

PNNL-38910

Grid Topology Discovery Algorithm Evaluation of Suitability for Utility Deployment (CRADA 606) Final Report

June 2026

Shiva Poudel (PNNL)
Andrew Fisher (PNNL)
Aaqib Peerzada (PNNL)
Arun Veeramany (PNNL)
Andrew P. Reiman (PNNL)
Vitalii Dubonos (Survalent)
Peter Jang (Survalent)
Young Ngo (Survalent)

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, **makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical
Information,
P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: <http://www.ntis.gov>

Grid Topology Discovery Algorithm Evaluation of Suitability for Utility Deployment (CRADA 606) Final Report

June 2026

Shiva Poudel (PNNL)
Andrew Fisher (PNNL)
Aaqib Peerzada (PNNL)
Arun Veeramany (PNNL)
Andrew P. Reiman (PNNL)
Vitalii Dubonos (Survalent)
Peter Jang (Survalent)
Young Ngo (Survalent)

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

CRADA Final Report Certification Memo

Memo Completion Date: 6/29/2026

CRADA Number: 606

CRADA Title: Grid Topology Discovery Algorithm Evaluation of Suitability for Utility Deployment

Responsible Technical Contact at DOE Lab (PNNL): Shiva Poudel (PNNL)

Provide a list of publications, conference papers, or other public releases of results, developed under this CRADA:

Poudel S., A.R. Fisher, A.A. Peerzada, A. Veeramany, A.P. Reiman, V. Dubonos, and P. Jang, et al. 2025. "Evaluating the Suitability of a Grid Topology Discovery Algorithm for Utility Deployment." In 2026 IEEE Power and Energy Society General Meeting. PNNL-SA-217966

Provide a detailed list of all subject inventions, to include patent applications, copyrights, and trademarks:

No subject inventions were generated under this CRADA

Executive Summary of CRADA Work

This project, conducted by Pacific Northwest National Laboratory (PNNL), in collaboration with Survalent and Peninsula Light Company (PenLight), focused on evaluating topology discovery algorithms for their suitability in utility deployment. The partnership was instrumental in providing real-world feeder models and operational data from PenLight, as well as Advanced Distribution Management System (ADMS) software and valuable technical input from Survalent, ensuring that the research addressed the practical needs of distribution utilities. Using Milsoft-exported planning distribution system models deployed in OpenDSS format, along with advanced metering infrastructure (AMI) and SCADA data received from the Survalent ADMS, the team developed, validated, and tested a workflow that integrates utility-provided models with measurement data to support accurate topology identification.

The project demonstrated a field-oriented approach by configuring and validating network models, integrating AMI and ADMS data, and conducting extensive simulations to evaluate tool performance under actual real-world conditions. Results showed that optimized sensor placement significantly improves accuracy and robustness of topology discovery, reducing misdetection and enhancing operational awareness. Final observations highlight both the potential and limitations of current algorithms, particularly under sparse measurement scenarios. The report concludes with recommendations for utilities to pursue incremental deployment of advanced sensors, prioritize data validation processes, and integrate topology discovery tools into operations platforms to improve situational awareness.

The report is structured as follows:

Chapter 1.0 introduces the topology identification problem and discusses the motivation behind this work. A brief background of the existing approaches, along with their limitations, is also given. Chapter 2.0 presents the formulation and details the developmental process of the proposed measurement-based topology identification algorithm. The algorithm was developed at PNNL, as part of the research on distribution model selection-based state estimation. Chapter 3.0 and 4.0 discuss the implementation, validation, and demonstration of the topology identification algorithm. This work was performed as part of a Cooperative Research and Development Agreement (CRADA) agreement between PNNL and Survalent, with the overarching objective of demonstrating the topology discovery algorithm in a field test environment.

Key Insights

Some key insights from the report are summarized below.

- This work offers actionable insights for integrating utility datasets and workflows into open-source simulators and optimization frameworks to enhance grid reliability and visibility.
- The results in the report highlight that sensor placement is critical to algorithmic accuracy. Many existing sensors in the distribution systems that are deployed for protection or operational purposes may not always provide the data that is needed for topology discovery.
- A cost-effective sensor deployment strategy that supports multiple objectives is important to ensure system observability.

Acknowledgments

The authors would like to express their sincere gratitude to Peninsula Light Company, for their invaluable collaboration and support throughout this project. We especially appreciate the opportunity to work with real feeder models and real-world sensor data, which significantly enhanced the practical relevance and impact of our research. This material is based upon work supported by the Department of Energy, Office of Electricity, Sensor Technologies and Data Analytics Program.

Acronyms and Abbreviations

AMI	advanced metering infrastructure
API	application programming interface
CRADA	cooperative research and development agreement
GIS	geographic information system
GSE	generalized state estimation
IEEE	Institute of Electrical and Electronics Engineers
MSDR	missed switch detection ratio
NP	nondeterministic polynomial (time)
PenLight	Peninsula Light Company
PNNL	Pacific Northwest National Laboratory
SCADA	supervisory control and data acquisition
SE	state estimation

Contents

CRADA Final Report Certification Memo.....	ii
Executive Summary of CRADA Work.....	iii
Acknowledgments.....	iv
Acronyms and Abbreviations.....	v
Contents	vi
1.0 Introduction to Topology Identification	1
1.1 Challenges for Distribution System Model Discovery	1
1.2 Distribution Operations Platform	1
1.3 Approaches for Topology Identification	2
2.0 Topology Identification Algorithm – Formulation and Development.....	4
2.1 Methodology	4
2.1.1 Model Reduction and Case Reduction.....	5
2.1.2 Model Selection.....	8
2.2 Metrics	8
2.2.1 Missed Switch Detection Ratio (MSDR).....	8
2.2.2 Average Percentile Ranking (APR).....	9
2.2.3 Absolute Objective Difference (AOD).....	9
3.0 Topology Identification Algorithm – Implementation and Validation	10
3.1 Development of Utility Evaluation Plan.....	10
3.2 Configure and Validate Network Model.....	10
3.3 Characterize Tool Sensitivity to Practical Constraints	10
3.4 Tool Finalize and Evaluation	10
3.5 Data Integration Workflow.....	11
3.5.1 Topology and Parameter Validation.....	11
3.5.2 AMI Data Integration.....	12
3.6 Correcting Model and Component Definitions	13
4.0 Topology Identification Algorithm – Field Demonstration	15
4.1 Interfacing with SCADA Host	15
4.2 Demonstration	16
4.3 Simulation Setup.....	16
4.4 Simulation Results	17
4.4.1 Base Case.....	17
4.4.2 Updated Sensor Configuration.....	19
4.5 Performance Analysis	20
5.0 Summary of Research Results	22
6.0 References.....	23

Figures

Figure 1.	Distribution operations and topology discovery.	2
Figure 2.	Proposed graph-theoretic approach to state estimation-based model selection. Different steps in implementing this approach are highlighted within the green box.	4
Figure 3.	Algorithm 1 for model reduction using graph minor.	5
Figure 4.	Connectivity matrix representation for two different switch combinations. Note that “xx” represents a non-zero entry in the connectivity matrix showing a possible path from the source node (“s”) to a particular node. The zero value in <i>Ami</i> represents a disconnected network.	6
Figure 5	Sample distribution circuit to illustrate switches statuses.	7
Figure 6.	Algorithm 2 for case reduction using graph theory.	7
Figure 7.	Workflow for model validation and data integration.	11
Figure 8.	Mapping AMI data to load object.	12
Figure 9.	Model development with appropriate component definitions.	13
Figure 10.	Workflow for interfacing between SCADA host and state estimation algorithm.	15
Figure 11	Representative one-line diagram of feeder with sensor location.	16
Figure 12.	Voltage profile.	17
Figure 13.	Objective function plot for the base case.	18
Figure 14.	Objective function value plot, where identified and actual topologies are different.	19
Figure 15.	Switching combination analyses for different cases.	19
Figure 16.	Objective functions plot for case 3 in Figure 15.	20
Figure 17.	Objective functions plot for case 6 in Figure 15.	20

Tables

Table 1.	Algorithm characteristics for base case.	17
Table 2.	Performance comparison for default and updated sensor configurations.	21

1.0 Introduction to Topology Identification

Advanced distribution management applications such as model-based state estimation, volt/VAR control, distributed energy resource management, outage management are highly contingent on maintaining an accurate grid model in the control center of the distribution system operator (Zamani and Nielsen 2020). The challenge of updating the system topology is complicated by the highly dynamic nature of electric distribution systems, which typically can undergo several reconfigurations in a single day (Li et al. 2018). These reconfigurations are driven by multiple operational objectives, including power equipment maintenance and service activities, restoring service after faults, minimizing technical losses, balancing feeder loading, etc.

