January 2017

Business & Technology Strategies

Application of AMI Data to Anomaly Detection and Dynamic Power Flow Analysis

Based on a paper written by:

D. W. Pinney and T. Kirk, NRECA
and
F. A. Eldali, Graduate Student, Colorado State University & Member of IEEE

NRECA
America's Electric Cooperatives

EXECUTIVE SUMMARY

What has changed in the industry?

Advanced Metering Infrastructure (AMI) has been deployed at over 70% of rural electric cooperatives, and this new data source offers opportunities for valuable applications beyond billing. NRECA has developed open source computer software to perform anomaly detection and dynamic power flow analysis with AMI data for cooperatives. Dynamic power flow analysis is an important area of development because it is able to model the time-varying behavior of emerging distributed energy resources.

What is the impact on cooperatives?

Cooperatives have the ability to get more value out of their AMI data by analyzing it for anomalies including faults, damaged meters, or energy theft. Finding these problems quickly can help cooperatives save money and maintain high member satisfaction.

What do cooperatives need to know or do about it?

Cooperatives need to consider how they are using their AMI data and if dynamic models provide an avenue for more value. In addition to the applications discussed in this paper, AMI data can also be used for asset monitoring, outage management, power quality assessment, voltage optimization, demand side management, minimizing energy theft, and better loss estimation.

I. INTRODUCTION

Grid modernization is a term commonly used to refer to the growth of distributed energy resources and networked controls on the electric system. One of the technologies central to this modernization is Advanced Metering Infrastructure (AMI). AMI deployment is increasing throughout the United States, and co-ops have the highest percentage of installations in the country compared to investor-owned and municipal utilities [1], [2]. More than 70 percent of the cooperatives in the U.S. are using AMI for its numerous benefits including remote meter reading, improved outage management, support for dynamic rate structures, distribution generation (DG) monitoring, improved load forecasting accuracy, and measurement and verification of demand side management programs [3],[4].

The data provided by meter readings has increased considerably since the end of the 18th century. At first, conventional meters were read once a month in-person to determine the electricity usage. Then, as technology progressed advanced meter read (AMR) systems provided remote energy readings and reduced the cost of billing. Recently, advanced meter infrastructure (AMI) provides more granular hourly (or sub-hourly) reads, increasing the amount of data to be transmitted and stored [5]. These data volumes, although historically not common in the electric utility industry, are not a significant challenge to modern information technology infrastructure. Among rural electric cooperatives, the largest is Pedernales which serves approximately 300,000 meters. If each meter is transmitting data every 15 minutes, the daily data volume is approximately 1 GB/day with a sustained DB insertion rate of 12 Kbps. At the current prices for storage available on demand from cloud hosting providers, supporting a data workload of that volume would cost less than $2,700 annually [6], a small expense for a utility. Using purchased instead of on-demand IT infrastructure would lower that amount significantly.

The data collected by the meters can be sent through one of multiple channels, including power line carrier (PLC), broadband over power lines (BPL), copper or optical fiber, or wireless radio frequency (RF). The rate of transmitting data depends on the communications medium, and bandwidth limits there constrain how much interval data can be retrieved by the utility for analysis. Many cooperatives use PLC because it is well suited for remote, sparsely populated service territories due to being the lowest cost medium for a given network size [4],[5].

One of the major problems still facing AMI data research is the lack of a common data format for AMI data. Without a widely adopted standard, sharing data generated by different vendors' products is very challenging and stymies potential research efforts. The ANSI C12.19 standard used in AMR meters could be a good candidate for developing a matching standard for AMI data, but support by vendors is incomplete

[5]. In the absence of such a standard, in this work we defined our own simple standard based on a comma separated value format.

Once the utility has acquired AMI data, they can use a Meter Data Management System (MDMS) or other database to store and manage AMI data. The main advantage of using an MDMS is the functionality for validating, editing, and estimating (VEE) AMI data for billing and other purposes [4]-[7]. Many new MDMS capabilities are emerging including (i) aggregating data using “virtual meters” (ii) facilitating information sharing between generation and transmission (G&T) and distribution utilities, (iii) managing direct load control events.

