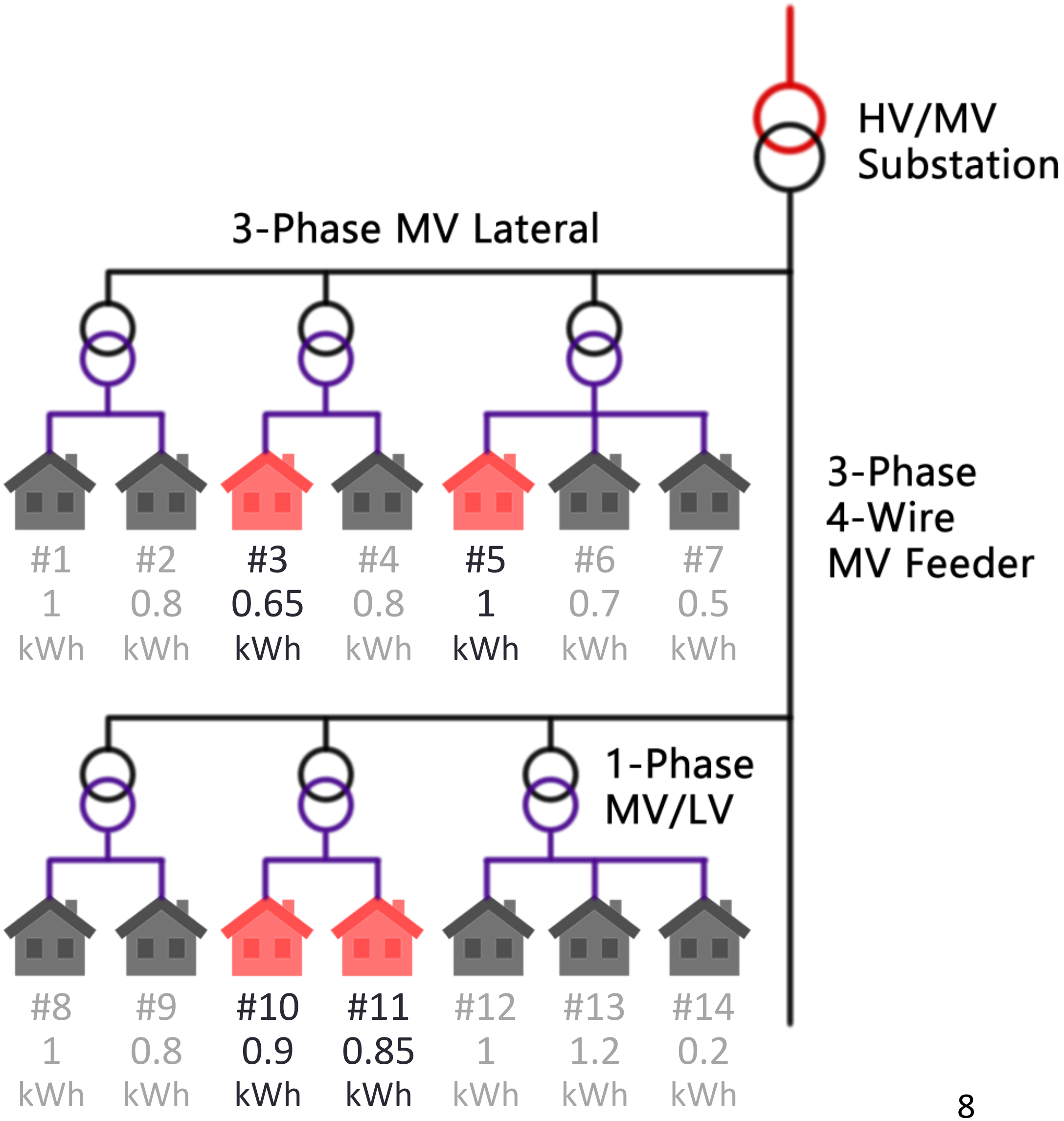
Solution

For Phase A:

Substation Measurement:
- 3.5 kWh

Aggregated Consumptions of #3, #5, #10, #11:
- 3.4 kWh

Distribution Loss:
- 0.1 kWh



Solution

For Phase B:

Substation Measurement:

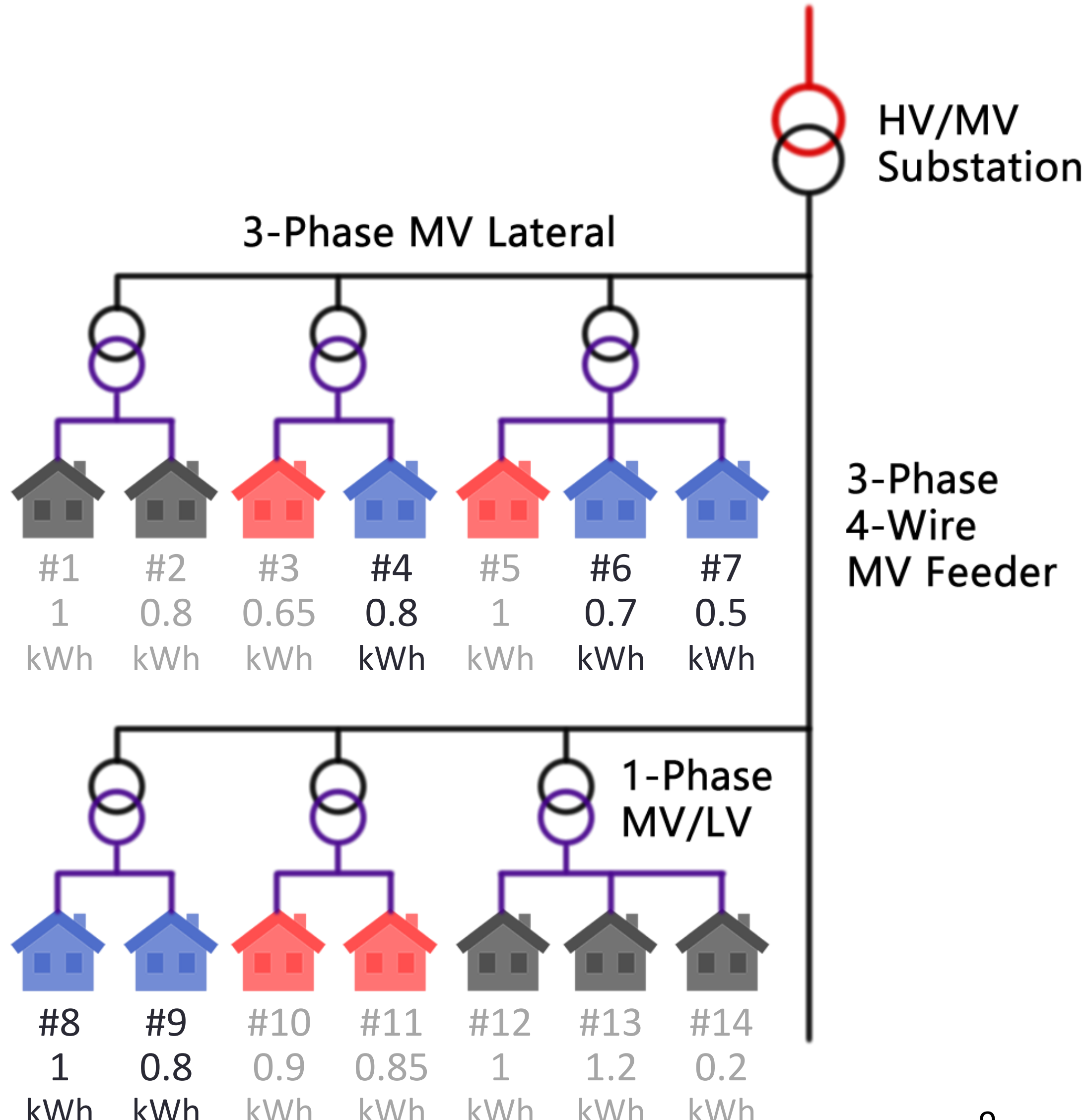
- 4 kWh

Aggregated Consumptions of #4, #6, #7, #8, #9:

- 3.8 kWh

Distribution Loss:

- 0.2 kWh



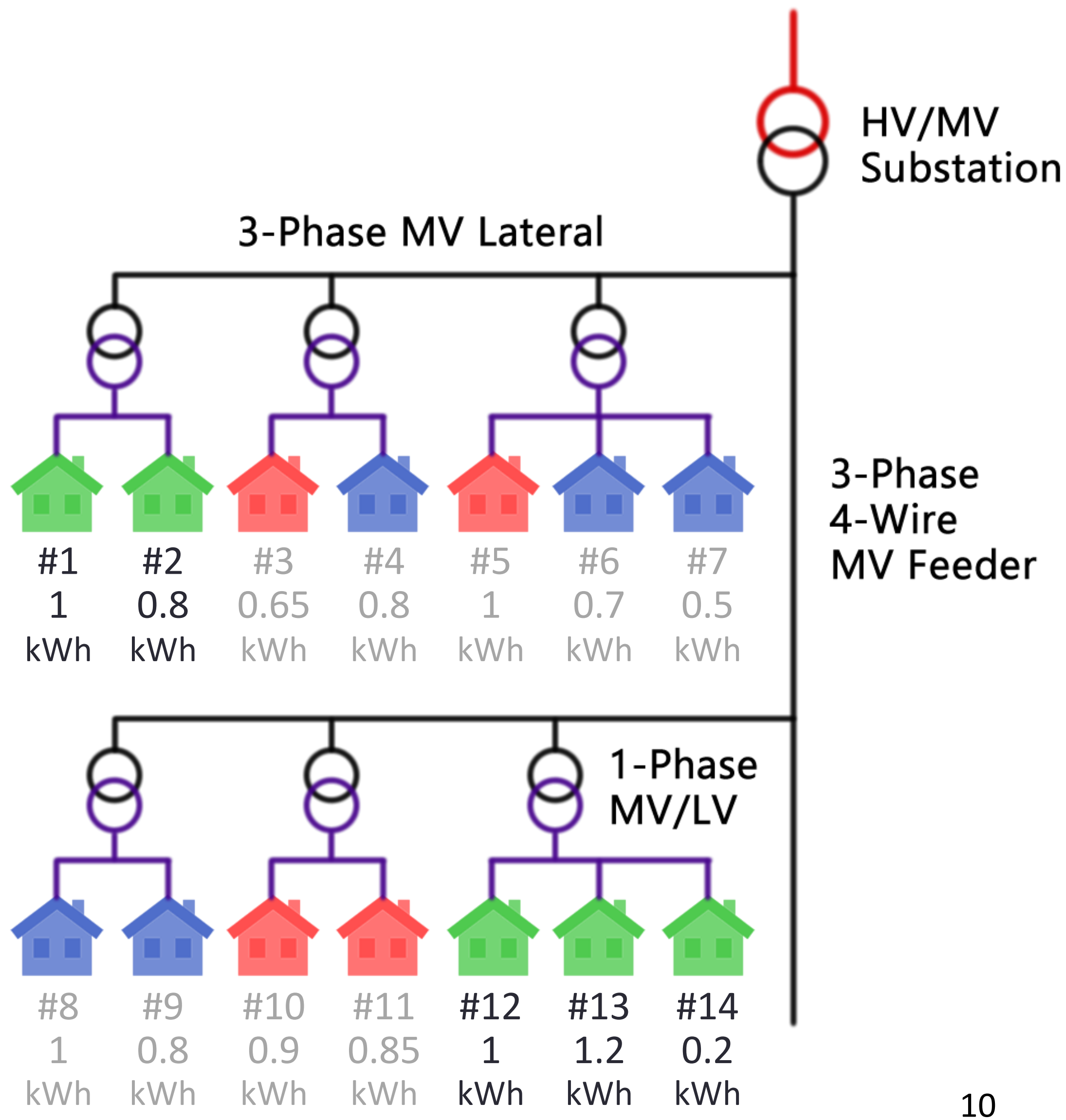
Solution

For Phase C:

Substation Measurement:
- 4.5 kWh

Aggregated Consumptions of #3, #5, #10, #11:
- 4.2 kWh

Distribution Loss:
- 0.3 kWh



Proposed Model

Mixed-Integer Quadratic Programming Model:

Find connectivity of the loads by minimizing the squared difference between consumed and lost load with distributed load, for each phase, at each measurement.

$$\min_{x} \sum_{q=1}^{Q} \sum_{t=1}^{T} \left(\sum_{n=1}^{N} a_{tn} x_{qn} + p_{qt}^{L} - p_{qt}^{S} \right)^{2}$$

s.t. $\quad x_{qn} \in \{0, 1\} \qquad \forall n \in N, q \in \{a, b, c\}$

$$\sum_{q=1}^{Q} x_{qn} = I_n \qquad \forall n \in N$$

Proposed Model

Expectations:

- A Data-driven method that employs smart meter and substation measurements
- Accounting for measurement loss and noise
- Accounting for distribution loss
- Accounting for load profiles homogeneity
- Applicable to large-scale networks
- Applicable to networks with lower than 100% AMI penetration rate
- Independent of the number of consumers in each phase
- Assuming that all consumers' phases are unknown

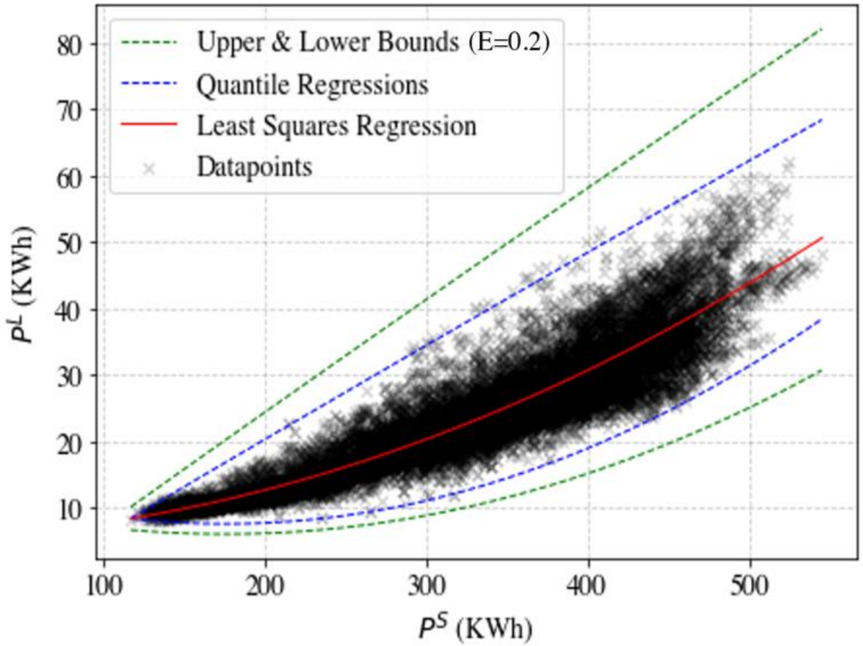
Distribution Loss

Analyzing distribution loss of the network:

- Least squares regression algorithm to estimate distribution Loss
- Quantile regression to find the lower and upper bounds of the estimation
- The Boundaries can be fed to the model as constraints
- A coefficient (E) to relax the constraints to account for outliers

Distribution loss estimation for 21K demand and loss datapoints

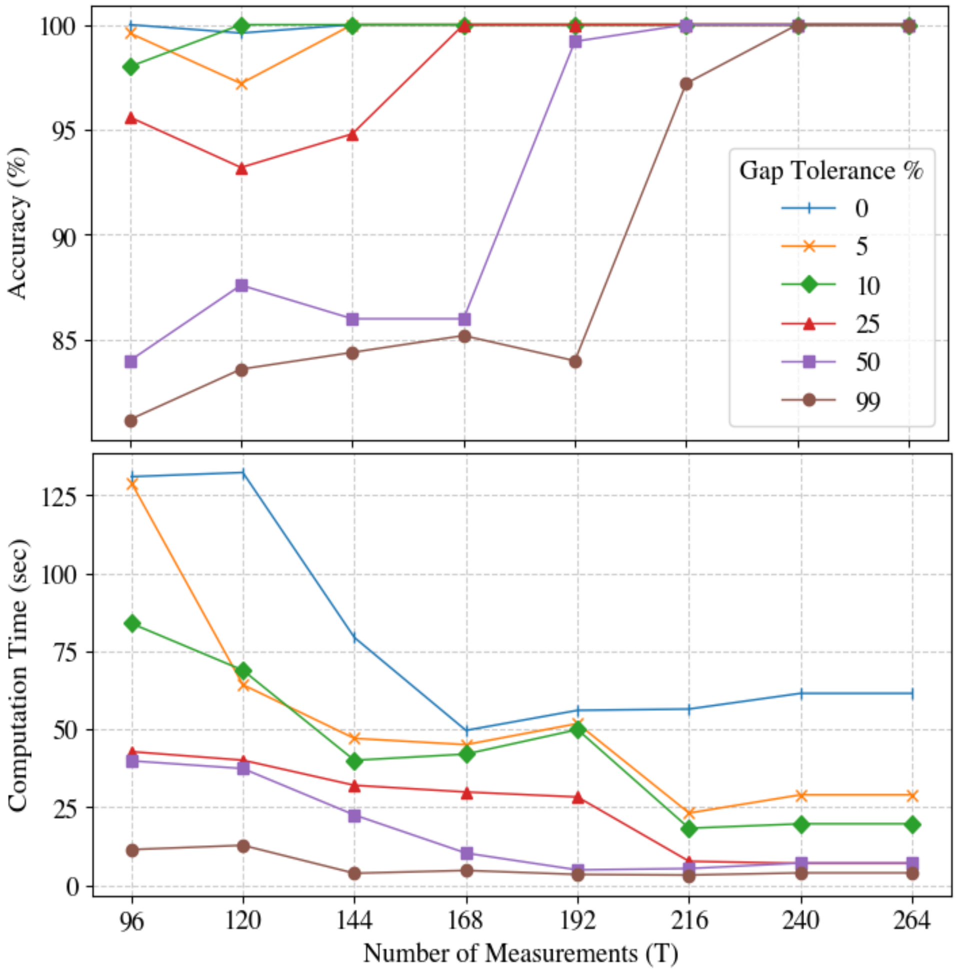
Distribution loss



Distributed Load

Measurements Window

Analyzing a system with 250 loads



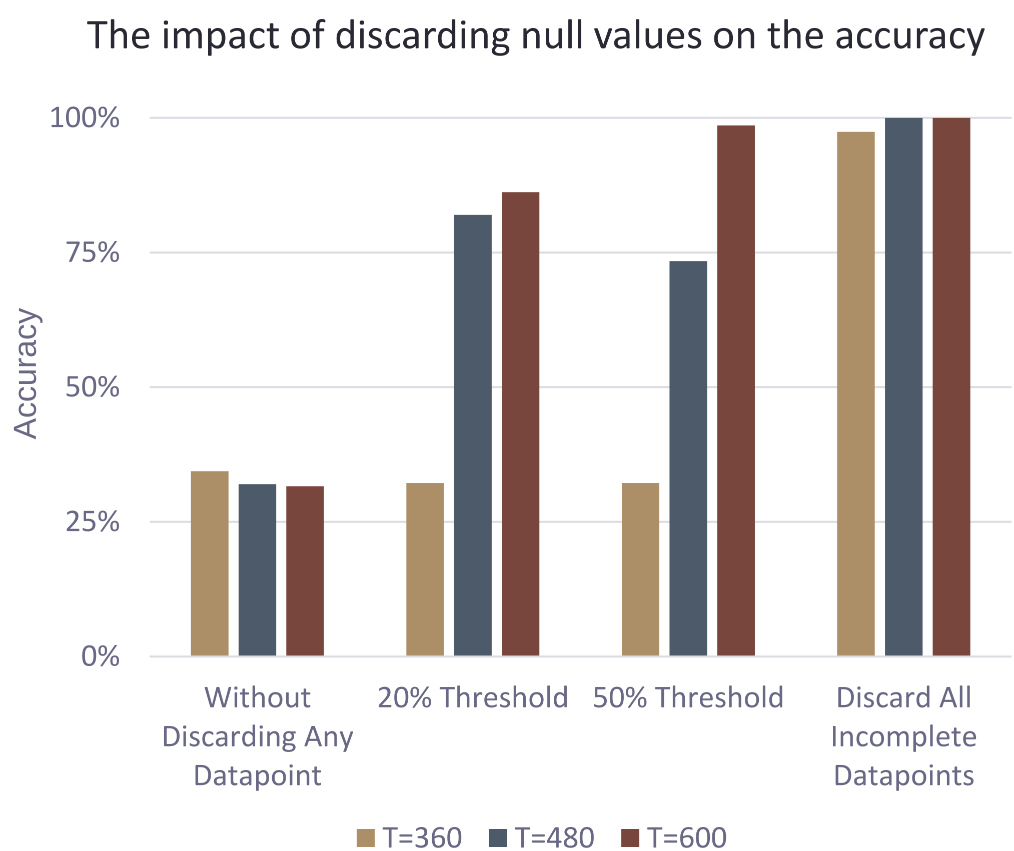
- A wider measurement range increases the accuracy.
- A higher optimality gap requires more measurements.
- The number of consumers strongly affects both parameters.

Measurement Loss

Analyzing a system with 500 loads

(30% null values - 360, 480, and 600 available measurements)

- Unrecorded values in the consumption measurement that produce outliers
- This will result in a disparity between the amount of distributed load and the consumed load
- The impact of null values decreases as the window of measurements increases
- The best approach is to discard all uncomplete datapoints



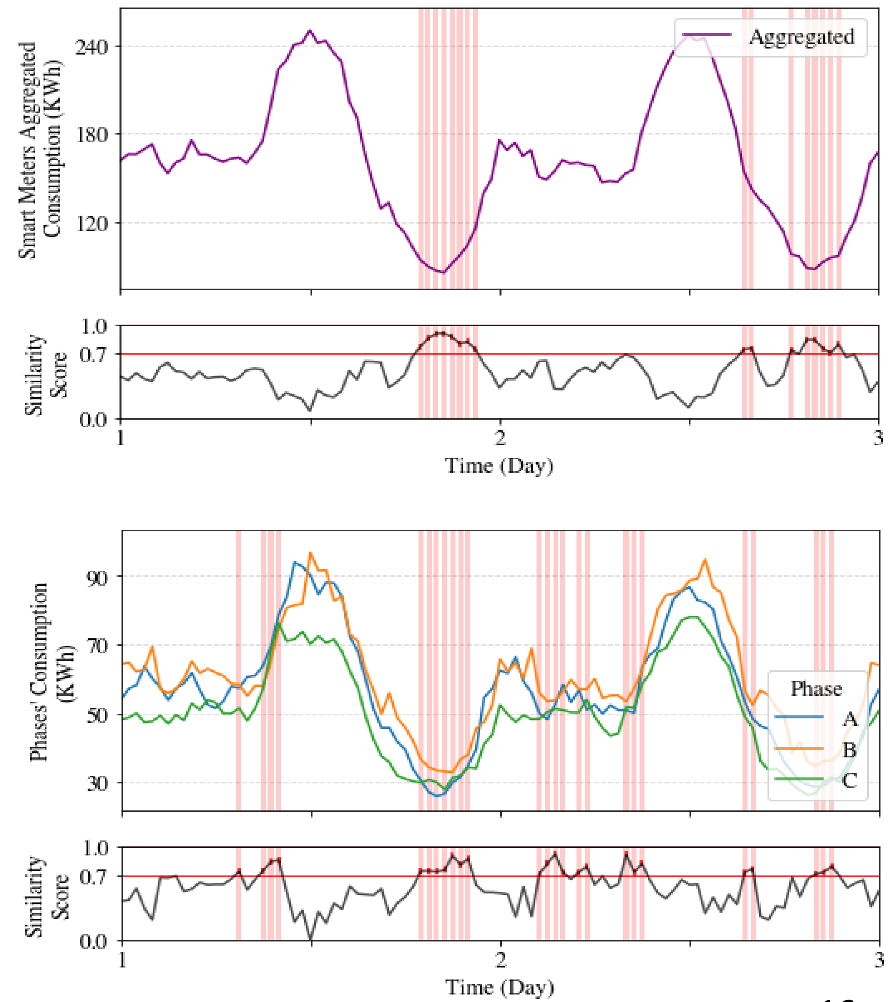
Load Profile Similarity

- Load profiles usually have similar patterns
- Datapoints with close values of measurements will cause the solver to reach the wrong solution
- Each data point is labeled with a similarity score

$$S_t = 1 - \frac{\sigma_t - \min(\sigma)}{max(\sigma) - \min(\sigma)} \quad \forall t \in T$$

- Data points with a high similarity score should be discarded

The similarity score of load profiles in a system with 500 loads over 96 measurements

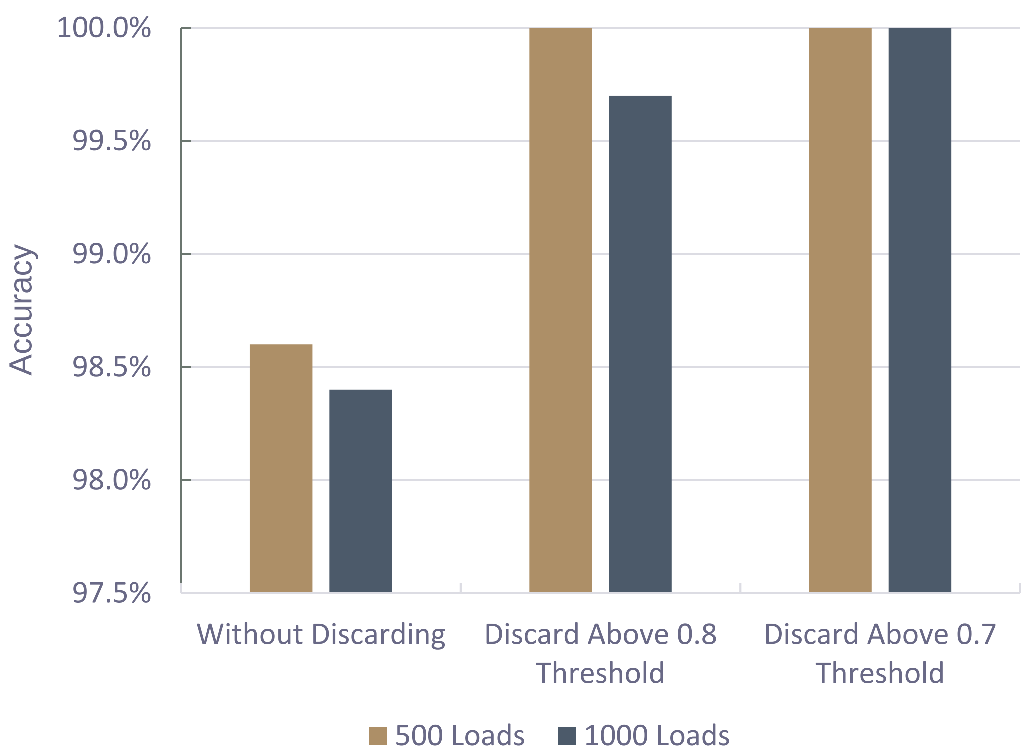


Load Profile Similarity

Analyzing two systems with 500 and 1000 loads

(960 available measurements)

The impact of discarding datapoints on the accuracy



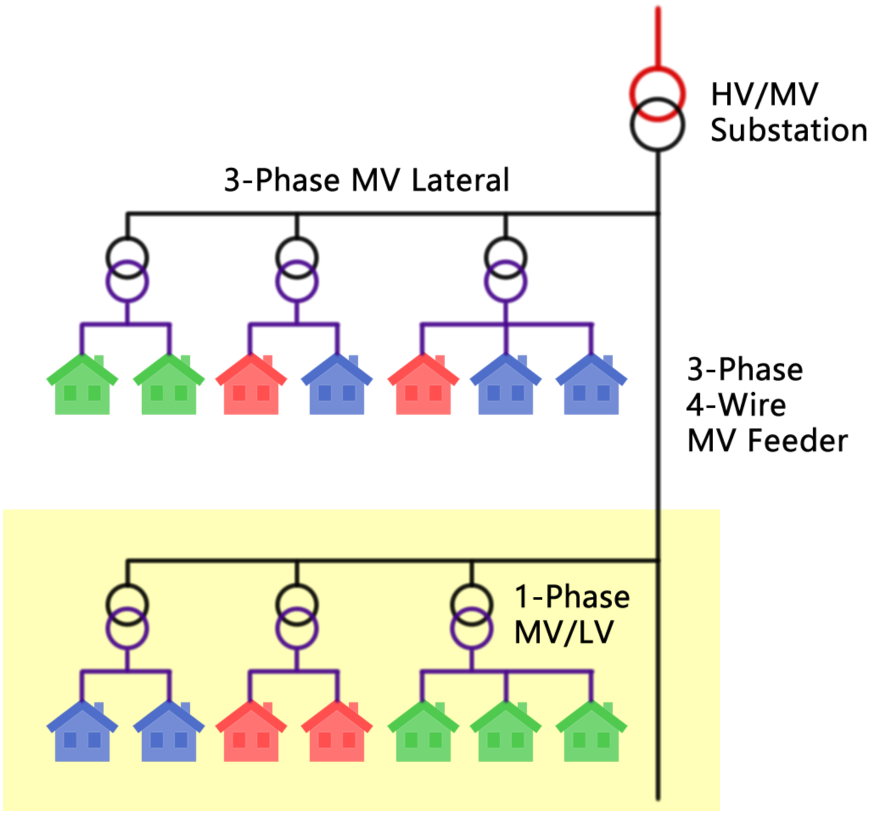
- The impact of profile similarity increases with the size of the network.
- Choosing the data points with the lowest similarity score is the most effective approach.

Clustering Single-Phase Loads

To reduce the complexity of the problem single phase loads on the same phase can be considered as one load.

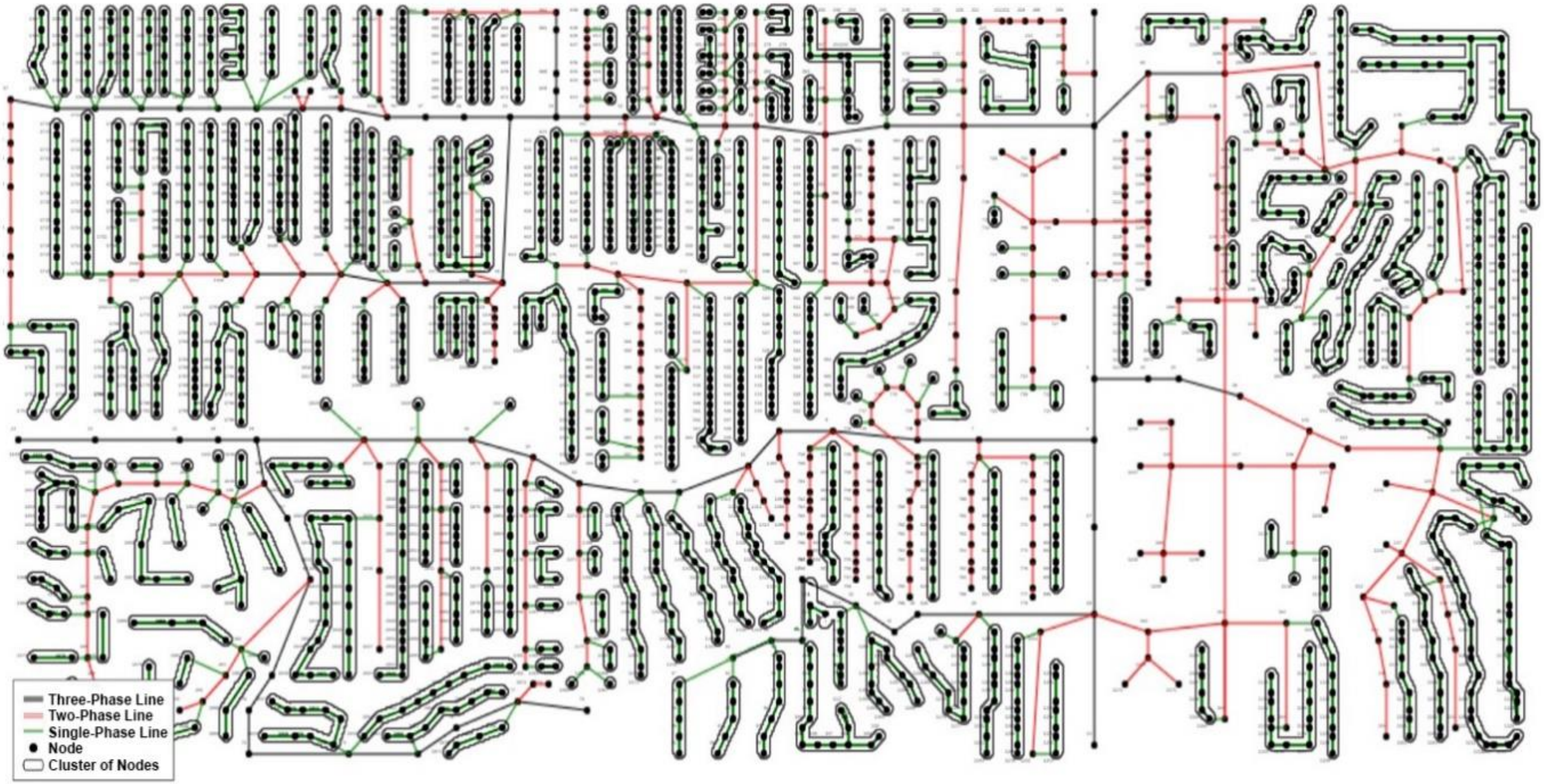
- Using minimal topology data:
  - Geographical Coordinates
  - Type of the line
  - Nearby Loads

- DBSCAN Clustering algorithm



Clustering Single-Phase Loads

Allocation of single-phase nodes in a 2500-load-bus system



Clustering Single-Phase Loads

Analyzing a system with 2500 loads

# Of Clusters	N After Clustering	T Measurements	Computation Time (s)	Accuracy
Without Clustering		1920	~5 hours	99.56%

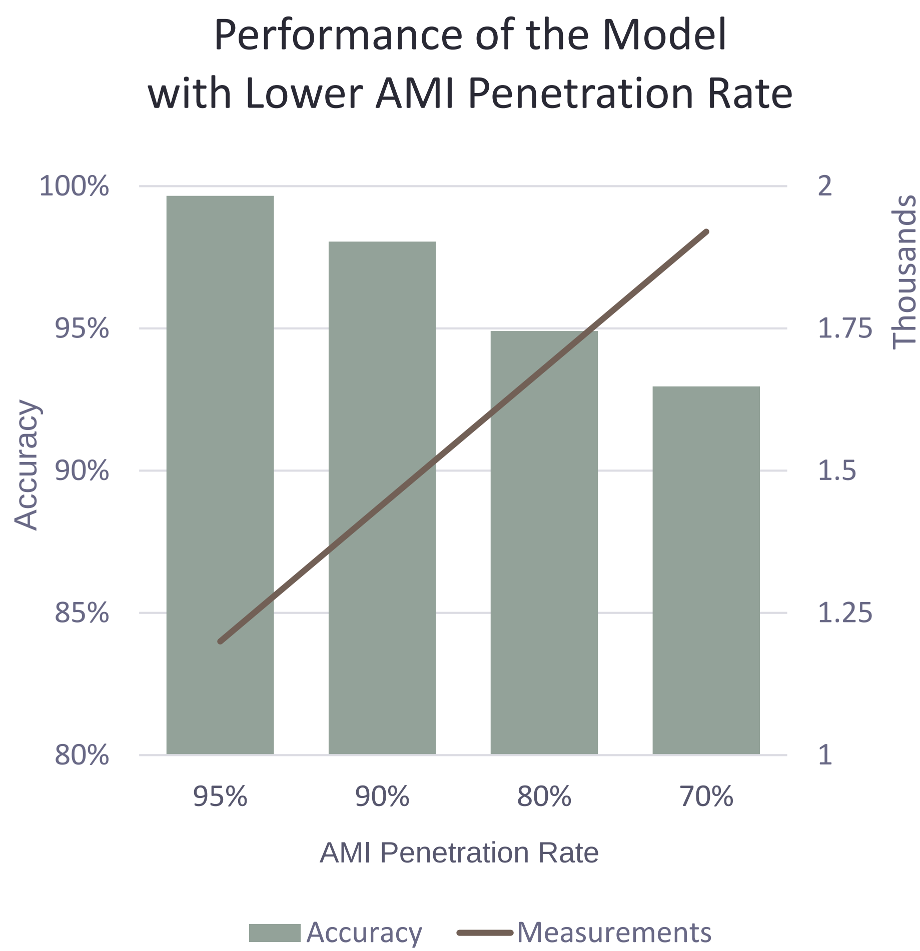
- The complexity of the model reduces
- Clustering can be done with only minimal information from topology data.
- The proposed model is independent of the topology data

Clustering Single-Phase Loads

2500-load-bus system with AMI penetration rate < 100%

(1205 Loads After Clustering)

- Performance is improved by the distribution loss lower and upper bounds
- This method is still effective for older networks that are not fully equipped with smart meters
- It is possible to identify the phase of the unmetered loads



Conclusion

- A scalable data-driven method for phase identification in distribution grids
- The model was challenged by the appearance of noise and outliers with increasing network size
- To increase performance, a regression algorithm was utilized to estimate the distribution loss
- The optimality gap and measurement interval should be selected appropriately
- To choose optimal data points, two criteria were evaluated. Similarity Score and null value ratio
- Using 40 days of measurements and only consumption data, the phase of 2500 loads was identified with high accuracy
- Single-phase loads were clustered using the DBSCAN algorithm, which enhanced the performance in networks with a low AMI penetration rate.

Thank You

Behrouz.Sohrabi@du.edu

References

[1] A. Madureira, J. Pecas Lopes, A. Carrapatoso and N. Silva, "The new role of substations in distribution network management," CIRED 2009

[2] B. Liu, K. Meng, Z. Y. Dong, P. K. C. Wong and X. Li, "Load Balancing in Low-Voltage Distribution Network via Phase Reconfiguration: An Efficient Sensitivity-Based Approach,"

[3] L. Marrón, X. Osorio, A. Llano, A. Arzuaga and A. Sendin, "Low voltage feeder identification for smart grids with standard narrowband PLC smart meters," in 2013 IEEE 17th International Symposium on Power Line Communications and Its Applications, 2013.

[4] A. Sendin, A. Llano, A. Arzuaga and I. Berganza, "Strategies for PLC signal injection in electricity distribution grid transformers," in Power Line Communications and Its Applications (ISPLC), IEEE International Symposium, 2011.

[5] "Boston Spa Energy Efficiency Trial - Smart Meter Voltage Measurement Performance," Boston Spa, December 2020. Available: https://www.northernpowergrid.com/beet. [Accessed 2022].

[6] "Smart Grid Legislative and Regulatory Proceedings," U.S. Energy Information Administration, November 2011.

[7] D. Alahakoon and X. Yu, "Smart Electricity Meter Data Intelligence for Future Energy Systems: A Survey," IEEE Transactions on Industrial Informatics, vol. 12, no. 1, pp. 425-436, 2016.

[8] B. Pinte, M. Quinlan and K. Reinhard, "Low voltage micro-phasor measurement unit (μPMU)," 2015 IEEE Power and Energy Conference at Illinois (PECI), pp. 1-4, 2015.

[9] M. H. Wen, R. Arghandeh, A. von Meier, K. Poolla and V. O. Li, "Phase identification in distribution networks with micro-synchrophasors," 2015 IEEE Power & Energy Society General Meeting, pp. 1-5, 2015.

[10] K. J. Caird, "Meter phase identification". U. S. Patent US8143879B2, 2008.

[11] S. Bolognani, N. Bof, D. Michelotti, R. Muraro and L. Schenato, "Identification of power distribution network topology via voltage correlation analysis," in 52nd IEEE Conference on Decision and Control, 2013.

[12] A. B. Pengwah, L. Fang, R. Razzaghi and L. L. H. Andrew, "Topology Identification of Radial Distribution Networks Using Smart Meter Data," IEEE Systems Journal, pp. 1-12, 2021.

[13] W. Wang, N. Yu, B. Foggo, J. Davis and J. Li, "Phase Identification in Electric Power Distribution Systems by Clustering of Smart Meter Data," in 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), 2016.

[14] A. D. Ruchkina, M. S. Falaleev and K. V. Mankov, "Phase Identification in Low-Voltage Network Using Smart Meters Data," 2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), pp. 858-860, 2022.

[15] A. J. Victor, W. Adrian and R. Sebastian, "Phase identification and substation detection using data analysis on limited electricity consumption measurements," Electric Power Systems Research, vol. 187, 2020.

[16] Z. S. Hosseini, A. Khodaei and A. Paaso, "Machine Learning-Enabled Distribution Network Phase Identification," IEEE Transactions on Power Systems, vol. 36, no. 2, pp. 842-850, 2021.

[17] A. Hoogsteyn, M. Vanin, A. Koirala and D. Van Hertem, "Low Voltage Customer Phase Identification Methods Based on Smart Meter Data," 2022.

[18] S. J. Pappu, N. Bhatt, R. Pasumarthy and A. Rajeswaran, "Identifying Topology of Low Voltage Distribution Networks Based on Smart Meter Data," IEEE Transactions on Smart Grid, vol. 9, no. 5, pp. 5113-5122, 2018.

[19] V. Arya, D. Seetharam, S. Kalyanaraman, K. Dontas, C. Pavlovski, S. Hoy and J. R. Kalagnanam, "Phase identification in smart grids," 2011 IEEE International Conference on Smart Grid Communications (SmartGridComm), pp. 25-30, 2011.

[20] S. P. Jayadev, A. Rajeswaran, N. P. Bhatt and R. Pasumarthy, "A Novel Approach for Phase Identification in Smart Grids Using Graph Theory and Principal Component Analysis," 2016.

[21] M. Vanin, T. Van Acker, R. D'hulst and D. Van Hertem, "Phase Identification of Distribution System Users Through a MILP Extension of State Estimation," 2022.

[22] "How much electricity is lost in electricity transmission and distribution in the United States?," U.S. Energy Information Administration, [Online]. Available: https://www.eia.gov/tools/faqs/faq.php?id=105. [Accessed 2022].

[23] "SmartMeter Energy Consumption Data in London Households," UK Power Networks, [Online]. Available: https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households. [Accessed 2022].