Unlike static planning models, the as-operated distribution network often differs from its designed (as-built) configuration due to frequent switching operations, both automatic and manual. As a result, maintaining an accurate and up-to-date representation of the network topology is essential. Without it, system operators risk making operational decisions based on outdated or incorrect models, which can compromise operational safety, reliability, resiliency, and efficiency of the distribution grid. As such, the need to identify time-varying operational topology is a central motivation for this work, which proposes a graph-theoretic approach to model selection.

1.1 Challenges for Distribution System Model Discovery

In transmission systems, the topology identification task is relatively straightforward due to the widespread deployment of telemetered devices, supervisory control and data acquisition (SCADA) systems, and phasor measurement units, which provide rich and high-resolution data for processing topology changes (Srivastava et al. 2022). Unlike transmission systems, which benefit from extensive redundancy in their metering infrastructure (redundancy ratios of 1.7–2.2), distribution networks are sparsely equipped with real measurements. Distribution networks employ a wide variety of switching devices—such as circuit breakers, disconnecter switches, load-break switches, tie switches, sectionalizing switches, fuses, and reclosers—to enable circuit reconfiguration and isolation. Due to uneven levels of automation and communication infrastructure, many of these switching devices cannot transfer their open/closed status in real time to the Distribution Management System (Dalavi et al. 2024).

1.2 Distribution Operations Platform

Figure 1 illustrates the flow of information and control between the grid, devices, and the distribution operations platform, with a particular focus on topology discovery. Sensor and meter data are collected from the grid as a function of ongoing operations and capture the effects of any actions taken by controllable devices. These measurements serve as a critical input to the model enhancement process, where topology discovery—conducted in collaboration with Survalent and PenLight—ensures that the system representation reflects the as-operated network rather than just the planned configuration. The enhanced model is then maintained within the distribution operations platform and supports applications that rely on an accurate, up-to-date understanding of the system for monitoring, situational awareness, and control.

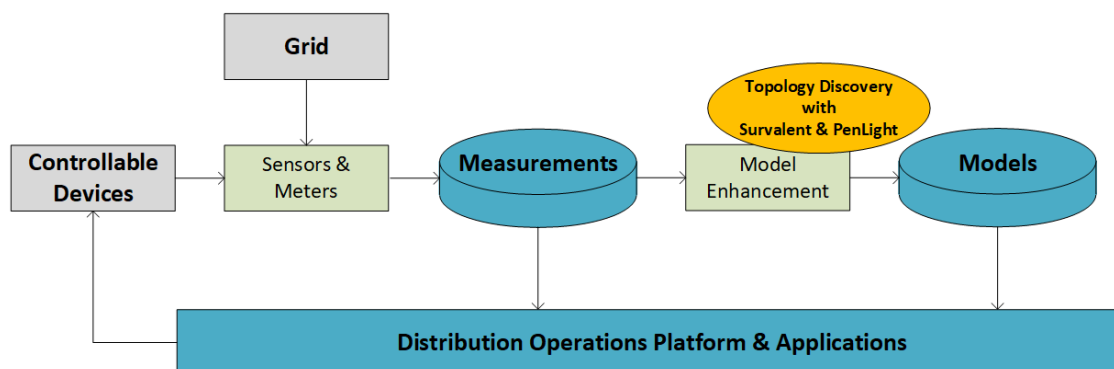


Figure 1. Distribution operations and topology discovery.

1.3 Approaches for Topology Identification

Theoretically, topological identification is a nondeterministic polynomial time (NP)-hard combinatorial problem that involves searching for a feasible tree within a graph that aligns with the measurement data. NP-hard problems that include discrete and continuous measurements are not tractable for emerging power distribution systems with many switches and possible operating topologies.

Different categories for solving the topology identification problem are based on 1) generalized state estimation (GSE), 2) optimization-based methods, and 3) data-driven approaches. While approaches to GSE allow for topology and error identification while simultaneously computing the distribution grid state, there are limitations to these approaches; some rely on forecasted or pseudo-measurements to work correctly (Korres and Manousakis 2012), and some algorithms must solve the GSE for each possible switch topology. A few approaches leverage mixed-integer-based optimization to solve the topology identification problem (Tian et al. 2016; Gandluru et al. 2020); however, mixed-integer problems are NP-hard and computationally complex (Kannan and Monma 1978). Data-driven methods based on regression, expected-maximization, and unsupervised machine learning are used to infer the topology of the grid. While data-driven approaches for GSE are promising, they pose a few implementation challenges. First, they rely on voltage and current measurements at high frequencies from micro phasor measurement units or power flow meters at unrealistic deployment levels. Second, most of these methods estimate topology within data windows, with window lengths on the scale of days or much longer. Similarly, some algorithms assume only one switch change at a time, thus restricting the applicability for a real-world distribution system. While many methods have been proposed to solve this problem, none have found practical use in the electric utility industry. Furthermore, changes in grid structure and function, as well as changes in grid requirements, make the problem more complex than past approaches could handle. Therefore, there is a need for an algorithm that can detect topology by accounting for practical constraints in the distribution grid. To this end, the project team at Pacific Northwest National Laboratory (PNNL) hardened existing approaches to ensure the effective evaluation of algorithms for their suitability in utility deployment.

A two-stage approach for topology identification of power distribution systems is demonstrated in this work. A graph-theory analysis is implemented in the first stage to reduce the sample space of the given combinatorial problem. The system connectivity matrix and system measurements identify a set of candidate topologies. The second stage solves an estimation problem to infer the grid state of the candidate topologies. The model with the least objective value (absolute difference between measurements and grid state) overall is selected as the identified topology for the given set of measurements. Unlike mixed-integer-based formulations,

we adopt a graph-theoretic approach to shrink the sample space of possible switch combinations, thus avoiding an exhaustive search among all possible topologies. For solving the model selection problem, we leverage sparse measurements from the lines and buses where the distribution network is typically instrumented, such as reclosers. Furthermore, the algorithm implemented in this work is not restricted to the degree of topology change as determined by the number of switch statuses that have changed from the last known topology.

2.0 Topology Identification Algorithm – Formulation and Development

The proposed methodology for topology identification is summarized in Figure 2 and is based on the prior work done at PNNL (Poudel et al. 2024). PNNL conducted research on the applicability of graph-theoretical approaches for topology identification. The measurement-based topology identification algorithm leverages measurements from the distribution network to infer the current operational topology from a given planning model. The planning model of a distribution grid consists of power delivery elements such as lines, transformers, normally closed sectionalizing switches, and normally open tie switches. These components together form a cyclic graph that can be operated in numerous topologies by opening and closing switches within the planning model. Thus, theoretically, it is a combinatorial problem that searches for a feasible tree or sub-trees with a graph that satisfies given measurements. In this work, the topological identification problem is based on two types of measurements: 1) sensor measurements (line flows and/or nodal voltages) and 2) pseudo-measurements in the form of measured or otherwise forecasted nodal injections at all buses.

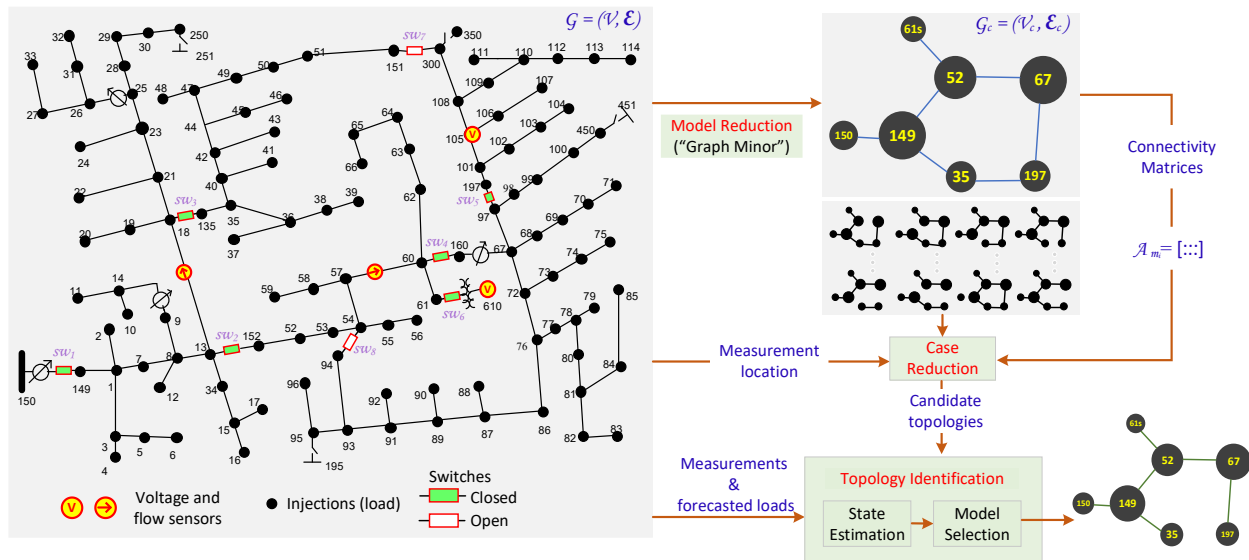


Figure 2. Proposed graph-theoretic approach to state estimation-based model selection. Different steps in implementing this approach are highlighted within the green box.