This paper describes the developed open source computer software to perform anomaly detection and dynamic power flow analysis with AMI data for the cooperatives. The rest of the paper is organized as follows:

- Section II explains the application of AMI data to anomaly detection model.
- Section III explains how we applied AMI data to perform dynamic power flow simulations using GridLAB-D.
- Section IV discusses further potential applications such as theft detection and technical losses estimation.
- Section V concludes the paper and provides ideas for future work.

II. Applying AMI Data to Anomaly Detection

We define an anomaly as a period of unusual energy consumption when compared with the normal (average) usage for a set period of time (e.g., a month). An anomaly can be quantified by measuring its deviation from the average during that time, and can be caused by bad data, damaged or malfunctioning meters, faults, or energy theft. In this section, we describe the software we developed for detecting anomalies in AMI datasets.

In the literature, there has not been a particular work that examines anomaly detection on its own. Research related to anomaly detection has been done as part of non-technical losses (NTL) estimation and energy theft detection. A key area of research focuses on the problem of data analytics in MDMS for energy theft detection. Even though some vendors have been offering theft detection functionality, the methodologies are unavailable publicly. Reference [8] examines theft detection through a threat model and with higher resolution data. The key contributions of [8] are: (i) considering an attacker model for theft detection in MDMS, (ii) developing a metric that considers the classification accuracy of the theft detector, and (iii) utilizing real AMI data to evaluate the performance of the theft detection by deploying an autoregressive moving average (ARMA) detector.

Excluding and quantifying NTL from TL has been an important topic in AMI analytics research. Reference [7] uses a state estimation methodology to estimate consumption at the distribution transformer (DT) to detect and quantify NTL. It uses the data of individual consumers and the aggregated data at the DT to localize the usage anomaly. The proposed algorithm has two main phases. First, it detects the type of anomaly or meter fault. Second, it finds the location of the customer with either the suspected theft or low voltage level. The results show the usefulness of using the distribution state estimation in the localization of an anomaly. Lastly, the analysis of the variance is important to detect individual anomalies. Reference [9] examines the error source and the methodologies to distinguish NTL from the total losses in a microgrid using two approaches, model-based and data-based. The model-based uses optimal power flow (OPF) from the Gridlab-D simulator. The data-based model uses a pre-trained regression model. The paper also proposes the localization of NTL at the distribution aggregators for future work [9].

The electricity consumption of a given residential or commercial consumer follows a highly predictable energy usage pattern that can be represented statistically [10]. Thus, relating unusual consumption to the average consumption is a simple, straightforward and practical way to detect irregularities in energy consumption. Our work uses this technique to detect anomalies from a specific meter. The model input consists of AMI data for multiple meters, the minimum energy deviation from the average to detect as a percentage (threshold), and the minimum deviation length to detect in hours as shown in Fig.1.

