
Peer reviewed version

Link to published version (if available):
10.1109/ISGTEurope.2013.6695379

Link to publication record in Explore Bristol Research

PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/
Inferring Low Voltage Transformer State Using Only Smart Metering Data

Qipeng Chen
and Dritan Kaleshi
Centre for Communications Research
Dept.of Electrical & Electronic Eng.,
University of Bristol,
Bristol, BS8 1UB, United Kingdom
{Qipeng.Chen, Dritan.Kaleshi}@bristol.ac.uk

Zhong Fan
Toshiba Research Europe Limited,
Telecommunications Research Laboratory,
32 Queen Square
Bristol, BS1 4ND, United Kingdom
Zhong.Fan@toshiba-trel.com

Abstract—Previous works on Distribution State Estimation (DSE) utilizing Smart Metering (SM) data assume the availability of sufficient measurements from distributed generators and Low Voltage (LV) transformers as well. However, both distributed generators and LV transformers are seldom equipped with advanced meters due to the high expenses. This paper investigates how DSE is performed using only SM data when inferring the states at LV transformers. SM data usually is captured synchronously but reported over a given time interval. The impact of the errors introduced by non-perfect meter clock synchronisation is analysed using real SM data, and is confirmed to be not significant. The experimental results further show the possibility of using SM data alone to characterize the states at LV transformers.

I. INTRODUCTION

Low Voltage (LV) transformers are seldom equipped with sensors [1]. Their voltage and power are typically inferred from historical load records or statistics, according to which the operation conditions can be determined and the decisions regarding system control are made subsequently. Using the inferences as the states of transformers is effective only when the historical load profiles are reliable, while the introduction of distributed and renewable generations and electrical vehicles will bring in more load dynamics in the near future. Moreover, in order to improve human’s living qualities and to reduce power wastes, new grid applications will be introduced. For example, Demand Response updates electricity tariff in real time and send the price signal to smart meters at the consumer ends, in order to adjust users’ power consumption behaviours [2]. Demand Side Management is acting in households which arranges the usage of the appliances in an optimal manner, in order to reduce the power consumptions and to avoid peak uses of power [3]. Utilizing any of these technological advances will also give rise to difficulties in system control.

Smart Meters will provide raw measurement data at the edge of the power distribution system. And the introduction of Advanced Metering Infrastructure (AMI) provides an opportunity to use the Smart Metering (SM) data. Voltage and current magnitudes, active and reactive power and the power factor at consumer ends are provided to the data centres constantly [4]. A unique standard specifying the sampling rate of the smart meters does not exist yet, while the measurement intervals of every 15, 30 or 60 minutes are widely adopted. Various literatures have explored the use of the data, in particular for Distribution State Estimation (DSE). DSE is a premise to other grid applications, that it filters out the errors from the provided measurements and estimates voltage magnitudes and phase angles at the positions where the measurements are not directly provided [5].

The use of SM data for DSE has been reported already. Reference [6] shows the participation of real time SM data improves the performance of DSE. Reference [7] shows the possibility of running DSE when smart meters cannot provide measurements in time, which is to use the previous day’s smart metering data as a replacement. Reference [1] overcomes the same issue from a machine learning perspective. A machine learning algorithm is developed, which predicts a transformer’s load from the historical data when smart meters cannot provide data in real time. Otherwise, the load of the transformer is obtained by aggregating the smart meter readings. Both of the predicted and the real loads are put into a robust medium voltage (MV) state estimator. The predicted value is adjusted and fed back to the machine learning algorithm for further training. After sufficient time of training, the closed-loop procedure is able to infer the states at the transformers accurately, even when no real time measurements are provided.

This paper evaluates the possibility of estimating power and voltage values at the local transformer using only SM data, while the aforementioned references assume the availability of the measurements from sensors at distributed generators or transformers themselves.

The aggregation of smart meter measured power is an important input to DSE and other grid applications, while it would deviate from the real value if the clocks of the meters are not well synchronized. Clocks equipped at the smart meters ensure all measurements are with a same time stamp. However, the emerged abnormal electricity behaviours like power peak may cause clock deviations. Therefore certain techniques are utilized to maintain an acceptable synchronisation rate. Typically, a reference clock is deployed at the data centre, which periodically sends correction signals to meter clocks.