2.1 Methodology

The process is divided into two stages: 1) graph-based candidate topology enumeration and 2) state estimation (SE)-based model selection (SEMS) on the candidate topologies. The first stage is carried out in two steps: a) model reduction and b) case reduction. The model reduction step is done offline using a static planning model of the distribution network. The case reduction step and SEMS stage are implemented in real time using graph-theory analysis and optimization, respectively. Note that these processes could be started in parallel to improve performance.

Reducing candidate topologies before invoking the model selection stage can greatly improve the computational efficiency of the topology identification process. Since there are several possible switching combinations for a traditional SE to calculate a prediction from, it is also more likely that an inaccurate topology prediction would occur over the increased sample space

compared to the reduced sample space. To this end, the proposed graph-theoretic approach can improve accuracy and computational performance.

2.1.1 Model Reduction and Case Reduction

A distribution planning model comprises multiple feeders and can be operated in numerous operational topologies by choosing different switch combinations, as shown in Figure 2. The planning model is represented by an undirected graph $G = (V, E)$; the set V , with cardinality $|V| = n$ representing buses, and the set E , with cardinality $|E| = m$ representing edges associated with the electrical lines, transformers, and switching devices. For the topology identification problem, it is the position (ON/OFF) of switching devices that needs to be discovered. In this work, we implement a graph-theory-based approach for case reduction to eliminate several topologies before invoking the SE algorithm. To implement the case reduction technique, only switch-level information is required; thus, the graph-minor technique is implemented to consolidate all non-switch edges (Figure 3).

Algorithm 1 Model reduction using graph minor

Given: $\mathcal{G} = (\mathcal{V}, \mathcal{E}), \mathcal{E}_s$
 $\mathcal{F} \leftarrow \text{TRUE}$
while $\mathcal{F}$ **do**
 $\mathcal{F} \leftarrow \text{FALSE}$
 for e **in** $\mathcal{E}$ **do**
 try
 if $e \notin \mathcal{E}_s$ **then**
 $\mathcal{G} \leftarrow \text{contracted_edge}(\mathcal{G}, e)$
 $\mathcal{F} \leftarrow \text{TRUE}$
 catch Continue
Output: $\mathcal{G}_c = (\mathcal{V}_c, \mathcal{E}_c)$

Figure 3. Algorithm 1 for model reduction using graph minor.

Given a graph G and a set of switches E_s , the graph minor algorithm (Demaine et al. 2005) will contract all the edges except the ones with switches, preserving the number of collapsed nodes as weights of the remaining nodes. An edge contraction is an operation that removes an edge from a graph while simultaneously merging the two vertices it connects. Model reduction using graph minor is accomplished using Algorithm 1 as implemented by Python's "networkx" package (Hagberg et al. 2008).

The reduced model preserves the information necessary for the case reduction procedure. Figure 2 shows how the original IEEE 123-bus feeder is reduced into a 7-bus, 7-edge system, where the size of the bus is proportional to the number of contracted edges around it. For example, the edge connecting buses 149 and 1 is not a switch, so it collapsed into bus 149. The output is a new graph object G_c with endpoints of non-switchable edges ($e \notin E_s$) identified in a single node.

There are 2^{n_s} possible combinations of switches for a distribution grid with n_s switches. This sample space is too large to efficiently identify the correct topology through enumeration. In this work, we implement a graph-theory-based case reduction technique to avoid massive enumeration.

Let $M = \{m_1, m_2, m_3, \dots, m_{|M|}\}$ represent the set of possible switch combinations with cardinality $M = 2^{n_s}$. For each switch combination, a connectivity matrix (A_{m_i}) is computed and stored by

summing powers of the adjacency matrix between one and the length of the longest path in G_c . In the worst case, when there exists a path that takes every intermediate node, the longest path in the network equals $V_c - 1$. The connectivity matrix for each switching combination is computed using Equation (1) and stored in a set A as shown in Equation (2).

$$A_{m_i} = A_{m_i}^1 + A_{m_i}^2 + A_{m_i}^3 + \dots + A_{m_i}^r \quad \forall m_i \in M \quad (1)$$

where $r = |v| - 1$

$$A = \{A_{m_1}, A_{m_2}, A_{m_3}, \dots, A_{m_{|M|}}\} \quad (2)$$

Even though calculating the connectivity matrix for each possible topology is a computationally complex process, it does not impact the algorithm's real-time performance because it is performed during pre-processing. An example of a connectivity matrix is shown in Figure 4. Two switch combinations, T-1 and T-2, representing two topologies, are selected for a given network. The topology T-1 is a connected network, and, hence, every entry of A_{m_i} is non-zero. In contrast, in the topology T-2, the source node is only connected to node "1," and the rest of the network is disconnected, thus giving several zero entries in the A_{m_i} .

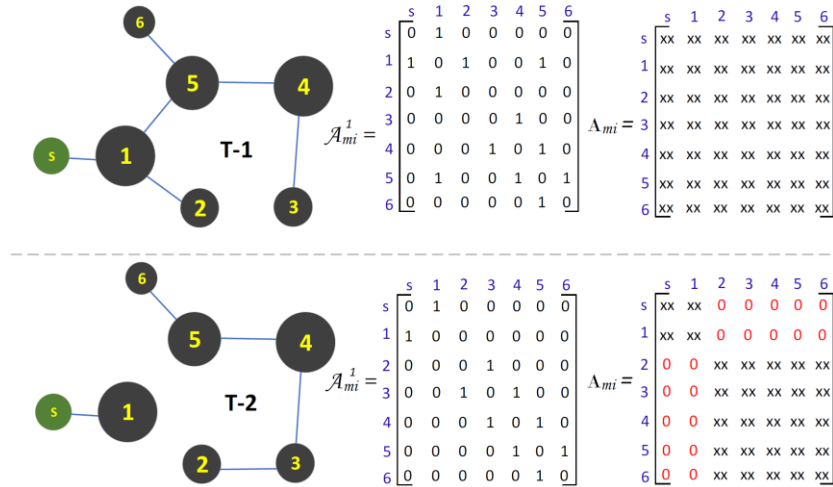


Figure 4. Connectivity matrix representation for two different switch combinations. Note that "xx" represents a non-zero entry in the connectivity matrix showing a possible path from the source node ("s") to a particular node. The zero value in A_{m_i} represents a disconnected network.

We first identify and lock the status of switches whose open position would isolate a measurement location. This step is justified because if a sensor on a line or node reports a measurement, there must be a continuous path to that sensor from the source. Accordingly, the nearest switch must remain closed whenever it serves as the only possible connection to that sensor. For example, in Figure 5, if the sensor recorded a measurement, then it is confirmed that switch s_3 should be in CLOSED position. Therefore, this switch can be removed from the set of unknown switches.

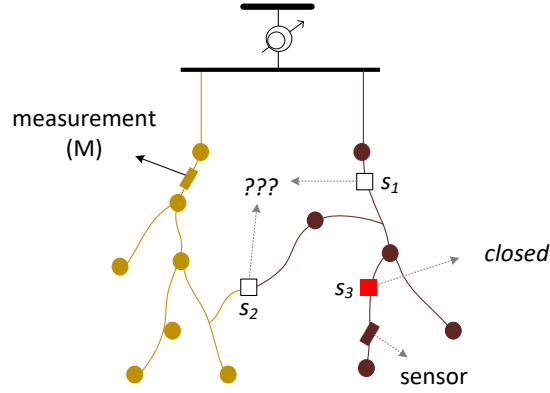


Figure 5 Sample distribution circuit to illustrate switches statuses

Similarly, the sensor status (energized or de-energized) is used to disregard several switch combinations that are invalid. For example, in Figure 5, if the sensor recorded a measurement, then it is confirmed that either of switch s_1 or s_2 must be closed. In other words, any switch combination that has s_1 and s_2 as OPEN is an invalid combination for this example. A proper mathematical approach to handle this is explained next.