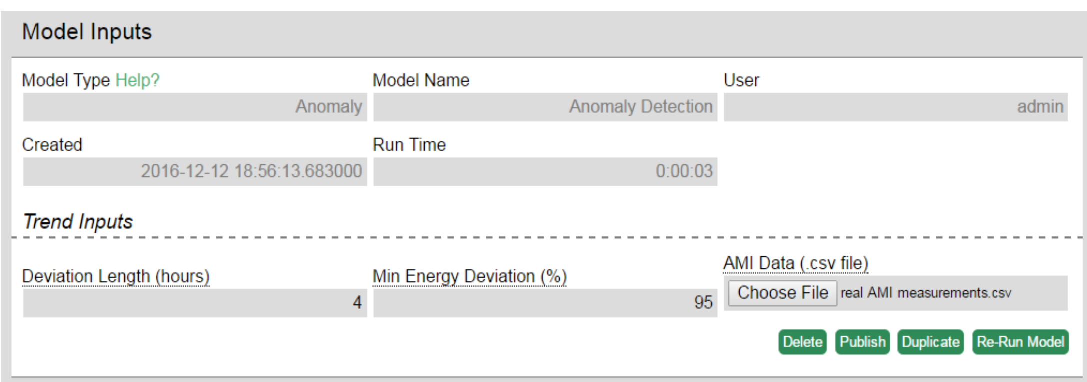

Fig. 1. Inputs of the anomaly detection model

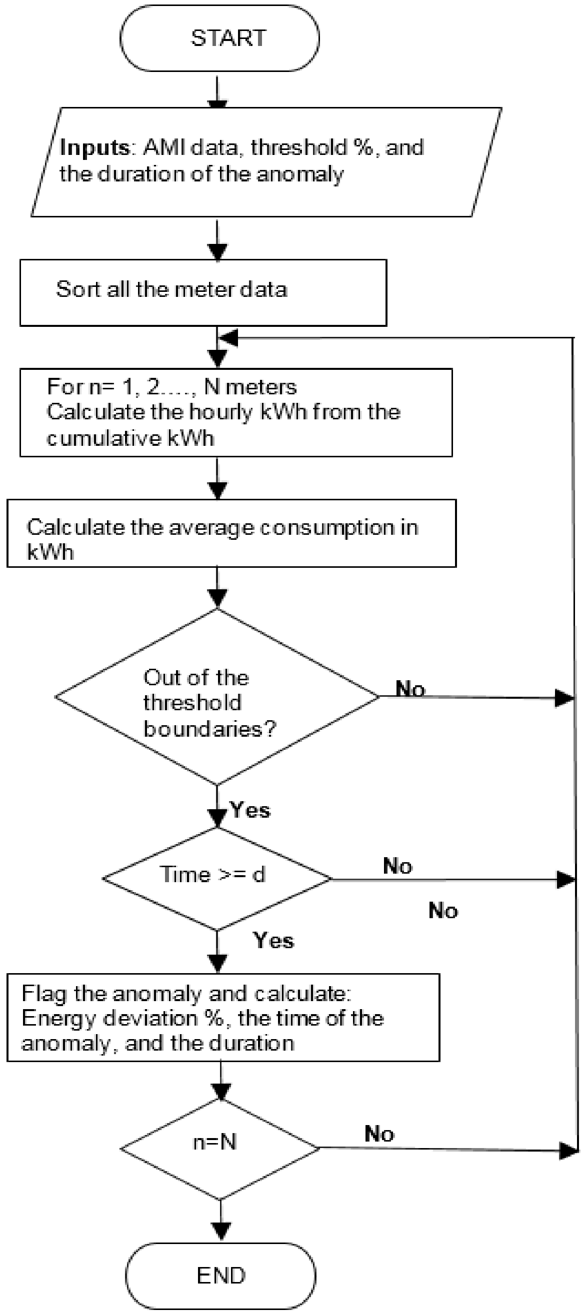

Fig. 2. Anomaly detection algorithm

Fig.2 explains how the code works to detect anomalies for multiple meters. Let N be the number of meters, d the deviation length as chosen by the user. For each meter the average consumption is calculated as $\mu = \frac{1}{T}\sum_{t}^{T} E(t)$, where T is the period under consideration (e.g., a month), E(t) is the energy consumption in MWh, and t is the time step (e.g., an hour, or 15 minutes). The full open source code for the model is available on Github [11].

Figures 3 and 4 show the outputs of the model. Fig. 3 shows the record of all the anomalies. Note that one meter could have multiple anomalies during the period of consideration. Fig. 4 allows the user to choose any of the meters with an anomaly from a dropdown menu and observe their load profile.