Wireless communication will be adopted in smart grid infrastructure, which brings in difficulties when utilizing typical synchronisation techniques [8]. It will be time consuming to transmit the correction signals to all smart meters, particularly
when traffic congestions happen. At present, a GPS strategy has been applied to maintain the clocks in Phasor Measurement Units (PMUs), which leads to a $\pm 1 \mu s$ accuracy [9]. However, adopting the advanced strategy for smart meters will incur far more costs. In this paper, the negative impact of meter clocks’ synchronisation errors are assessed as well. The result will indicate if it is necessary to utilize the advanced synchronisation strategy for smart meters.

The paper is structured as follows. The procedure of estimating transformer power and voltage values from SM data alone is described in Section 2. Impact of errors introduced by lack of synchronisation of smart meters is analysed in Section 3. The simulation results are summarized and described in Section 4. Conclusions are provided in Section 5.

II. INFERRING THE STATES OF LV TRANSFORMERS

The 33 node network introduced in [10] is shown in Fig. 1. Node 1 is the substation bus. The voltage at the bus can be measured and reported by the equipped sensors. All of the other nodes are buses of LV transformers where no measurement devices are installed. The substation and multiple LV transformers are interconnected as a MV power circuit. Each LV transformer is the starting point of a LV power network. It supplies electricity to around 150 households. Each of the households is equipped with at least one smart meter [11].

![Fig. 1. A 33 node medium voltage power network.](image)

Let us look at the problem of estimating power and voltage values at the LV transformers using only SM data. To simplify the problem we assume there are no distributed generators utilized, and the smart meters measure and report active power $P$, reactive power $Q$ and voltage magnitude $V$.

As a LV transformer is the unique power source of the LV circuit which is downstream from it, the power at the transformer is equivalent to the sum of the loads and the power losses of the circuit. The active power $P_t^b$ and the reactive power $Q_t^b$ of transformer $b$ at time $t$ are given by (1a) and (1b) respectively, where $n$ is the index of a household and $S_n$ is the total number of the households connecting with the transformer.

$$P_t^b = \sum_{n=1}^{S_n} p_{t,n}^b + P_{t,\text{lost}}^b$$  \hspace{1cm} (1a)

$$Q_t^b = \sum_{n=1}^{S_n} q_{t,n}^b + Q_{t,\text{lost}}^b$$  \hspace{1cm} (1b)

The line impedances along a LV circuit is hard to obtain in realistic cases, which leads to computational difficulties when taking into account the power losses. Besides, the power lost is much smaller than the loads of the LV circuit. Therefore the active power $\tilde{P}_t^b$ and reactive power $\tilde{Q}_t^b$ at a LV transformer are initially represented by the aggregation of smart meter readings. Their representations are given by (2a) and (2b) respectively. The same formula is also applied in [6], [7], [1] and [12]. Following from our assumption of not having local generators, the nearest household from the transformer has the maximum voltage which is taken to represent the secondary voltage at the transformer (2c).

$$P_t^b \approx \tilde{P}_t^b = \sum_{n=1}^{S_n} p_{t,n}^b$$  \hspace{1cm} (2a)

$$Q_t^b \approx \tilde{Q}_t^b = \sum_{n=1}^{S_n} q_{t,n}^b$$  \hspace{1cm} (2b)

$$\tilde{V}_t^b = \max_{n \in [1:S_n]} v_{t,n}$$  \hspace{1cm} (2c)

$\tilde{P}_t^b$ and $\tilde{V}_t^b$ include errors due to power losses and voltage drop. Once the initial representations at all LV transformers of the MV system are known, we use the Weighted Least Squares (WLS) state estimator introduced in [5] for the MV system.

$$\min_x j(x) = \sum_{i=1}^{Z_n} (z_i - h_i(x))^2 / R_{ii}$$  \hspace{1cm} (3)

$$= [z - h(x)]^T R^{-1} [z - h(x)]$$

The state estimator aims at obtaining the states of the MV system which minimizes the overall errors involved in all measurements. This is an optimisation problem and is formulated by (3), where $z$ is the state of the system including the phase angles (θ) and voltage magnitudes (V) of all of the nodes. In the 33 node system, $x$ at time $t$ is represented as the vector of $[\theta_{t,1}^1...\theta_{t,1}^{33} \ V_{t,1}^1...V_{t,1}^{33}]$. $z$ is the set of the available measurements of the system. And $z$ at time $t$ is given by $[\tilde{P}_{t,1}^1...\tilde{P}_{t,1}^{33} \ \tilde{Q}_{t,1}^1...\tilde{Q}_{t,1}^{33} \ V_{t,1}^1...V_{t,1}^{33}]$, where $V_{t,1}$ is the voltage magnitude at the substation and $i$ is the index of a node. $h_i(x)$ in (3) is a measurement function relating $x$ to a measurement $z_i$, so the residual $r_i = z_i - h_i(x)$ reflects the difference between the provided measurement and the measurement calculated from the estimated states. $R$ is the variance of the measurements. It emphasizes the respective belief degrees of different measurements.

As (3) is non-convex and is hard to solve, Newton’s method is applied to derive the optimal solution. It initializes all nodes’ voltage magnitudes to 1 and phase angles to 0 and iteratively updates $x$ through $x^{k+1} = x^k + \Delta x^k$ in each iteration $k$ until $j(x)$ in (3) converges. The increment $\Delta x^k$ is calculated through (4), where $H$ is the measurement Jacobian, and $G(x^k)$ is called the gain matrix which equals $H^T R^{-1} H(x^k)$.

$$[G(x^k)] \Delta x^{k+1} = H^T (x^k) R^{-1} [z - h(x^k)]$$  \hspace{1cm} (4)

Strong assumptions on good synchronisation between meter clocks are made in the procedure, which may not be possible in practice, therefore contributing errors to $\tilde{P}_t^b$ and $\tilde{Q}_t^b$. The impact of these errors is evaluated in the next section. And the robustness of the state estimator to power losses and voltage drop is evaluated in Section 4.
III. Synchronisation Evaluation

As described before, the power at a LV transformer is initially obtained by aggregating smart meter readings, and the aggregation of the power will deviate from the real value if smart meters’ clocks are not well synchronized. In this section, the impact of errors introduced by lack of synchronisation of smart meters is evaluated. We use the Electricity Customer Behaviour Trial database from the Commission for Energy Regulation (CER) [13]. The database records 6444 households’ half-hourly power consumption between 14-07-2009 and 31-12-2010. We only select a subset from the database which is sufficient to complete the evaluation. The subset consists of 92 days’ records from June, July and August of 2010 which yields 4416 \((48 \times 92)\) measurements of each household. Among all the households, those with significant missing records are removed. As a consequence, 6007 households’ records are retained. We divide all households into 32 groups randomly, and take each of the groups as a LV system. Finally, the data is converted to active power \((\text{kWh})\) by multiplying with 2, because the original records are half-hourly power consumption and are with the unit of \(\text{kWh}\).

The load at a transformer equals the sum of the power of all households connected with it. The load profile of a transformer is the transformer’s time varying loads. We call the load profile of a transformer as Proper Aggregation Power (PAP), if there is no synchronisation error involved in the meters’ clocks. PAP of a transformer \(b\) is represented as \(P_T^{\text{b}} = [P_T^{\text{b}}_1, P_T^{\text{b}}_2, \ldots P_T^{\text{b}}_{4416}]\). As the smart meters used by CER are assumed to be well synchronized, PAP of all transformers can be calculated.

A smart meter’s readings will deviate from the original records, if certain time error is incorporated into the meter’s clock. The daily load profile of a meter is represented by \([p_{1,\nu}^b, p_{2,\nu}^b, \ldots p_{48,\nu}^b]\), where \(\nu\) is the index of the household, and \(b\) is the transformer which is directly connecting with \(\nu\). The measurements should be generated at exactly \([00:30, 01:00\ldots 24:00]\) of the reference clock (The time vector is converted to \([1800, 3600\ldots 86400]\) in the following contexts with second as the unit). However, if the clock of \(\nu\) is \(\Delta t_{\nu}\) seconds faster than the reference clock, meter \(\nu\) would provide measurements of \([\hat{p}_{1,\nu}^b, \hat{p}_{2,\nu}^b, \ldots \hat{p}_{48,\nu}^b]\) at \([1800 - \Delta t_{\nu}, 3600 - \Delta t_{\nu}, \ldots 86400 - \Delta t_{\nu}]\).

Given the original measurements \([p_{1,\nu}^b, p_{2,\nu}^b, \ldots p_{48,\nu}^b]\), the time error \(\Delta t_{\nu}\) of \([\hat{p}_{1,\nu}^b, \hat{p}_{2,\nu}^b, \ldots \hat{p}_{48,\nu}^b]\) can be estimated by a local regression method. As shown in Fig. 2, \([\hat{p}_{1,\nu}^b, \hat{p}_{2,\nu}^b, \ldots \hat{p}_{48,\nu}^b]\) are plotted as red dots. The regression method fits a curve \(f_\nu^b(t) = \hat{p}_{1,\nu}^b(t)\) to the measurements. Therefore the value at any time point including \([1800 - \Delta t_{\nu}, 3600 - \Delta t_{\nu}, \ldots 86400 - \Delta t_{\nu}]\) can be estimated. For example, \(\hat{p}_{1,\nu}^b = f_\nu^b(1800 - \Delta t_{\nu})\).

The aggregated power at a transformer will be different from the transformer’s PAP, if time errors exist. The difference indicates the impact of the time errors on power aggregation. In order to evaluate the impact, we introduce different levels of time errors to the smart meters. It is assumed meters’ time errors follow a Gaussian distribution: \(N(0, \sigma^2)\), where \(\sigma\) is the standard deviation. Different \(\sigma\) indicates different levels of time errors. We choose \(\sigma\) from \([1s, 2s, 5s, 10s, 15s, 20s, 30s, 45s, 60s]\) or \([100s, 200s, 300s, 400s, 500s, 600s, 700s, 800s, 900s]\). With each \(\sigma\), we use the corresponding Gaussian distribution function to generate time errors for all smart meters.

The time errors of all smart meters belonging to transformer \(b\) are represented as \(\Delta T = [\Delta t_1, \Delta t_2, \ldots \Delta t_n]\), where \(n\) is the total number of the smart meters. For each smart meter \(\nu\) with time error \(\Delta t_{\nu}\), the local regression method is applied which changes its power records of \(P_T^{\nu} = [p_{1,\nu}^\nu, p_{2,\nu}^\nu, \ldots p_{4416,\nu}^\nu]\) to \(\hat{P}_T^{\nu} = [\hat{p}_{1,\nu}^\nu, \hat{p}_{2,\nu}^\nu, \ldots \hat{p}_{4416,\nu}^\nu]\). \(\hat{P}_T^{\nu}\) of all households in \(b\) are aggregated as \(\hat{P}_T^{b} = [\hat{p}_1^b, \hat{p}_2^b, \ldots \hat{p}_{4416}^b]\). The error of the aggregated power relative to the true value at time \(\tau\) is given by (5).

\[
\Delta P_T^b = (\hat{P}_T^b - P_T^b) / P_T^b
\] (5)

For each \(\sigma\), the process is repeated 100 times. The standard deviations of the relative errors at every time point are calculated. The averaged results over multiple days and multiple groups are derived and shown in Fig. 3 and Fig. 4.

As shown in both figures, the x-axis denotes different time intervals of a day. The y-axis is the standard deviation of the relative errors resulted from aggregating non-synchronized power data. Different levels of time errors are shown as the legends. The curves of the cases when time errors equaling 1s, 2s and 5s overlap each other, so only 7 curves are shown in Fig. 3. As it can be seen, the increase of the time error leads to a decrease of the accuracy of power aggregation. However, the
influences are not significant, particularly when \( \sigma \) is relatively small. For example, the maximum synchronisation error among meter clocks are specified as 20s in [14]. And for this case, the accuracy is under 0.1\% over the whole day.

IV. Simulations

In this section, the same subset of data is used to access the impact of power losses, voltage drop and synchronisation errors on DSE when characterizing the voltage magnitudes at LV transformers.

A. Test System

The 33-node power circuit shown in Fig. 1 is used as the test network. The voltage magnitude at node 1 is provided. The divisions of the 32 groups in Section 3 are still used in the simulation. Each group is related to one of the transformers. The time varying active power of a transformer is represented by the PAP of the related group. As the database only records power consumption, we set a constant value as the ratio of reactive to active power \( \frac{Q}{P} \) of each transformer, in order to derive the reactive power. MATPOWER is then applied to calculate the voltage magnitudes and phase angles at all of the nodes. MATPOWER is a Matlab simulator for studying electrical engineering, which is powerful to solve power flow problems [15]. All of the data is considered to be the true states of the system.

B. Experimental Results

The impact of different synchronisation rates on DSE is evaluated. We firstly introduce errors into the true states, in order to simulate the realistic scenarios. The standard deviations of the relative errors resulted from aggregating non-synchronized meter readings are summarized in Fig. 3 and Fig. 4. They are represented as \( \omega \) in the following contexts. \( \omega \) varies with time and different synchronisation rates. We use \( \omega \) to introduce errors into transformers’ active and reactive power. The active power \( P_r \) of a transformer at time \( \tau \) changes to \( \tilde{P}_r \) through \( \tilde{P}_r = P_r \times (1 + R_{\text{normal}}(\omega_{\tau})) \), where \( R_{\text{normal}}(\omega_{\tau}) \) generates Gaussian distributed variables with the mean of 0 and the standard deviation of \( \omega_{\tau} \). The introduction of errors into transformers’ reactive power is in the same manner. Moreover, the voltage magnitude at node 1 is assigned an error with the standard deviation equal to 0.2\%. The voltage magnitudes of the other nodes are assigned a 5\% error.

For each synchronisation rate, the step above introduces errors into the true states over the whole period. The new measurements of different time points are put into the WLS state estimator in turn. The Mean Absolute Percentage Error (MAPE) of nodes’ voltage magnitudes is calculated. It quantifies the performance of DSE of a specific time. It is given by \( V_{\text{MAPE}} = \frac{1}{n} \sum_{i=1}^{n} |V_i - \tilde{V}_i| \), where \( n \) is the number of nodes. \( V_{\text{MAPE}} \) along the whole period are calculated. And the averaged daily \( V_{\text{MAPE}} \) is obtained. Finally, we have 18 daily \( V_{\text{MAPE}} \) for different synchronisation rates. It is found the maximum \( V_{\text{MAPE}} \) value of a day mostly appears between 11:00 and 22:30. This is the period when transformers’ load profiles are relatively dynamic. Instead of presenting these daily \( V_{\text{MAPE}} \) themselves, we only show the minimum, maximum and mean values of each daily \( V_{\text{MAPE}} \) in Fig. 5 for a more intuitive demonstration of the impact.

As shown in the figure, the x-axis denotes different synchronisation rates. The minimum, mean and maximum values of the daily \( V_{\text{MAPE}} \) of each synchronisation rate are distinguished by different colour. As it can be seen, with the increase of time errors, the minimum \( V_{\text{MAPE}} \) keeps stable, while the maximum \( V_{\text{MAPE}} \) increases obviously, which means time errors only influence DSE when power consumption behaviours are dynamic. However, the decline of the performance is limited. As for the case of the synchronisation error equalling 20s, even the maximum \( V_{\text{MAPE}} \) is smaller than 0.06 \%. Therefore the errors involved in meter clocks do not decrease the performance of DSE significantly.

Using only SM data to represent transformers’ voltage magnitudes and active power are not accurate due to voltage drops and power losses. This experiment access their impact on DSE. The same subset of data is used, which contains smart meters measured half hourly power consumption of 92 days (or 2208 hours). We obtain the true states at different time in the same manner as before. Three cases of voltage drops are considered, which are 5\%, 10\% and 15\% respectively. For each case, transformers’ active power losses over the whole
period increase linearly from 5% of hour 1 to 20% of hour 2208. As only resistances exist along the transmission line, the reactive power losses should be relatively small, which is ignored in the experiment. A 3% accuracy is assigned to reactive powers due to measurement errors or errors resulted by poor synchronisation rates. We modify the true states by introducing the pre-defined power losses and voltage drops. The WLS state estimator is then applied to the data, and the MAPE of the voltage magnitudes ($V_{MAPE}$) are shown in Fig. 6. The x-axis denotes different time points with hour as the unit. The y-axis denotes the value of $V_{MAPE}$. From top to bottom, the three graphs correspond to voltage drops equalling 5%, 10% and 15% respectively. The $V_{MAPE}$ remains under 1% when the voltage drop is set to 5%. As it is recorded in [16], the total active power loss at a distribution power system is between 3% and 9%. We consider an extreme case when all of the 9% power is lost in the LV system. And in this experiment, the 9% power loss happens at hour 600. As shown in the figure, as for this case, even when the voltage drop is as large as 15%, the $V_{MAPE}$ is smaller than 0.5%.

V. CONCLUSION

This paper describes how MV DSE can be performed using only SM data, and evaluates the potential impacts of different meter clock synchronisation errors. The experimental results show the impact of time errors among smart meters are not significant, and the use of only smart metering data is able to characterize the states at LV transformers in realistic scenarios.

ACKNOWLEDGMENT

We thank the Commission for Energy Regulation (CER) for supplying the Electricity Customer Behaviour Trial database. The authors would also like to express sincere thanks to Toshiba Research Europe Limited for funding this work.

REFERENCES