Each connectivity matrix ($a \in A$) is passed through a processor that checks if a path exists from a source node to an energized sensor. Suppose S represents a set of sensors reporting measurements (i.e., energized). For a given energized sensor (s), the corresponding entry of a connectivity matrix is checked for a valid path $a[s]$. If no path exists from the source node to one of the sensor locations contained in S , then that combination is not included in the candidate topology list (Ω). If a connectivity matrix has non-zero entries corresponding to all energized sensors S , the switch combination represented by the connectivity matrix is included in the candidate topology list. The overall procedure is summarized in Algorithm 2 (Figure 6). The final output of this step is a set of candidate topologies Ω to be fed into model selection step.

Algorithm 2 Case reduction using graph theory

Given: $\mathcal{G}_c = (\mathcal{V}_c, \mathcal{E}_c)$, $\mathcal{S}$, $\mathcal{M}$, $\mathbf{A}$

$\Omega \leftarrow []$

for $k \leftarrow 1$ **to** $|\mathcal{M}|$ **do**

$\mathcal{F} \leftarrow 1$

$\mathbf{a} \leftarrow \mathbf{A}[k]$

for s **in** $\mathcal{S}$ **do**

 // find the entry in the connectivity matrix

$path \leftarrow \mathbf{a}[s]$

if $path = 0$ **then**

 // if no path exists from source to s , the given switch combination, $\mathcal{M}[k]$ is not valid

$\mathcal{F} \leftarrow 0$

break

if $\mathcal{F} = 1$ **then**

 // k^{th} switch combination in $\mathcal{M}$ is a candidate for SE

$\Omega \leftarrow [\Omega, \mathcal{M}[k]]$

Output: Ω

Figure 6. Algorithm 2 for case reduction using graph theory.

2.1.2 Model Selection

Once candidate topologies are identified, system states are calculated based on given measurements and network model. The objective function (or fitness function) denoted by O is calculated for each candidate topology where objective function is defined as the difference between the measurements and system state. Mathematically,

$$O_{\omega} = \sum_{i \in B} w_i (h_i(S) - z_i)^2 + \sum_{j \in V_c} w_j (h_j(v) - z_j)^2 \quad \forall \omega \in \Omega \quad (3)$$

where B is the set of sensors measuring line flow; V_c is the set of sensors recording node voltages (such as AMI); S represents the line flow; and v represents nodal voltage. Since these measurements have different errors associated with them, corresponding weights are applied to bias the objective toward a specific measurement. First, a topology is extracted from the candidate topology list (Ω) using the original planning model (G). A system admittance matrix for a candidate topology is then constructed leveraging the switching combinations M corresponding to the candidate topology. State estimation is solved for the candidate topology to determine the system states. In this work, we used a deterministic estimation, where a load flow analysis is conducted to evaluate the theoretical value of line flow and node voltages that are consistent with the observed measurements z_i and the values are plugged into Equation (3) to evaluate the objective function. The candidate topologies are ranked according to the optimal objective value, and the one with the least objective is the identified topology.

2.2 Metrics

The effectiveness of the proposed approach is thoroughly tested for different sets of unknown switches and varying sensor configurations. The metrics defined are capable of providing a measure of how far, on average, the identified topology is from the accurate topology if it is mis-detected.

2.2.1 Missed Switch Detection Ratio (MSDR)

Note that an incorrect estimation of more than one switch renders the given topology incorrectly identified. Thus, we define a metric, missed switch detection ratio (MSDR), that quantifies the algorithm performance in correctly estimating switch statuses. Mathematically,

$$MSDR = \frac{S_c}{S \times M}$$

where S_c is the number of switches identified incorrectly, S is the total number of switches, and M is the total number of trials.

2.2.2 Average Percentile Ranking (APR)

For a given topology, a percentile ranking (PR_i) of the accurate topology is calculated using the objective value at each trial. If $\mathbf{O} = \{O_1, O_2, \dots, O_N\}$ is the set of objective values of possible N topologies, the percentile rank of the correct topology with objective value O_i is given by:

$$PR_i = \left(1 - \frac{L}{N}\right) \times 100$$

where L is the number of values in $\mathbf{O}$ that is lower than O_i and N is the total number of possible topologies. Once PR_i is obtained for each trial, an average percentile ranking is calculated using:

$$APR = (\sum_{i=1}^M PR_i) / M.$$

2.2.3 Absolute Objective Difference (AOD)

During a topology misdetection, a model with the smallest objective function value is not the correct topology. In such a case, there is a difference between the objective value of a correct topology and the topology with the least objective value. This metric quantifies the absolute difference between the residuals of the actual topology and the identified topology. Mathematically, AOD is given as,

$$AOD = |\text{Min}(\mathbf{O}) - O_i|$$

where O_i is the objective value of the correct topology.

3.0 Topology Identification Algorithm – Implementation and Validation

The project team at PNNL established a Cooperative Research and Development Agreement (CRADA) with Survalent to implement and validate the topology identification algorithm, developed as part of the earlier research at PNNL. The goal of the CRADA was the validation of the model identification algorithm, in a controlled field test environment. Under the CRADA, PNNL provided a tool package, including research-grade code, and documentation to Survalent in a format that facilitates incorporation of the functionality into a commercial solution. Models and data for demonstration were provided by Peninsula Light Company (PenLight) through their relationship with Survalent.

3.1 Development of Utility Evaluation Plan

The team at PNNL began planning for the utility evaluation by identifying the minimum network model and sensor/AMI requirements necessary to support the deployment of the topology discovery algorithm. In parallel, PNNL and Survalent, in collaboration with PenLight, defined the network scope for evaluation, determining the portions of the system that would be included in the study. The parties also identified the data to be shared among collaborators, agreed on formats for the available data, and established data management and sharing protocols, including the provision of base electric network model information. As a part of this process, PNNL received the network model representing multiple feeders from two distribution systems. AMI measurements and SCADA data were exported and provided in CSV file format.

3.2 Configure and Validate Network Model

Survalent provided the data required for model creation, which PNNL used to develop an electric network model at the level of detail specified in the evaluation plan. This model was validated in collaboration with Survalent. For this evaluation, a Milsoft-exported OpenDSS model for the area of interest was provided. PNNL performed load flow analyses using the integrated AMI data to verify the model's validity, confirming that the simulated voltage levels and line flows aligned with expected operational values.

3.3 Characterize Tool Sensitivity to Practical Constraints

The team then evaluated the topology discovery tool's sensitivity to practical operational constraints. Sensitivity studies were performed by varying the number of unknown switches and sensor locations, and the results were carefully reviewed. The analysis revealed that the algorithm could not handle all possible configurations of unknown switches. In response, the PNNL team enhanced the graph-theory component of the algorithm (lock status of a few unknown switches, as described in 2.1.1), enabling it to handle at least 20 unknown switches. Additionally, it was observed that some AMI data was missing, and it was not fully time-synchronized with SCADA measurements. To address this, the team implemented a data-filling approach using a moving average, ensuring proper synchronization between AMI and SCADA data.

3.4 Tool Finalize and Evaluation

Because a direct comparative analysis in the field was impractical, a simulated environment was created in OpenDSS to emulate the existing sensor configuration. Multiple simulations were conducted within this environment to characterize the performance of the topology discovery

tool. Additionally, the team used a new sensor configuration that could improve the tool's performance compared to the default setup, using the relevant metric.

3.5 Data Integration Workflow

The proposed workflow for modeling and data integration is illustrated in Figure 7. The process began with extracting the distribution network model from Geographic Information System (GIS) data. PenLight utilizes Milsoft as its primary distribution system simulation tool, where the network model was maintained in Milsoft's proprietary format. Milsoft offers an export feature for OpenDSS, enabling the conversion of the network model into a set of files compatible with the OpenDSS simulation environment. These exported OpenDSS files, provided by PenLight, served as the foundation for tool development used in this study.