Consumption Anomalies

Meter ID #	Anomaly Start Date-Time	Duration (hours)	Max Deviation from Avg (%)
111569742	2012-09-30 23:00:00	744	Zero Demand
113363106	2012-10-01 00:00:00	144	-100.0
113363106	2012-10-14 00:00:00	72	-100.0
113363106	2012-10-27 00:00:00	48	-100.0
113363474	2012-10-11 00:00:00	120	-100.0
113363474	2012-10-23 00:00:00	48	-100.0
113368277	2012-10-03 23:00:00	96	114.91
113559360	2012-09-30 23:00:00	744	-100.0
113559627	2012-09-30 23:00:00	744	-100.0
113559891	2012-10-02 23:00:00	96	144.69
113559939	2012-10-02 23:00:00	120	324.91
113560340	2012-10-29 23:00:00	48	-100.0

Fig. 3. Detected anomalies for each meter

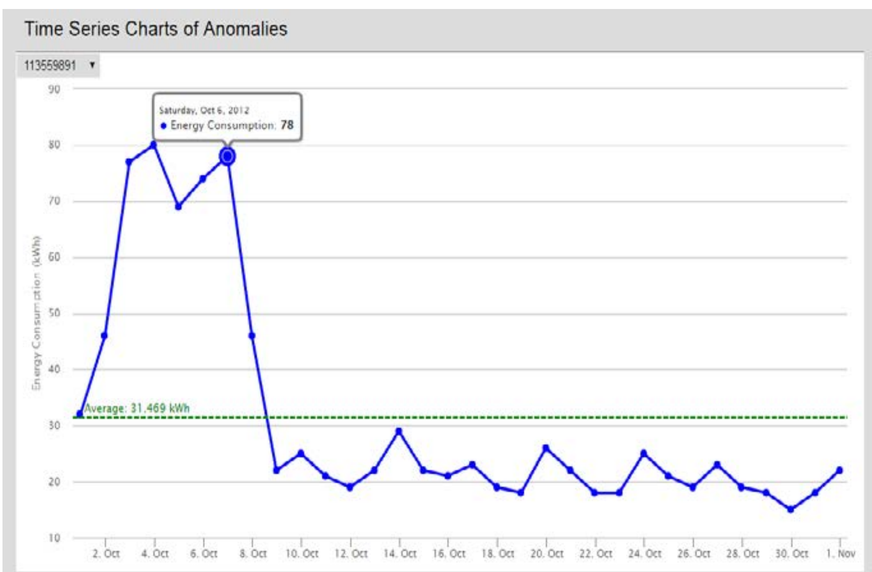

Fig. 4. Consumption of a particular meter where an anomaly was detected

Anomalies could be due to a fault, a damaged meter, or energy theft, and fixing the problem achieves both monetary and non-monetary benefits. Non-monetary benefits, such as consumer engagement and satisfaction–and in the case of a fault, consumer safety–are important to utilities. Monetary benefits of anomaly detection are also important and can be calculated more specifically. For energy theft, the value is avoiding lost sales. Lost sales due to an energy theft can be calculated as $L_S = \alpha \times \pi/100$ where L_S is the monthly lost sales in dollars due to energy theft, α is the energy deviation percentage due to theft, and π is the average monthly bill of that particular customer. For example, the average monthly bill of a customer in the United States is $120 according to the U.S. Energy Information Administration. In the test dataset of a small sample of meters we collected from a rural electric cooperative, we were able to find 10 anomalies. Assuming these are all non-technical losses and the average amount lost is 50 percent, the total revenue lost was $600 per month.

An anomaly caused by a damaged meter is another straight-forward value proposition. A damaged meter will read zero consumption, meaning the utility is losing out on all energy sales to that consumer. Thus, replacing the meter with a new one is much less expensive than 100 percent lost sale for several months. Detecting multiple anomalies at a particular location can also identify failing distribution equipment.

III. APPLYING AMI DATA TO DYNAMIC POWER FLOW SIMULATION

Static power flow analytical methods and tools, at both the transmission and distribution levels, lack the ability to model the time-varying behavior of emerging distributed energy resources [12]-[14]. Dynamic power flow simulations (DPFS), also referred to as quasi-static time series analysis (QSTS) by some researchers, can model changes in power flow over time. This approach enables: (i) enhanced detection of voltage problems within the distribution grid, (ii) development of dispatch strategies for energy storage, (iii) assessing distribution system problems such as theft or non-technical losses, (iv) modeling the behavior or load tap changers, voltage regulators, and switched capacitor banks, and (v) quantifying problem duration [11].

DPFS have been used in studies for integrating solar PV on the distribution grid. The study by Broderick et al. [12] used DPFS to quantify (i) the variation in PV impact based on time, feeder type and point of interconnection and (ii) the variable power output impacts on the operations of voltage regulators, capacitor banks, and switches, which can results in flicker and voltage violation. Similar to [10], reference [13] examines the impact of solar PV integration on the IEEE 34 node distribution test feeder. The paper uses high temporal resolution (1 min) solar PV data for an entire year. The impacts investigated include the violation of the voltage level, frequency of the voltage regulator operation, and characteristics of the power losses in the system. One of the disadvantages of the study is that the load profile data was unavailable. Therefore, static load equal to the peak load was assumed.

Several tools are in development to perform DPFS. Report [14] compares three of the most commonly used distribution system simulators that support DPFS analysis. These are CYMDIST (commercially available), OpenDSS (free and open source), and GridLAB-D (free and open source). GridLAB-D, which is used in our work, has several advantages over the others including support for more powerflow solvers (Gauss-Seidel, Newton Raphson, and Forward-backward sweep), models for capacitor bank control, voltage regulating transformers and many forms of distributed generation, and broader operating system support. GridLAB-D only supports a command-line interface, but NRECA has developed a graphical interface with many analytical enhancements, the Open Modeling Framework (OMF), available at https://www.omf.coop/.

Dynamic simulations require historical energy consumption data for all the meters in the distribution feeder. There are multiple ways to generate these load profiles, including using SCADA data from substation and downline regulators, or using purely synthetic models, but one of the simplest and clearest methods is to use AMI data. We developed software to translate for processing AMI data and creating these load models [14]. In addition to Wh consumption data, we require meter ID, phase, and time stamp (YYYY-MM-DD HH:MM:SS) information in a comma-separated value format.

IV. Additional Applications

There are other applications that can be developed using AMI data. These include asset monitoring, outage management, power quality assessment, voltage optimization, demand side management, minimizing energy theft, and better loss estimation. This section explains applications that we did not implement, but will potentially investigate in future work.

Illegal consumption of energy (energy theft) is an ongoing problem for utilities. Utilities are fighting back with new theft detection techniques that leverage AMI data. Reference [15] integrated both AMI and SCADA with state estimation methodology for theft detection. One of the disadvantages of the data requirements is that the method needs several measurements which might be challenging to obtain. The work in [16] depends on a statistical estimation approach to separate the technical losses and NTL before and after the DT. This method is similar to another method known as “totalizing” that calculates the mismatch between the total energy supplied by the DT and the sum of all legal energy consumption.

The theft detection technique proposed in [17] consists of two phases. First, the illegal use event is identified within the low voltage distribution grid by a simple power imbalance test. The second is the localization at the point of connection (POC). The localization is based on examining the voltage error, which is the difference between the estimated (simulated) voltage and the measured voltage by the AMI meter.

Typical distribution losses in North America are 6-10%. This percentage varies based on the distribution system's topology [18]. Reference [18] proposed evaluation techniques employing AMI data and geographical information systems (GIS). First, system components information, hourly consumption at the substation and individual meters are collected from GIS and AMI. Second, total losses, secondary losses, and DT losses are calculated. This straightforward approach can have a significant impact in determining the time and location of losses.

V. Conclusion and Future Work

AMI data has potential uses beyond just energy sales. Cooperatives are actively investigating how to this data to enhance their distribution systems and better serve their consumer members. The work described above offers two software applications that perform anomaly detection and distribution load modeling for dynamic power flow simulations.

Future work to be done in this area includes:

(i) developing a graphical user interface to the distribution load modeling code,

(ii) implementing a theft detection,

(iii) implementing loss estimation, and

(iv) measuring and verifying results against a second meter data source.

Contacts for Questions

Thomas Kirk and David Pinney
Business and Technology strategies
National Rural Electric Cooperative Association
Emails: Thomas.Kirk@nreca.coop David.Pinney@nreca.coop

Fathalla Eldali
Department of Electrical and Computer Engineering
Colorado State University
Email: Fathalla.Eldali@colostate.edu

References

1. C. Federal Energy Regulatory, "Assessment of Demand Response and Advanced Metering. 2015," ed.
2. (2016, March). National Rural Electric Cooperative Association, . The Consumer-Centric Utility Future-The 51st State [Online]. Available: www.sepa51.org/submissions/Roadmap_Reports/NRECA_51s (Accessed on: 12/13/2016)
3. F. Eldali, T. Hardy, C. Corbin, D. Pinney and M. Javid, "Cost-benefit analysis of Demand Response programs incorporated in Open Modeling Framework," 2016 IEEE Power and Energy Society General Meeting (PESGM), Boston, MA, USA, 2016, pp. 1-5.
4. (2013, December). L. Moorefield and D. Houghton. National Rural Electric Cooperative Association, . Advanced Meter Interval Data: Now You Have it, so What Should You Do with It.
5. S. E. McHann, "Grid analytics: How much data do you really need?," 2013 IEEE Rural Electric Power Conference (REPC), Stone Mountain, GA, 2013, pp. C3-1-C3-4.
6. https://calculator.s3.amazonaws.com/index.html
7. S. C. Huang, Y. L. Lo and C. N. Lu, "Non-Technical Loss Detection Using State Estimation and Analysis of Variance," in IEEE Transactions on Power Systems, vol. 28, no. 3, pp. 2959-2966, Aug. 2013.
8. D. Mashima and A. A. Cárdenas, “Evaluating electricity theft detectors in smart grid networks,” in Research in Attacks, Intrusions, and Defenses. Berlin, Germany: Springer-Verlag, 2012, pp. 210–229.
9. M. Buevich et al., "Microgrid Losses: When the Whole Is Greater Than the Sum of Its Parts," 2016 ACM/IEEE 7th International Conference on Cyber-Physical Systems (ICCPS), Vienna, 2016, pp. 1-10.
10. P. Jokar, N. Arianpoo and V. C. M. Leung, "Intrusion detection in advanced metering infrastructure based on consumption pattern," 2013 IEEE International Conference on Communications (ICC), Budapest, 2013, pp. 4472-4476.
11. github.com/dpinney/omf/tree/master/omf/scratch/amiAnomalyDetection
12. R. J. Broderick, J. E. Quiroz, M. J. Reno, A. Ellis, J. Smith, and R. Dugan, "Time Series Power Flow Analysis for Distribution Connected PV Generation," Sandia National Laboratories SAND2013-0537, 2013.
13. B. A. Mather, "Quasi-static time-series test feeder for PV integration analysis on distribution systems," 2012 IEEE Power and Energy Society General Meeting, San Diego, CA, 2012, pp. 1-8.
14. R. Ayyanar, “Distribution System Analysis Tools for Studying High Penetration of PV with Grid Support Features,” PSERC Final Report, September 2013.

15. W. Luan, G. Wang, Y. Yu, J. Lin, W. Zhang and Q. Liu, "Energy theft detection via integrated distribution state estimation based on AMI and SCADA measurements," 2015 5th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), Changsha, 2015, pp. 751-756.

16. D. N. Nikovski, Z. Wang, A. Esenther, H. Sun, K. Sugiura, T. Muso, and K. Tsuru, “Smart meter data analysis for power theft detection,” in Machine Learning and Data Mining in Pattern Recognition. Springer, 2013, pp. 379–389.

17. P. Kadurek, J. Blom, J. F. G. Cobben and W. L. Kling, "Theft detection and smart metering practices and expectations in the Netherlands," 2010 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe), Gothenburg, 2010, pp. 1-6.

18. J. Triplett, S. Rinell and J. Foote, "Evaluating distribution system losses using data from deployed AMI and GIS systems," 2010 IEEE Rural Electric Power Conference (REPC), Orlando, FL, 2010, pp. C1-C1-8.