The Data Challenge

The distribution circuit model that is maintained by the electric utility often terminates at the service transformer, neglecting the secondary circuit. The effects of missing and/or inaccurate data in the geospatial information system (GIS) model are compounded when the model information is required by other applications and systems to study or operate the grid.

Solution Overview

Robust algorithms and applications must be developed to predict and confirm the impedance and therefore the characteristics of the conductor representation in the GIS and to automatically update the GIS with limited or no user input.

Potential Methods for Solving the Problem

The manual method of having utility or contract staff physically walk out the distribution line and document the secondary conductor characteristics is an option. This endeavor would ultimately lead to an accurate secondary model. However, the process would be costly and time consuming, and it would lend itself to entry errors.

A more automated approach is to use available information along with reasonable assumptions to estimate the important characteristics of secondary conductors with each transformer-to-meter connection. Based on the secondary conductor design practices associated with the date of transformer installation, a reasonable assumption could be made for the type of conductor likely installed. Furthermore, the size of the conductor could be estimated based on the size of the transformer and/or the service ampacity rating, and the average length could be assigned to all transformer-to-meter connections. However, this method could potentially lead to unacceptable modeling errors. To determine the length of the conductors, the GIS data set of the utility assets could be merged with a data set having locational information about the buildings and homes. With the transformer pole location and the approximate location of the meter, an algorithm could be developed to calculate the length of the secondary conductor. This system of assumptions can approximate the secondary circuits with reasonable accuracy.

The third method builds an impedance model of the secondary circuits using data analytics and requires no manual data entry. The estimation of conductor impedance is based on a linear regression model to estimate the resistance and reactance of two conductors that share a common voltage point: either a distribution service transformer or common secondary point [1]. The known input values would be the voltage at each of the smart meters and the real and reactive components of the current, which can be calculated using the real and reactive power measurements of the smart meters. Once meters served by the same distribution service transformer are associated, the resulting impedance values are known for each secondary conductor. Therefore, the complete secondary model can be developed for each distribution service transformer. This method can be expanded to confirm existing models for primary circuits.
Available Data Sets

The data sets highlighted in the following figure are available in the EPRI Data Repository to solve this data analytics case.

Classifications of Data:

- **Traditional Data Set**
- **New Data Set**
- **Structured Data**
- **Un-structured Data**
- **Format of Data Varies**

Frequency of Measurement:

- **C** Cycles
- **S** Seconds
- **M** Minutes
- **H** Hours
- **D** Days
- **Y** Months to Years
- **E** Event Driven

Denotes a primary data set used to solve this data analytics case.