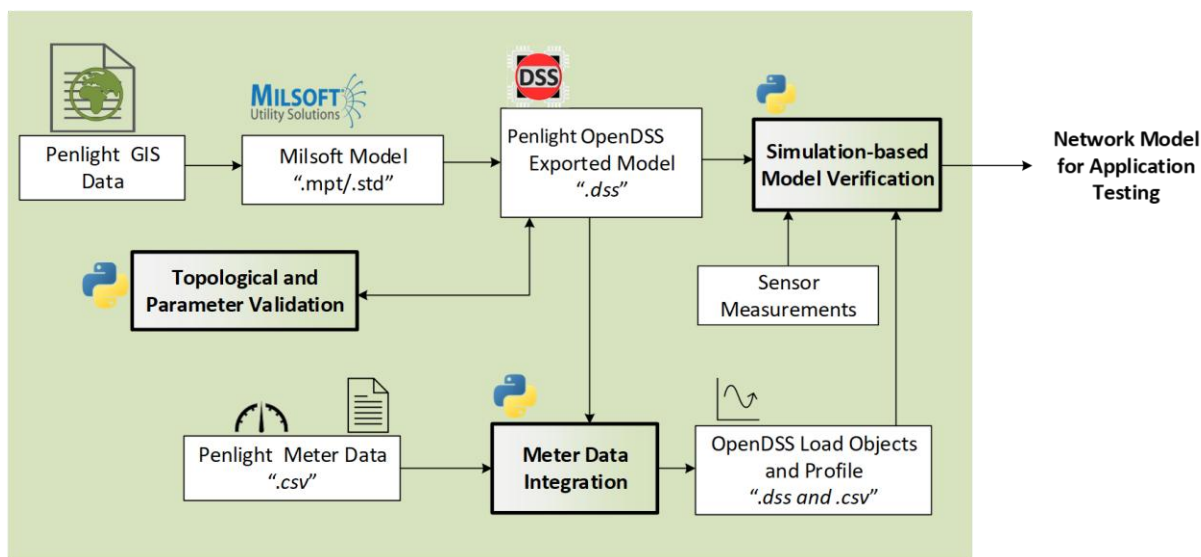


Figure 7. Workflow for model validation and data integration.

In the remainder of this section, we describe the process used to validate the topology of OpenDSS models exported from Milsoft, eliminate gross modeling errors, and integrate AMI data into the network model, enabling load flow analysis.

3.5.1 Topology and Parameter Validation

This function thoroughly checks topology, base voltages, and component parameters to ensure the operational accuracy of the network representation.

Topology validation involves both visual inspection and graph-theory-based analysis. Visual inspection helps identify obvious modeling errors such as isolated components, phase inconsistencies, or disconnected buses. Graph-theory analysis is used to detect isolated networks or unintended partitions within the system systematically. Where missing connectivity is identified, discussions with utility partners are initiated to update the model with the appropriate components and ensure a fully connected network representation. Additionally, the radiality of the system is tested by confirming the network's as-operated model with utility operators, ensuring that the model accurately reflects the operational configuration of the real-world distribution system.

Nominal voltage validation focuses on confirming that the primary and secondary voltages of transformers, as well as the nominal voltages assigned to other components—such as

capacitors, loads, distributed energy resources, and voltage regulators—are correctly specified. Special attention is given to the control modes of regulators to ensure consistency with field operation.

Component parameter validation includes verifying that critical physical and electrical parameters are reasonable and correctly assigned. This includes verifying that line impedance parameters, such as insulation values and geometric mean radius, transformer reactance values, and line lengths, are non-negative and non-zero where applicable, and that their units are consistent and accurate. Cross-verification is performed against engineering standards or utility-provided specifications to ensure fidelity.

This validation process ensures that the distribution network model is configured correctly and is ready to conduct a load flow study.

3.5.2 AMI Data Integration

Integrating the AMI measurements from individual sites requires three steps. First, the data must be mapped to the correct transformer to which the meter would be connected. Second, the site power must be applied to one of the potential feeders connected to the secondary side of the transformer. Finally, the integrated sensor data must be verified to be a reasonable measurement.

Figure 8 illustrates the process of integrating AMI data from a meter database into an OpenDSS distribution network model. This integration begins with the collection of time-series data from individual customer meters. Although this AMI data may not be available in real time, it can typically be exported in a structured format, such as a CSV file. To associate this meter data with the physical distribution network, ADMS model is leveraged. This allows us to map each asset ID (meter ID) to its corresponding service transformer as well as line segment the transformer is connected to, enabling us to determine which transformer each customer meter is connected to. As multiple meters can be associated with a single service transformer, each meter is linked to a specific triplex node within the network model. This creates a direct connection between meter readings and precise locations in the network topology.

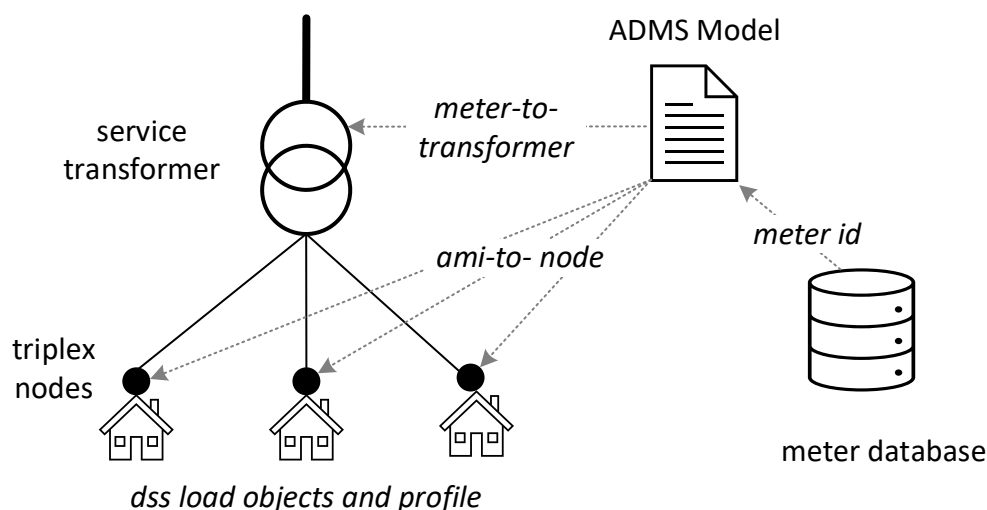


Figure 8. Mapping AMI data to load object.

Using this mapping, each AMI meter is associated with a corresponding triplex node in the OpenDSS model, which represents the customer's point of connection on the secondary side of

the transformer. A load object is then created at each of these triplex nodes within the OpenDSS model. The AMI data for each customer is used to define a time-series load profile for these load objects. These profiles reflect the actual energy usage patterns of individual customers and serve as dynamic inputs for power flow simulations.

3.6 Correcting Model and Component Definitions

A major objective of the model development and the correction process (Figure 9) is to identify the switch elements and the associated switch properties. The process begins by parsing the switch data and identifying all the switch types (breakers, reclosers, fuses, reactors). The data parsing is performed in Python 3.12, which uses a “py-dss-interface” package to interface with the OpenDSS solver engine. The properties of the switch objects are retrieved from the OpenDSS models using the Application Programming Interface (API) calls to connect with the compiled model of the circuit.

In the second step of the switch mapping process, the “fuse” objects are identified and removed. The switches that are identified are added to the OpenDSS model. With these changes, a new OpenDSS switch file (Switch.DSS) is created, and the corresponding “Reactor.DSS” file that originally contained all the switch types is modified. Appropriate modifications are made to the “SwtControl.DSS” files that contain the control logic of the reactors in the circuit. The process is repeated for the second distribution system, and a new switch file (with all the switches except fuses) is created, along with new reactor and switch control files.

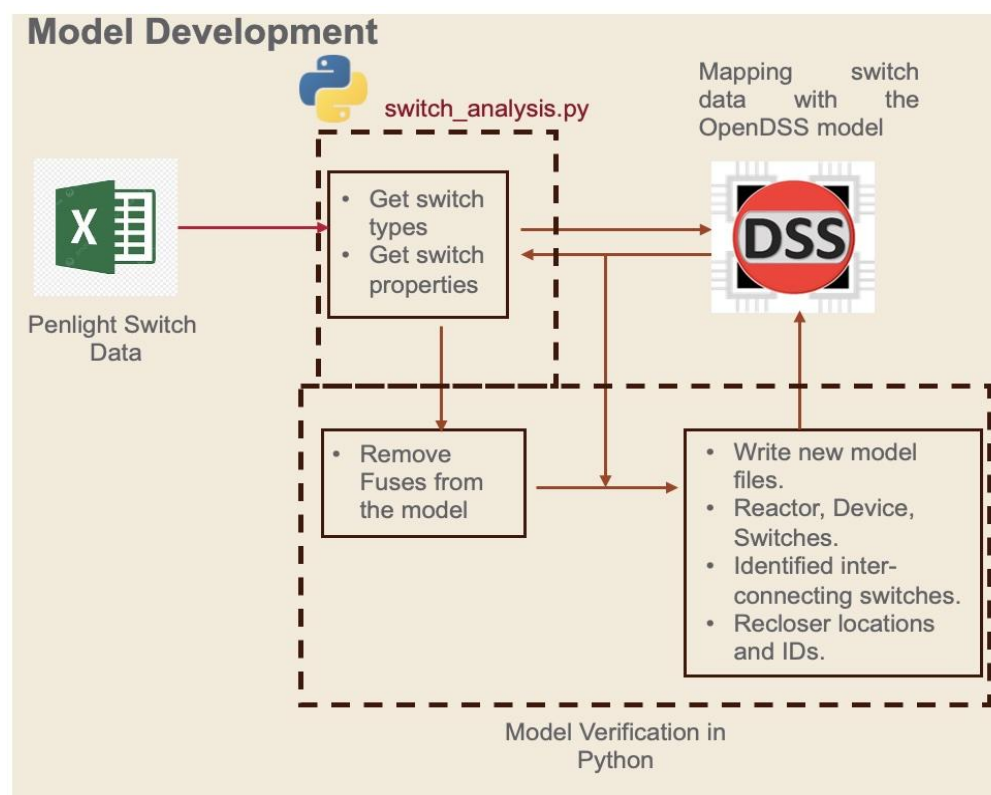


Figure 9. Model development with appropriate component definitions.

