IMPROVING THE DATA QUALITY OF THE LV-CONNECTIVITY

Tom PYCKE
Eandis - Belgium
tom.pycke@eandis.be

Dirk COSTROP
Eandis - Belgium
dirk.costrop@eandis.be

Luc HENDERIECKX
Eandis - Belgium
luc.henderieckx@eandis.be

ABSTRACT

Knowing the exact connectivity of a distribution grid leads to major savings on the long term. In this paper we propose a method to calculate this connectivity solely based on the data we receive from smart meters. After elaborating on the mathematical background of the algorithm, we prove the applicability by illustrating its usage with a real-life example. We conclude by giving some ideas about possible future work around this topic.

INTRODUCTION

The term low voltage connectivity refers to (a) which customers are connected to which low voltage feeder, and (b) which feeder is connected to which distribution substation. GIS systems typically try to determine relation ‘a’, while the most interesting relation is a combination of ‘a’ and ‘b’: to which distribution substation is a customer connected. For this relation an algorithm using smart metering data is proposed.

In April 2010, Eandis started the installation of 4000 smart meters to conduct a proof of concept. The main topic to be investigated was the performance of a powerline communication concept (Frère *et al.*, 2011). Several other topics were also investigated, amongst others different ways to determine the low voltage connectivity of our customers.

Business drivers

Knowing the exact connectivity on the low voltage grid is getting more important every year.

Electricity outage
Before a planned outage of the LV-electricity grid, customers are informed about the upcoming maintenance. Knowing the exact connectivity avoids informing a customer that will not be affected. It also avoids not informing a customer, which would be worse.
In some countries, DGO’s have to pay a fee to customers when they suffered an unplanned electricity downtime. This requires the DGO to know the exact LV-connectivity.

Demand side management
Synergies between smart metering and smart grid initiatives allow demand side management for residential customers. Bad connectivity data decreases the efficiency of the used algorithms.

Fraud detection
Smart metering projects have a positive business case thanks to possibility to detect unmeasured electricity consumption. Bad connectivity data will cause a lot of false positive detections. The overall minimum detectable fraud will rise in case of unreliable connectivity data.

Maintenance
Knowing the exact connectivity of a distribution grid allows the DGO to use this grid more optimal in regard to balancing and preventive maintenance.

Connectivity issues

Until now, the only way to know the real connectivity of a customer was to use special equipment. With this equipment and a lot of manual effort, the exact connectivity can be determined. Unfortunately time and cost stop us from using these tools to determine the connectivity for a large area.
As an alternative to this method, utilities often use geographic data and a simple heuristic to determine the LV-connectivity on a larger scale. The cable segment that lies geographically closest to the customer or connection in question, is considered the connected cable. In various situations, this heuristic fails:

- Streets with more than 1 LV-cable
- Junction of 2 cables
- Near distribution substations
- Parallel streets close to the connection

In urban areas, these situations occur more often than in rural regions. On top of these failures, the geographic location of old infrastructure is not always known and is hard to localize with cables running underground.

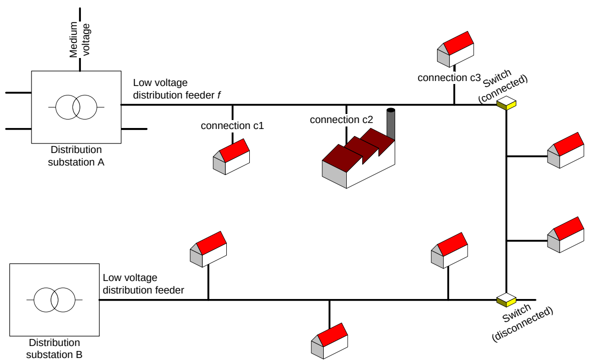

Figure 1: Simplified representation of a distribution grid and the nomenclature used in this paper

PROPOSED ALGORITHM

Because of a research on fraud detection using the concept of energy balance, every distribution feeder in the substations has a smart meter connected to it using a current transformer. This allows us to measure the energy going through every feeder. Every connection on this feeder is also measured.
The unit of measurement is kWh (energy). This unfortunately means that copper/cable losses on the distribution feeders are included in the energy measurements. A current instead of energy measurement would therefore increase the precision of the algorithm.