It is important to note that the switch analysis uncovered some fictitious buses that were found in both circuits with random connections that were inconsistent with the design of typical radial and meshed distribution systems. Such buses and the associated line connections were deleted

to preserve the original layout of the circuits. It is our understanding that these buses with inconsistent connections were added to the circuit during the model conversion process from Milsoft's Windmill platform to OpenDSS format. Such conversions are known to introduce errors and necessitate a model cleanup process that checks the viability of the model to produce acceptable power flow results.

A second major objective of the model development and correction process is to identify the interconnecting switches that serve to interconnect two different feeders of the same distribution system. This was accomplished by tracing the connectivity paths from the candidate interconnecting switches to the substations.

A series of power flow programs was run, each with a different potential combination of the interconnecting switches with associated switch states. It was found that the first distribution circuit consists of four interconnecting switches that can be used to transfer load between feeders (reconfigurations) in the event a feeder is disconnected from the substation. A similar analysis was done for the second distribution system, which hosts two interconnecting switches. The normal state of the interconnecting switches is set to open in the OpenDSS model. This is done to create separate radial feeders for the model discovery analysis.

4.0 Topology Identification Algorithm – Field Demonstration

The application developed and later found not suitable is about interfacing with the Survalent ADMS system to obtain and process telemetry data from various electrical components. The system works by connecting to a SCADA host, which is based on the Network Data Acquisition (NDA) API framework. Through this connection, the system can retrieve and analyze essential data points like voltage, current, and status information for specific electrical elements. An application was developed to interface with the Survalent ADMS system for obtaining and processing telemetry data from various electrical components. It operated by connecting to a SCADA host through the Network Data Acquisition (NDA) API framework, enabling the retrieval and analysis of key data points such as voltage, current, and status information for specific elements in the network. While a dedicated application itself was not required to demonstrate the tool's suitability, the effort helped uncover the actual mechanisms required for potential integration between research environments and production-grade ADMS systems.

4.1 Interfacing with SCADA Host

The workflow for interfacing the SCADA subsystem of the Survalent ADMS host with the state estimation algorithm is shown in Figure 10 and discussed further in the rest of this section.

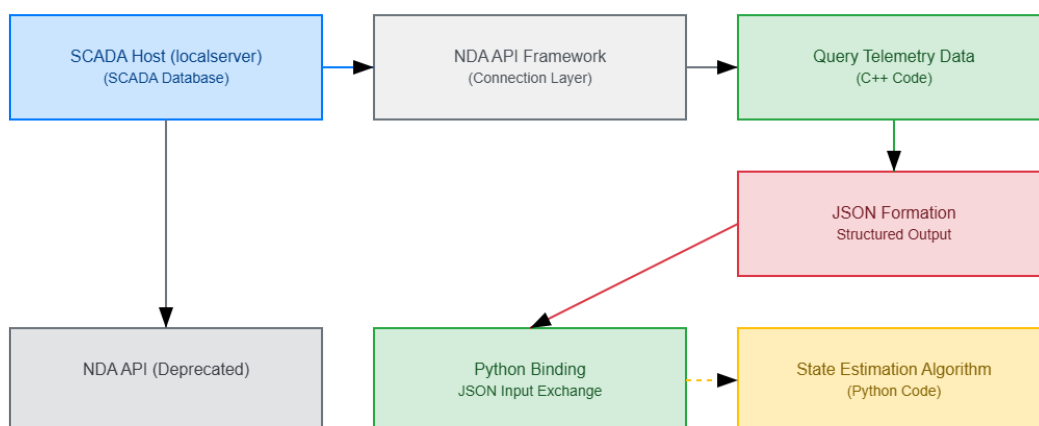


Figure 10. Workflow for interfacing between SCADA host and state estimation algorithm

The system enables communication with the Survalent ADMS environment hosted on a local server (“localhost”) and utilizes the element names (reclosers, switches) given as input to identify and retrieve corresponding telemetry data. Specifically, it constructs derivative point identifiers for key metrics like amperage (current), voltage, and operational status for breaker switches. These metrics are then collected from the ADMS system, which provides the real-time operational information about the actual state of the electrical infrastructure for the state observation algorithms. A key feature of this system is its ability to interface the Survalent ADMS with the Topology Discovery tool through a JSON-based data exchange, effectively replicating real-world conditions within the laboratory environment. Each element is associated with its corresponding status, voltage, and current values, which are collectively formatted to be easily interpretable by downstream systems or users. The structured data could potentially serve in monitoring, diagnostics, and reporting tasks aimed at ensuring the reliability and efficiency of the electrical network.

Upon completion of data retrieval tasks, the application performs cleanup operations to release resources and maintain optimal functionality. This includes disconnecting from the ADMS server and cleaning up its environment. The functionality of this system is centered around providing a seamless interface for acquiring, processing, and structuring telemetry data from critical electrical infrastructure elements. The system enables real-time monitoring and delivers enhanced operational insights to stakeholders responsible for managing electrical systems and maintaining network stability.

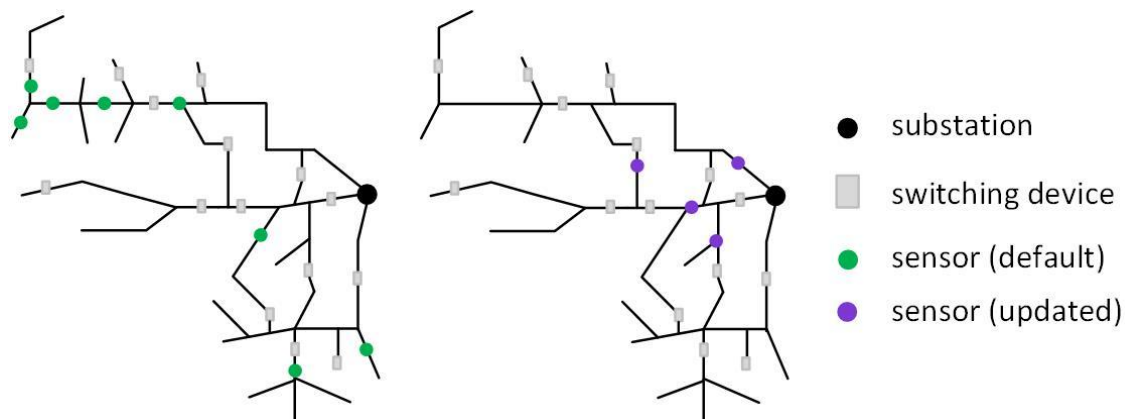


Figure 11 Representative one-line diagram of feeder with sensor location

4.2 Demonstration

The topology discovery tool is demonstrated using one of the distribution systems and associated feeders. Three feeders are selected for the demonstration purpose. AMI meters associated with these feeders were filtered from the database to ensure that only relevant devices were included in the analysis. Figure 11 shows the one-line diagram that closely represents the feeder topology, including switching devices and sensor locations. While the number of switching devices and sensors is not exact, their placement and density are shown to approximately reflect those in the utility's model.

4.3 Simulation Setup

The AMI meter readings and SCADA measurements were aligned by selecting a timestamp at which most AMI data were available. Since AMI readings were recorded at 1-minute intervals, any missing values were filled by averaging the readings immediately before and after the target timestamp. From this process, the timestamp 2025-06-25 17:30 was chosen, where SCADA data were available and approximately 95 percent of AMI readings were present. For the remaining 5 percent, interpolation using adjacent timestamps was applied to create a complete AMI dataset.

With these AMI meters integrated as load objects, a load flow analysis was performed to confirm power flow convergence and verify that the solution was reasonable. Figure 12 shows the voltage profile of the feeder, and it is observed that the per unit voltages at all the nodes are within the bounds. At this snapshot, the loading of the system was about 5 megawatts.

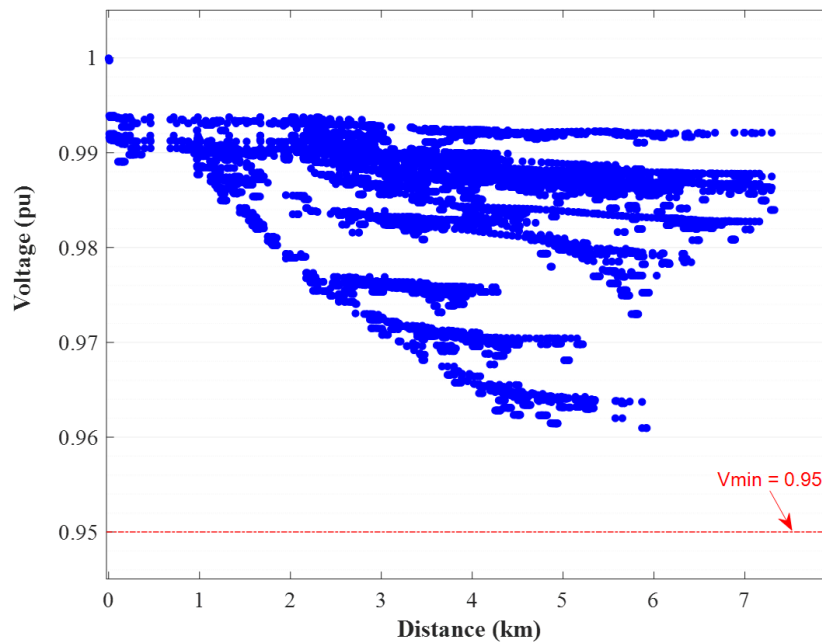


Figure 12. Voltage profile.

For the simulation purpose, a set of 20 different unknown switches were identified. This set is chosen arbitrarily for different trials to ensure that the tool can identify the status of diverse sets of unknown switches.

4.4 Simulation Results

In this section, we discuss the simulation results from two different studies. First, we identify the topology of the system for the given measurements collected from SCADA at the chosen timestamp. Next, we create a hypothetical sensor configuration to quantify the performance of the tool and compare against the base sensor configuration.

4.4.1 Base Case

In this case, we selected a set of 20 unknown switches, and with the given sensor configuration (telemetered reclosers in PenLight feeders and AMI readings), we were able to deduce 7 switches from the list of unknown. This brings the total number of unknowns to 13. Now, out of 8,192 switching combinations, the case reduction stage provided with only 256 valid combinations, as shown in Table 1. The objective function value was computed for these 256 candidate topologies.

Table 1. Algorithm characteristics for base case.

Parameters	Value
Total Unknown Switches	20
Switches Deduced	7
Remaining Unknown Switches	$20 - 7 = 13$
Total Combinations	$2^{13} = 8192$
Total Valid Combinations	256

Figure 13 shows the plot for the objective functions for all 256 valid combinations. The minimum objective function was observed for 39th index, and that switching configuration represents the most likely topology. It is also observed that the objective function values for different candidate topologies are close to each other. This means that the identified topology might not always represent the actual topology, as measurements have errors. Since these measurements were obtained for a particular time, it was not practical to validate this output with the actual topology in the field as the system's topology is not known or recorded for that time.

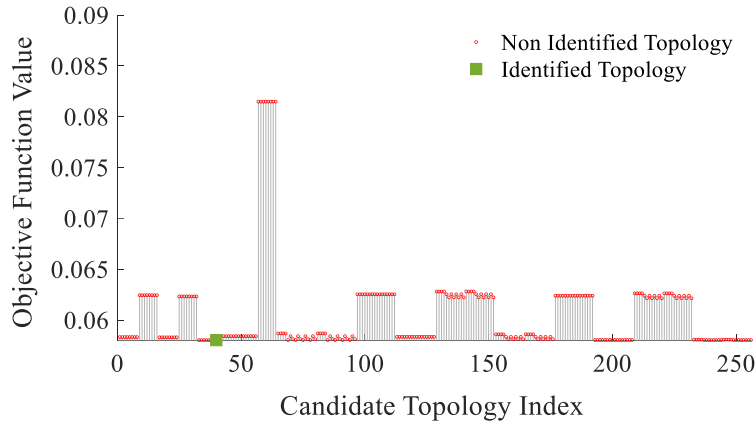


Figure 13. Objective function plot for the base case.

To test a known scenario, we created an operational topology in OpenDSS and generated measurements using the base sensor configuration. With this, we would be able to compare the identified topology with the actual topology. Figure 14 shows such a case, where the identified topology and actual topology were different. As noted earlier, when measurements cannot clearly differentiate between topologies, incorrect identification may occur. Because errors are inherent in measurements and some topologies are very similar—meaning sensor readings are minimally affected by different switching configurations—misdetectors are likely. This underscores the need for optimal sensor placement to improve the accuracy of the topology discovery tool. The minimum number of measurements required to guarantee topological identifiability (assuming error-free measurements) is based on spanning tree and fundamental cycle (Sevlian and Rajagopal 2015). The condition specifies that the measurement set must be chosen such that every possible spanning tree of the network differs from the others by at least one flow measurement, ensuring that each tree can be uniquely identified. Consequently, this implies that at least one flow measurement is needed along each fundamental cycle (Gandluru et al. 2020).

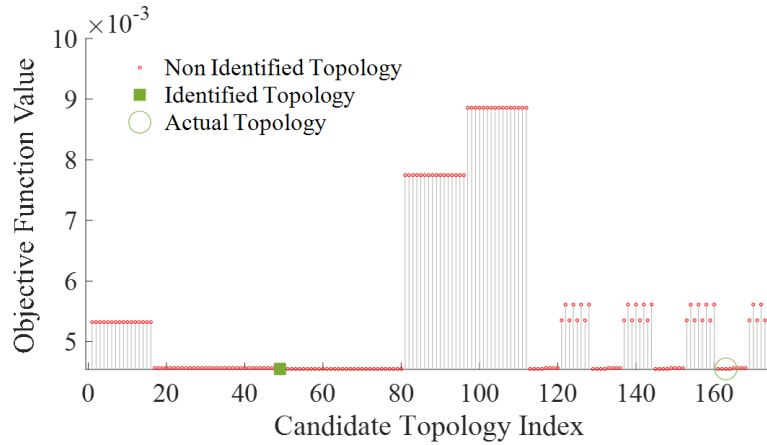


Figure 14. Objective function value plot, where identified and actual topologies are different.

4.4.2 Updated Sensor Configuration

In this section, we provide simulation results by updating the sensor placement. We placed sensors in various locations to ensure that the configuration of candidate topologies would affect the sensor reading (see Figure 11). It should be noted that we did not use an optimal sensor placement approach; instead, we applied a heuristic method to position sensors so that changes in topology would affect their readings. Flow readings were extracted from these sensor locations to create input for the tool, like in the previous case.

Different cases were created, where each case represents a different set of unknown switches. The number of unknown switches is fixed to 20. Figure 15 shows the plot for different cases where the performance of the first stage of the algorithm is highlighted. In each case, the solid blue curve represents the number of unknown switches that were deduced during the first stage, while the solid red curve indicates the number of valid switch configurations retained for further analysis. For example, in Case 3, 8 unknown switches were deduced, and 12 unknown switches were considered for generating candidate topologies. Out of the 4,096 total possible combinations (illustrated by the dashed red curve), only 176 valid combinations were retained for the computation of objective functions (see solid red curve).

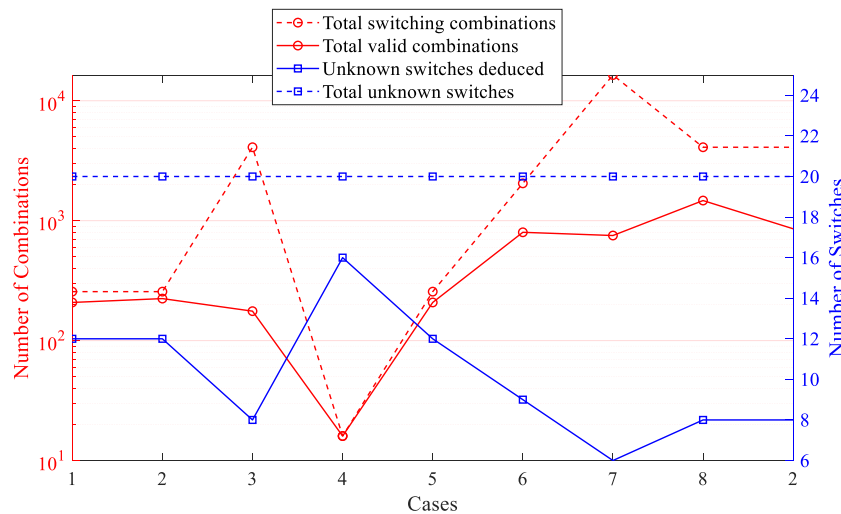


Figure 15. Switching combination analyses for different cases.