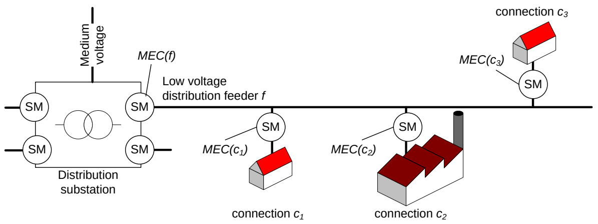

Figure 2: On every distribution feed and every connection, a smart meter (SM) is installed

The sample rate used is 1 energy measurement every 15 minutes. However, in the proposed algorithm the number of samples and not the sample rate is important.

Mathematical model

Assume

- c_i : connections on distribution feeder f
- $EC(x)$: Actual energy consumption
- $MEC(x)$: Measured energy consumption
- Every smart meter has a measurement error, which is considered constant over a longer sample interval.

Then

$$MEC(f) = \left(CableLosses + \sum_i EC(c_i) \right) \cdot MeasurementError(f)$$

We assume that every connection has an impact on the cable losses that is directly proportional with its consumption:

$$CableLosses \approx \sum_i \alpha_i \cdot EC(c_i)$$

This gives us:

$$\begin{aligned} MEC(f) &\approx \left(\sum_i (1+\alpha_i) \cdot EC(c_i) \right) \cdot MeasurementError(f) \\ &= \sum_i \frac{(1+\alpha_i) \cdot MEC(c_i) \cdot MeasurementError(f)}{MeasurementError(c_i)} \\ &= \sum_i MEC(c_i) \cdot \beta_i \end{aligned}$$

Per connection c_i we need to find a constant β_i which contains the measurement error and the impact of this connection on the cable losses.
Because of the approximation of the cable losses and other side effects such as small deviations on the smart meter's system clocks, we need several samples to determine β_i.
The most straightforward method to calculate β_i would be a least-squares method:

$$MEC(f) = \sum_i MEC(c_i) \cdot \beta_i$$

Assuming we have for every connection a load profile of n samples, this becomes in a matrix representation :

$$\begin{pmatrix} MEC(f)_{t1} \\ \vdots \\ MEC(f)_{tn} \end{pmatrix} = \begin{pmatrix} MEC(c_1)_{t1} & \cdots & MEC(c_m)_{t1} \\ \vdots & \ddots & \vdots \\ MEC(c_1)_{tn} & \cdots & MEC(c_m)_{tn} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$$

The β_i values can be calculated by using a QR-decomposition or an iterative least squares method with constraints on the resulting coefficients.
In a perfect situation, β_i should be between 1 and 1.1, assuming a maximum cable loss of 10%.

Taking it one step further

Assume

- $c_{f,i}$: connections on distribution feeder f
- $c_{g,j}$: connections not on distribution feeder f
- $c_i = \{ c_{f,i} , c_{g,j} \}$

With other words, we extend our collection of connections with some extra ones which are not on feeder f.
The resulting coefficients β will be:

- $\beta(c_{f,i}) \geq 1$
- $\beta(c_{g,j}) = 0$

The least squares algorithm will assign a zero coefficient to the load profiles not connected to feeder f.

Wrapping it all up

Using the information given above, we can design an algorithm to determine the connectivity:

1. Select a feeder f.
2. Choose a period p which, depending on the sample rate, contains enough samples.
3. Select a collection of connections which are good candidates to be connected to feeder f. Good selection criteria can be street names, geographical coordinates, additional information from PL-communication, …
4. Retrieve the load profiles for period p, for both the feeder and the collection of connections.
5. Build the corresponding matrix representation and feed them into the least squares algorithm.
6. The load profiles with a corresponding coefficient β with a value around 1, is connected to feeder f.

Extension to phase identification

The same approach can be used for determining the phase(s) the meter is connected to. To support this, the tri-phase smart meters (in both the distribution cabinet and at our customers' premises) will need to measure the energy or current going through every phase. No additional features are required for the mono-phase meters.
In the algorithm every phase will need to be treated as if it was a separate feeder in the original algorithm. All load profiles will need to be added to the "collection of connections" of each phase.
A typical residential LV-distribution feeder contains both mono-phase and tri-phase connections. For mono-phase connections this means the load profile will be excluded on 2 phases and selected on 1 phase. For tri-phase connections there will be 3 load profiles. One will be selected per phase, and the other 2 will be excluded. Because we introduce more unknown coefficients into the model than we add measurements, the precision of the phase-identification algorithm will be lower than the connectivity determination algorithm.

EXPERIMENTAL RESULTS

The experimental results are based on the smart metering data collected during the proof of concept in 2010.
Because our smart metering infrastructure is not configured to measure each phase separately, we will only illustrate the connectivity determination algorithm by providing a real-life example.
For illustration purposes, the distribution feeder we will use is a small one (6 connections). Figure 3 shows the load profile of the distribution feeder (red), the sum of the load profiles of the connections on the feeder (according to our asset data – blue) and the green curve shows the difference between these two. The positive difference indicates one or more missing connections.
We need to find a collection of "candidate connections" for the selected distribution feeder. This was done by taking all connections in a radius of 100 meters around the distribution feeder (according to our GIS and asset data). Because this feeder is located in a rural area, this resulted in only 3 extra connections.
After retrieving the new load profiles from our Meter Data Management system, we fed them into an iterative least squares algorithm in Matlab and received the coefficients shown in figure 4. Connections c6, c7 and c8 are the ones we added. The algorithm clearly shows that c7 is also connected to our distribution feeder. The resulting load profiles can be seen in figure 5 and shows a perfect match between the energy that goes through the distribution feeder, and the energy that is consumed by the connections.

Experiments with larger distribution feeders showed similar results.

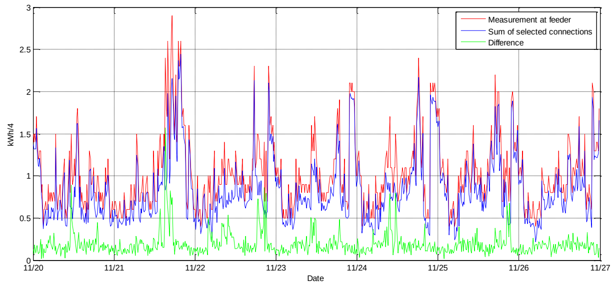

Figure 3: Non-zero difference indicates missing connections

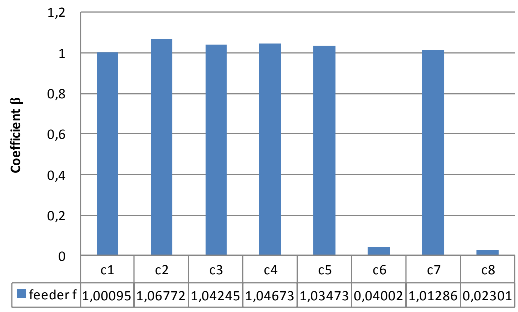

Figure 4: Coefficients as a result of the least-squares method

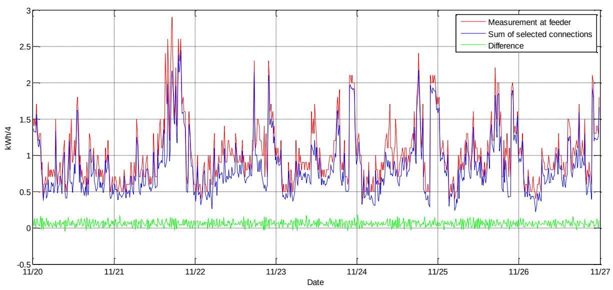

Figure 5: Zero difference indicating correct connectivity

FUTURE WORK

In our future work we will optimize the proposed algorithm so it can be applied on a large scale.
Smart meters measuring the energy consumption or current per phase will allow us to calculate the LV-connectivity per phase instead of per connection. This will bring advantages to people responsible for net-balancing, especially in areas with a lot of solar panels. Field tests will be required to
As already explained, energy measurements include cable losses and are not ideal for this application. It needs to be further investigated to which degree current measurements improve the accuracy of this method.

CONCLUSION

Smart meters have numerous advantages. We have shown that the energy consumption data that is collected by even the most basic smart meters can be used to determine the connectivity of the distribution grid.
The algorithm we proposed proved to be successful using real-life data.

MISCELLANEOUS

Acknowledgments

We gratefully thank our employer Eandis to give us the time and opportunity to work on this. We would also like to thank Dirk Costrop and Luc Henderieckx for their support and encouragements.

REFERENCES

[1] S. Frère, V. Vancaeyzeele and L. Henderieckx, 2011, "Technological breakthrough for real time smart metering via power line communications (PLC)", *Proceedings CIRED Conference*, #400