Figure 16 shows the objective functions plot for all 176 valid combinations for case 3. It is observed that the identified topology and actual topology were the same, thus indicating that the topology is discovered accurately.

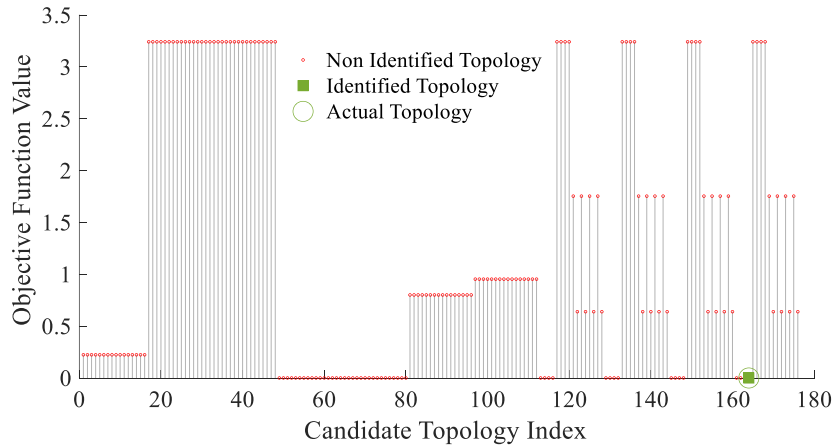


Figure 16. Objective functions plot for case 3 in Figure 15.

On the other hand, it was observed that the identified topology and actual topology were different for case 6, as shown in Figure 17. This indicates that some switches were identified incorrectly when compared to the actual topology. To investigate the cause of these discrepancies, a root cause analysis was performed. For the incorrectly identified switch, the observed flow was zero since no loads were connected downstream. As a result, the objective function was largely unaffected by whether the switch was open or closed. Consequently, the algorithm classified the switch as open, even though it was closed in the actual topology. Such misidentifications are possible but do not affect the intended application, as the misclassified switch has no operational significance.

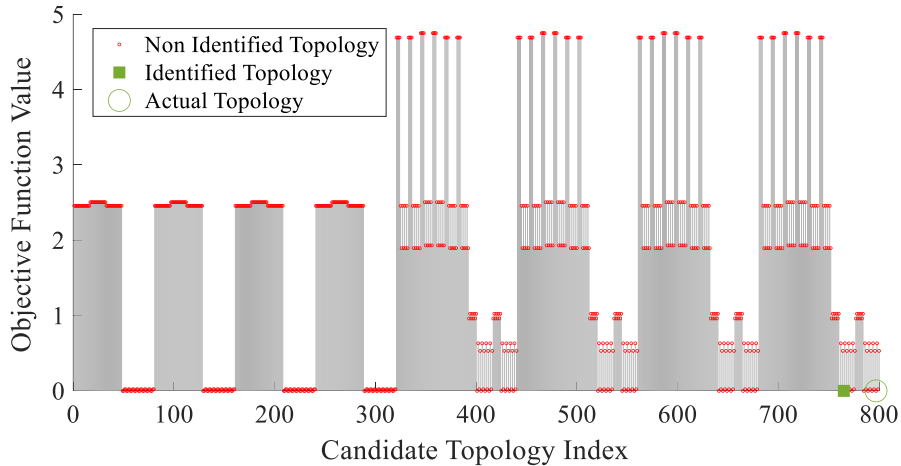


Figure 17. Objective functions plot for case 6 in Figure 15.

4.5 Performance Analysis

Three metrics described in 2.2 were evaluated for default sensor configuration and updated sensor configuration (Table 2). The purpose of these metrics is to evaluate the performance of the topology discovery tool, particularly in scenarios where the correct topology is not explicitly

identified. Even when the accurate topology is not determined, this performance analysis allows operators to examine the set of most likely topologies and understand the operational differences between the actual topology and these identified candidates.

Table 2. Performance comparison for default and updated sensor configurations.

Metrics	Default Sensor Configuration	Updated Sensor Configuration
Percentile Ranking	61.41	10.21
MSDR	0.26	0.12
AOD	0.000298	3.27e-7

For instance, the percentile ranking metric quantifies how many topologies exhibit an objective function value smaller than that of the actual topology. This statistical insight assists operators in narrowing down a list of the most plausible topologies. Similarly, the MSDR metric represents the proportion of switches that were incorrectly identified. The MSDR improves with an optimized sensor configuration, where optimal sensor placement minimizes the likelihood of misdetection. Finally, a lower AOD metric indicates that the identified topology is closely aligned with the actual topology in terms of objective function values. This similarity suggests that the identified topology has comparable operational characteristics to the true topology and does not significantly affect grid states.

5.0 Summary of Research Results

This project successfully performed a demonstration of a topology discovery algorithm in collaboration with Survalent and Peninsula Light Company, leveraging real feeder models and sensor data to ensure practical relevance. Through this demonstration, the team showed that advanced topology discovery tools can be applied in real-world environments by interfacing directly with the ADMS system. The approach relies on actual field data, with all its typical deficiencies and characteristics, highlighting the system's capability to handle real operational conditions. As a result, these tools can enhance situational awareness in distribution system operations.

Additionally, results highlighted that sensor placement is critical to algorithm accuracy, yet many existing sensors in the distribution grid—such as reclosers—are deployed for protection or operational purposes and may not always provide the data needed for topology discovery. Therefore, it is essential to conduct sensitivity analyses to determine cost-effective sensor deployment strategies that support multiple objectives while ensuring system observability. While limitations remain in handling sparse or imperfect data, the research provides a pathway for incremental improvements and field deployment. Overall, the study offers actionable insights for utilities seeking to modernize operations and lays the groundwork for integrating topology discovery tools into commercial platforms to enhance grid reliability and visibility.

6.0 References

- Dalavi, F., M. E. H. Golshan, and N. D. Hatziargyriou. 2024. "A review on topology identification methods and applications in distribution networks." *Electric Power Systems Research* 234: 110538.
- Demaine, E. D., M. T. Hajiaghayi, and K.-i. Kawarabayashi. 2005. "Algorithmic graph minor theory: Decomposition, approximation, and coloring," in 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05), pp. 637–646, IEEE.
- Gandluru, A., S. Poudel, and A. Dubey. 2020. "Joint Estimation of Operational Topology and Outages for Unbalanced Power Distribution Systems," *IEEE Transactions on Power Systems* 35 (1): 605–617, doi: 10.1109/TPWRS.2019.2935401.
- Hagberg, A. A., D. A. Schult, and P. J. Swart. 2008. "Exploring network structure, dynamics, and function using networkx," in Proceedings of the 7th Python in Science Conference (G. Varoquaux, T. Vaught, and J. Millman, eds.), (Pasadena, CA USA), pp. 11–15.
- Kannan, K., and C. L. Monma. 1978. "On the computational complexity of integer programming problems," *Optimization and Operations Research*, pp. 161–172, Springer.
- Korres, G. N., and N. M. Manousakis. 2012. "A state estimation algorithm for monitoring topology changes in distribution systems," *2012 IEEE Power and Energy Society General Meeting*, San Diego, CA, USA, pp. 1–8, doi: 10.1109/PESGM.2012.6345126.
- Li, Z., S. Jazebi, and F. de León. 2018. "Determination of the Optimal Switching Frequency for Distribution System Reconfiguration," *IEEE Transactions on Power Delivery*, 32 (4): 2060–2069, doi: 10.1109/TPWRD.2016.2594385.
- Poudel, S., T. Ramachandran, A. Veeramany, C. Francis, and A. P. Reiman. 2024. "Topology Identification Using Graph Theory Informed State Estimation-Based Model Selection for Power Distribution Systems," *IEEE Transactions on Industrial Informatics* 20 (3): 3563–3573, doi: 10.1109/TII.2023.3310740.
- Sevlian, R. A., and Rajagopal, R. 2015. "Distribution system topology detection using consumer load and line flow measurements." *arXiv preprint arXiv:1503.07224*.
- Srivastava, A., S. Chakrabarti, J. Soares, and S. N. Singh. 2022. "An optimization-based topology error detection method for power system state estimation." *Electric Power Systems Research* 209: 107914.
- Tian, Z., W. Wu, and B. Zhang. 2016. "A Mixed Integer Quadratic Programming Model for Topology Identification in Distribution Network," *IEEE Transactions on Power Systems* 31 (1): 823–824, doi: 10.1109/TPWRS.2015.2394454.
- Zamani, V., and T. Nielsen. 2020. "How to Create an Accurate Network Model and Dynamic State Data for an Advanced Distribution Management System (ADMS)."

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov