Parameter identification of unknown radial grids for theft detection — **Source link** ⧉

Sam Weckx, Carlos Gonzalez, Jeroen Tant, T. De Rybel ...+1 more authors

Related papers:

- Non-Technical Loss Detection Using State Estimation and Analysis of Variance
- Nontechnical Loss Detection for Metered Customers in Power Utility Using Support Vector Machines
- Electricity Theft Detection in AMI Using Customers' Consumption Patterns
- A Multi-Sensor Energy Theft Detection Framework for Advanced Metering Infrastructures
- Privacy-Preserving Energy Theft Detection in Microgrids: A State Estimation Approach

Parameter Identification of Unknown Radial Grids for Theft Detection

Sam Weckx, *Graduate Student Member, IEEE,* Carlos Gonzalez, Jeroen Tant, *Graduate Student Member, IEEE,* Tom De Rybel, Johan Driesen, *Senior Member, IEEE*

Abstract—This paper proposes an algorithm to detect the stealing of electricity by illegal connections in smart grids with unknown or uncertain cable lengths. Many new applications in the rising smart grid context will require information of the grid topology. We show that with measurement data of smart meters, the grid can be identified, as well as the phase of connection. One of the applications requiring grid information is the detection of electricity theft by double feeding. Electricity theft is a problem faced by all power utilities. Financial impacts are a reduced income for the system operator and the necessity to charge more to other customers. Stealing of electricity by double feeding can not be detected by the smart meter or by analysing the load profiles. Therefore it is suggested in this paper to use measurements of smart meters to identify the grid parameters and detect irregularities of specific customers.

Index Terms—Smart metering, theft detection, grid parameter identification

I. INTRODUCTION

High penetration of distributed generators and electric vehicles in low voltage distribution network challenges the future grid operation. More intelligent methods should be used for better utilization of the distribution network, in order to maintain, or even to improve, the power-supply reliability and quality. A lot of these new control methods are based on accurate network information of the distribution grid. Grid topology information is required for example when including grid constraints in the coordinated charging of vehicles [1], for Volt-Var optimization [2] and State-Estimation [3] in distribution grids as well as for electricity theft detection based on a load flow algorithm [4]. This grid topology information is however often not available in an easy accessible digital structure, as the requirements for distribution grids were low in the past, especially for low voltage grids. Therefore there is a need towards automated grid topology identification. This paper presents an algorithm to extract a linearised load flow model out of smart meter data and applies this to the detection of electricity theft.

Electricity theft is one of the main concerns of utility companies. The amount of theft is small (1-2%) in terms of electricity generated in many systems, however the corresponding financial loss is high due to the large amount of electricity distributed [5]. It is shown that electricity theft is increasing in most regions of the world [5]. Lost earnings might result in decreased investments in power system improvement or

S. Weckx, C. Gonzalez, J. Tant, T. De Rybel and J. Driesen are with the Department of Electrical Engineering ESAT-ELECTA, KU Leuven, Kasteelpark Arenberg 10, bus 2445, 3001 Leuven, Belgium.
Corresponding author's email address sam.weckx@esat.kuleuven.be

in increased tariffs for other customers by rising distribution tariffs for electricity. Besides that, stolen energy might be used for illegal activities, like the production of narcotics [4].

Different types of theft exist: Fraud, stealing of electricity and unpaid bills. When committing fraud, a customer tries to deceive the utility by tampering the meter so that a lower consumption is measured. Stealing of electricity can be performed by double feeding, bypassing the meter so that the consumption for additional feeding is not registered. The introduction of smart meters and the modernization of measurement systems will make electricity theft harder [6]. Different features are included in smart meters itself to make them tamper proof. Another way to detect fraud with smart meters is by analysing the measured load profiles with supervised learning methods like support vector machines [7], extreme machine learning [8] or decision trees [9]. Stealing of electricity by double feeding remains hard to detect without a visual inspection, as this does not involve the smart meter itself, it will not be detected by it. Also the load profile might not be influenced.

Radial distribution networks are either overhead or underground. In both of these cases it is hard to make an illegal connection directly to the network. Therefore, electricity theft usually takes place in the customers house. By checking all customers equipment, the illegal lines can be detected. The illegal lines are often easy to detect as they are highly visible. If they are not visible, Time Domain Reflectometry (TDR) techniques can be used to detect the illegal connection between the home service entrance and the metering point [10]. This remains a time consuming task and might require too much of the Distribution System Operators (DSO) human resources. Besides that, sometimes the illegal lines might be well hidden, especially when applied for illegal activities.

Theft localization can also be done based on a load flow algorithm [4]. Errors between measured and simulated voltages will point out the place of theft. Exact lengths between houses, or of the connecting cable between smart meter and distribution feeder are however normally not available. For these cases an algorithm is proposed in this paper that identifies the grid and points out the customer which steals electricity by double feeding or by tampering the meter. The proposed algorithm assumes that at every customer both active and reactive power are measured, as well as voltage magnitude by a smart meter. The system operator can access this data which might be stored at a database and can identify the time instances of theft by comparing the measured power at the substation with the sum of the measured consumed power. The low voltage network considered is a three-phase, four-wire radial system with a
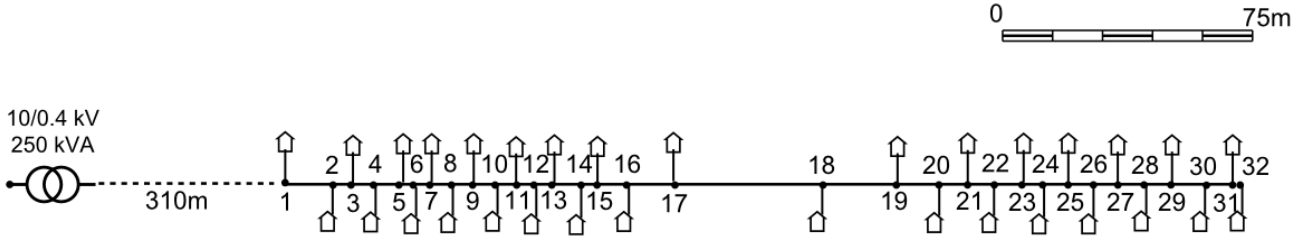
Fig. 1. The network used in the simulations. All lengths are drawn to scale.

TABLE I
PROPERTIES OF THE SIMULATED NETWORK

Properties	Value	Unit
Total feeder length	594	[m]
Average length between houses	9.2	[m]
R (resistance) of feeder	0.206	[Ω/km]
L (inductance) of feeder	0.248	[mH/km]
Connection length between house and feeder	10	[m]
Cross section of connection cable (copper)	16	[m^2]

TT earthing arrangement for residential customers. Only radial operated distribution networks are considered, as these are the strict majority for the connection of residential customers to the grid.

II. SIMULATION OF A DISTRIBUTION NETWORK

The network used for simulations is a 3-phase, 4 wire radial low voltage system with a TT earthing arrangement for residential customers out of a city-environment of Flanders. The city feeder has a short average length between two consecutive households, typically shorter than 10 meters. In total there are 32 customers connected. The first customer connected to the feeder is far from the substation. The topology of the feeder is depicted in Fig. 1. However, the algorithm can be applied to grids with side branches as well. Table I gives the grid properties.

III. PHASE AND GRID IDENTIFICATION

If distribution grid parameters are available, the voltages in all nodes of the grid can be calculated with a load flow algorithm. For radial distribution networks, the backward-forward sweep method can be applied [11]. The load flow problem can be represented as a non-linear system of equations of the complex powers of all customers:

$$f(V_{1,k}, ..., V_{N,k}, S_{1,k}, ..., S_{N,k}) = 0, \qquad (1)$$

where $V_{h,k}$ and $S_{h,k}$ are 3x1 vectors representing the three-phase voltage magnitude and the complex power consumption at house h for timestep k, and N is the number of houses connected to the considered grid.

This non-linear model is approximated with a linear model:

$$V_{h,k} = V_k^0 + \sum_{\tilde{h}=1}^{N} a_{h,\tilde{h}} P_{\tilde{h},k} + \sum_{\tilde{h}=1}^{N} b_{h,\tilde{h}} Q_{\tilde{h},k}, \qquad (2)$$

where
- V_k^0 is the voltage magnitude at the distribution transformer at timestep k
- $P_{\tilde{h},k}$ is the active power of customer $\tilde{h}$ at timestep k
- $Q_{\tilde{h},k}$ is the reactive power of customer $\tilde{h}$ at timestep k
- $a_{h,\tilde{h}}$ is the influence of active power of customer $\tilde{h}$ on the voltage magnitude of house h
- $b_{h,\tilde{h}}$ is the influence of reactive power of customer $\tilde{h}$ on the voltage magnitude of house h

In case the loads would be modelled as current sources, the system would behave linearly. Therefore, the linearisation would not introduce errors due to the superposition principle and (1) and (2) would be identical. In here, loads are modelled as PQ-sources and therefore linearisation errors will occur.

We assume that smart meters measure active and reactive power, as well as the voltage magnitude. In case a house is connected three-phase to the grid, voltage and power are measured in each phase separately. This data is stored in a database. If multiple time steps are available, an ordinary least squares can be recognised in (2) with $a_{h,\tilde{h}}$ and $b_{h,\tilde{h}}$ as unknowns. If the influence on the voltage of house h needs to be defined, one can solve the following problem:

$$\min_{a_{h,\tilde{h}}, b_{h,\tilde{h}}} \quad \sum_{k=1}^{n_t} \|V_{h,k} - V_k^0 - \sum_{\tilde{h}=1}^{N} a_{h,\tilde{h}} P_{\tilde{h},k} + \sum_{\tilde{h}=1}^{N} b_{h,\tilde{h}} Q_{\tilde{h},k}\|_2^2, \qquad (3)$$

where n_t is the total amount of time steps available for least squares.

This linear least squares problem does not require any information about the grid. Nor the location or the order of the customers matters. The data used for the identification in the grid is historic data of time steps at which there was no theft. These time steps can be found by comparing the sum of the measured power of all smart meters with the measurements at the substation transformer. The constants $a_{h,\tilde{h}}$ and $b_{h,\tilde{h}}$ allow to recompute the voltage in a house. For the theft location algorithm, all constants need to be defined, such that the voltage across the distribution feeder can be calculated with measurements of time steps with theft.

The constants $a_{h,\tilde{h}}$ and $b_{h,\tilde{h}}$ obtained by the LS method will give a measure of the location throughout the grid. They give also information about the phase connection. In case the smart meter that had measured the voltage V^h is connected to the same phase as the household $\tilde{h}$, the constant $a_{h,\tilde{h}}$ will be
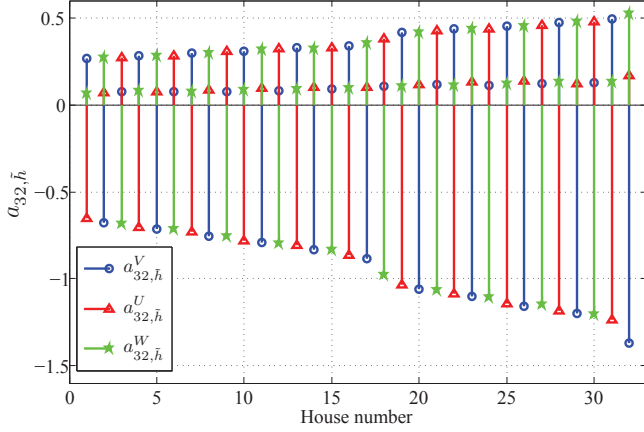
Fig. 2. The influence of active power of all houses on the phase voltages of house 32.

Fig. 3. Power consumption and voltage for all customers.

negative, as the active power has created a voltage drop. On the other hand, when the household is connected on another phase, $a_{h,\tilde{h}}$ will have a small positive value, due to the neutral point shift [12]. This allows to group all loads that are connected at the same phase, based on the constants $a_{h,\tilde{h}}$ and $b_{h,\tilde{h}}$. Fig. 2 shows the influence of loads of all houses on the phase voltages of house 32, when all households are connected by a single phase to the network. This is obtained by applying the least-squares method for 1000 time steps. It can be seen that the phase of connection applied to this case was an alternating sequence of U,V,W,U,V,W,...

This method of phase identification is less computational intensive as [13], where a mixed-integer program (MIP) is formed to identify the phase of connection of customers. [14] solves a combinatorial optimization problem by a Tabu search method. The phase of connection can also be identified by a unique signal injection, as proposed by [15]. This would require adapting the smart meter, which might be costly.

With the obtained linear load flow model, the voltages at the customers nodes can be computed by evaluating (2) without any knowledge about the cable lengths. Linearisation errors will occur in case customer nodes are assumed to behave as PQ-loads. An exact load flow with perfect knowledge of all cable lengths would have no voltage errors, however when uncertainty is added to each of the cable lengths, voltage errors will occur. To evaluate the linearisation errors occuring in a linearised load flow and the errors due to uncertainty in an exact load flow, voltages are calculated for a grid where an uncertainty of $\pm$ 10 % is added to each of the cable lengths. The linearised load flow model is obtained by a training set of 1000 time steps and the voltages are simulated for a validation set consisting of 1000 time steps. Fig. 3 presents the simulation parameters by a modified boxplot. The modified boxplot consists of an inner box that spans the 25th to 75th percentiles, the outer box spans the 5th to 95th percentiles, and the whiskers indicate the minimum and maximum values. The phase of connection, active and reactive power consumption/production as well as the customer voltage are depicted. Fig. 4 depicts the voltage error for an exact load flow with cable lengths that differ up to $\pm$ 10 % of the real
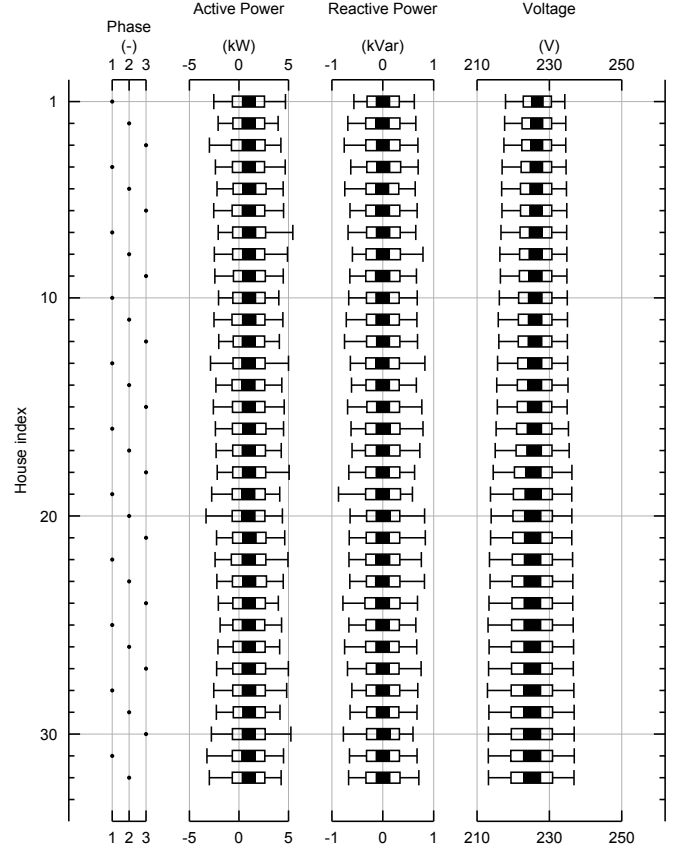
lengths and compares this with the linearised load flow. The errors in the linearised load flow are significantly smaller and remain below 1 V. The errors in the exact load flow are biased to one side due to overestimation of some of the cable lengths. Therefore we conclude that identifying the radial distribution grid is beneficial as soon as there is a small uncertainty on the cable lengths. In case there is no knowledge at all about the grid, it remains the only option.

When prior knowledge of the grid is available, this can be included by adding constraints to the LS problem and making it a convex optimization problem which can be efficiently solved [16]. When the order of houses is known, but not the cable lengths or cable type, this can be added to the optimization problem. For example the voltage drop in house h caused by a load in house 29 needs to be bigger than the influence of house 1 on the voltage of house h in case all of them are connected to the same phase. This can be written as:

$$\min_{a_{h,\tilde{h}}, b_{h,\tilde{h}}} \quad \sum_{k=1}^{n_t} \| V_{h,k} - V_k^0 - \sum_{\tilde{h}=1}^{N} a_{h,\tilde{h}} P_{\tilde{h},k} + \sum_{\tilde{h}=1}^{N} b_{h,\tilde{h}} Q_{\tilde{h},k} \|_2^2$$

$$\text{subj. to} \quad a_{h,29} \leq a_{h,1}$$

$$b_{h,29} \geq b_{h,1}.$$

$$(4)$$

Prior knowledge can sligthly improve the results of the obtained constants $a_{h,\tilde{h}}$ and $b_{h,\tilde{h}}$.

As European distribution networks are generally operated in
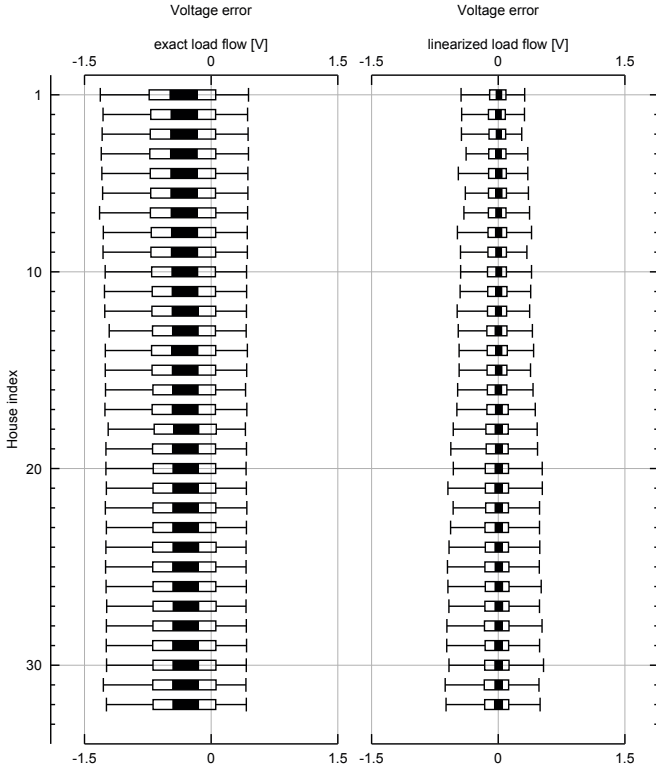
Fig. 4. Voltage estimation error for both the exact load flow as the linearised load flow for a radial grid with uncertain cable lengths (± 10 %).

Fig. 5. Flow chart of theft detection procedure.

a radial way, only these type of networks are evaluated. When applied to meshed or ring type of networks the linearisation errors by eq. (2) might increase.

IV. ELECTRICITY THEFT LOCALIZATION IN KNOWN RADIAL GRIDS

With the measured currents available by the smart meters, the voltage throughout the network can be estimated by running a load flow algorithm if the grid is known. The measured currents however do not coincide with the actual currents due to electricity theft. This result in a difference between the measured and the estimated voltage. In [4] it is proposed to use this estimation error to localize the place of electricity theft. If the estimation error does not clearly provide an accurate location, the localization will be extended for another time frame or the time frames can be aggregated. It is shown in [4] that this method is already robust against measurement errors. However, this method requires exact knowledge of the grid parameters. Cable lengths between subsequent houses are often not known or not known exactly, as well as the length of the cable connecting the house to the feeder.

V. ELECTRICITY THEFT LOCALIZATION IN UNKNOWN RADIAL GRIDS

When the grid is identified by least squares, all customer voltages at time steps of theft can be computed with (2). The stolen power will result in a lower voltage at the customer node then estimated by (2). This allows to identify the place
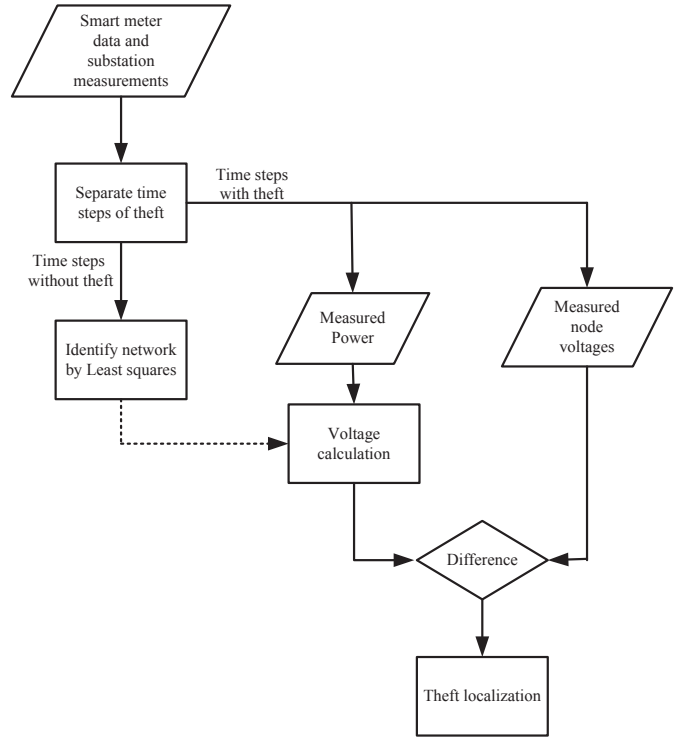
of theft. At each time step of theft, the difference between measured and estimated voltage is found. The location where this difference is most often the maximum is selected as the most likely place of theft. This is often also the place of maximal error. The flowchart of the theft detection procedure is shown in Fig. 5.

To test the algorithm, all customers are considered as unbalanced three phase loads. The active power consumed/produced by a customer in each phase is assumed to be random with a mean of 1 kW and a variance of 1 kW, the reactive power is assumed to be random with a mean of 0 kvar and a variance of 0.2 kvar. Customer 20 commits theft during 10 out of 1000 time steps. The average stolen power is 2 kW, the variance of the stolen power between different time steps is 20%. It is assumed that by measurements at the substation, the time instances of electricity theft are known. The typical cable used to connect the feeder with the houses is copper 16 mm^2, where the length is modelled as a gaussian distributed number with a mean of 10 m. The 3σ variance is assumed to be 2 m. Gaussian noise is added to both power (σ^P=5 W and σ^Q=5 VA) and voltage measurements (σ^V=0.1 V). In Fig. 6 the maximum voltage estimation error for all customers is plotted. The highest estimation error occurs at the customer committing theft.

One could argue that the linearisation errors made by evaluating (2) will result in an incorrectly identified location of theft. However, the voltage drop over the impedance of the cable connecting the smart meter with the distribution feeder is significantly higher then the linearisation errors for normal lengths of this cable. Fig. 7 shows the influence of the
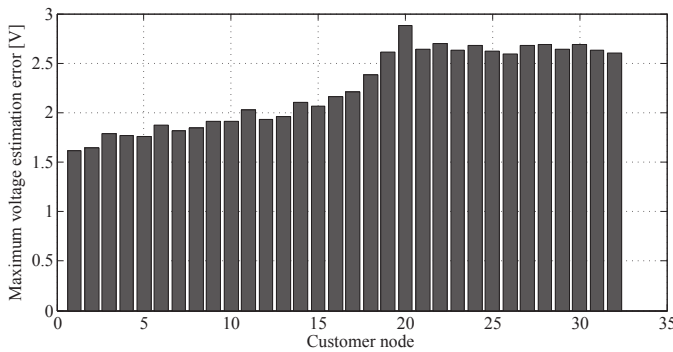
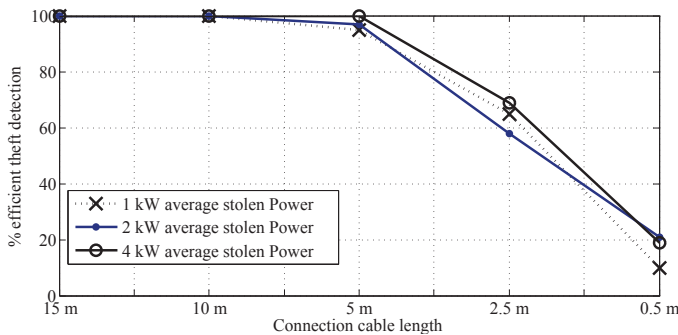Fig. 6. Maximum voltage estimation error for all houses in an identified grid.

Fig. 7. Succesfull theft detection as a function of average connection cable length.

length of this cable on the efficient detection of theft. Only for unlikely lengths shorter then 5 m, efficient detection might be a problem.

VI. Conclusion

Many applications in the rising smart grid context require information of the grid topology. In this paper it is shown how by means of smart meter data a linearised load flow model for unknown radial grids can be identified. The linearised load flow model provides information about both the location as well as the phase of each customer. It is shown that errors due to the linearisation are small. One of the applications of the linearised load flow model is electricity theft in radial grids. Electricity theft is a major problem in the distribution of electricity. The introduction of smart reading provides new opportunities to automate the detection of electricity theft. With information of the grid, the measured voltage can be compared with an estimated voltage that is the result of a load flow with the measured current. This method is inapplicable in case the length between houses is not known accurately. Therefore another method is proposed that improves the localization of electricity theft. The proposed algorithm combines grid identification and theft localization. Results show that the location of theft can be easily identified and can be applied to unknown grids.

VII. Acknowledgment

VIII. References

[1] K. Clement, E. Haesen, and J. Driesen, "Coordinated charging of multiple plug-in hybrid electric vehicles in residential distribution grids," in *Proc. IEEE/PES Power Systems Conf. and Exposition PSCE '09*, 2009, pp. 1–7.

[2] Z. Hu, X. Wang, H. Chen, and G. Taylor, "Volt/var control in distribution systems using a time-interval based approach," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 150, no. 5, pp. 548 – 554, sept. 2003.

[3] R. Gelagaev, P. Vermeyen, and J. Driesen, "State estimation in distribution grids," in *Harmonics and Quality of Power, 2008. ICHQP 2008. 13th International Conference on*, 28 2008-oct. 1 2008, pp. 1 –6.

[4] P. Kadurek, J. Blom, J. F. G. Cobben, and W. L. Kling, "Theft detection and smart metering practices and expectations in the Netherlands," in *Proc. IEEE PES Innovative Smart Grid Technologies Conf. Europe (ISGT Europe)*, 2010, pp. 1–6.

[5] B. Thomas and Smith, "Electricity theft: a comparative analysis," *Energy Policy*, vol. 32, no. 18, pp. 2067 – 2076, 2004.

[6] R. Alves, P. Casanova, E. Quirogas, O. Ravelo, and W. Gimenez, "Reduction of non-technical losses by modernization and updating of measurement systems," in *Proc. IEEE/PES Transmission & Distribution Conf. and Exposition: Latin America TDC '06*, 2006, pp. 1–5.

[7] A. H. Nizar, Z. Y. Dong, and Y. Wang, "Power utility nontechnical loss analysis with extreme learning machine method," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 946–955, 2008.

[8] J. Nagi, A. M. Mohammad, K. S. Yap, S. K. Tiong, and S. K. Ahmed, "Non-technical loss analysis for detection of electricity theft using support vector machines," in *Proc. IEEE 2nd Int. Power and Energy Conf. PECon 2008*, 2008, pp. 907–912.

[9] A. H. Nizar, Z. Y. Dong, J. H. Zhao, and P. Zhang, "A data mining based ntl analysis method," in *Proc. IEEE Power Engineering Society General Meeting*, 2007, pp. 1–8.

[10] K. Trupinic, M. Stojkov, and D. Poleto, "Reduction of non-technical losses based on time domain reflectometry (TDR) principles and function," in *Proc. 18th Int. Conf. and Exhibition Electricity Distribution CIRED 2005*, 2005, pp. 1–5.

[11] C. Cheng and D. Shirmohammadi, "A three-phase power flow method for real-time distribution system analysis," *Power Systems, IEEE Transactions on*, vol. 10, no. 2, pp. 671 –679, may 1995.

[12] L. Degroote, B. Renders, B. Meersman, and L. Vandevelde, "Neutral-point shifting and voltage unbalance due to single-phase DG units in low voltage distribution networks," in *Proc. IEEE Bucharest PowerTech*, 2009, pp. 1–8.

[13] V. Arya, D. Seetharam, S. Kalyanaraman, K. Dontas, C. Pavlovski, S. Hoy, and J. R. Kalagnanam, "Phase identification in smart grids," in *Proc. IEEE Int Smart Grid Communications (SmartGridComm) Conf*, 2011, pp. 25–30.

[14] M. Dilek, "Integrated design of electrical distribution systems: Phase balancing and phase prediction case studies," Ph.D. dissertation, Blacksburg Virginia, 2001.

[15] Caird, "Meter phase identification," US patent application Patent 20 100 164 473, 2010.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Mar. 2004.

IX. Biographies

Sam Weckx (S'11) received the M.Sc. degree in Electrical Engineering in 2009 and Mechanical Engineering in 2010 from the Katholieke Universiteit Leuven (KU Leuven), Belgium, where he is currently working towards the Ph.D. degree as a research assistant with the division ESAT-ELECTA. His research interests include the application of distributed optimization in smart grids and voltage control in distribution networks.

Carlos Gonzalez received his M. Sc in Electrotechnics in Polytechnic University of Catalonia (UPC, Barcelona Tech) in 2009. From 2006 till 2010, he was researcher in Citcea - UPC and IREC, Barcelona. Since 2010, he is Ph.D. student and research assistant in ESAT-ELECTA in the Katholieke Universiteit Leuven (KULeuven), Belgium. His research interests include power quality in smart grids and fault detection.

Jeroen Tant (S'08) received the M.Eng. degree in Electronics-ICT from the Hogeschool West-Vlaanderen, Kortrijk, Belgium, in 2007 and the M.Sc. degree in Mathematical Engineering from the Katholieke Universiteit Leuven (K.U.Leuven), Leuven, Belgium, in 2009. He is currently working as a Research Assistant at K.U.Leuven with the division ESAT-ELECTA. His research topics include simulation and modelling of power electronic systems and deterministic and stochastic numerical methods for the analysis of power distribution grids.

Tom De Rybel received the Industral Engineer degree in electronics design from Hogeschool Gent, Belgium, in 2002 and the M.A.Sc and PhD degrees in power systems from the University of British Columbia, Vancouver, Canada, in 2005 and 2010, respectively. His research interests, as a post-doctoral fellow at KU Leuven, Belgium, include smart-grid component design, high-voltage instrumentation, asset condition monitoring, power electronics, numerical acoustics, and hardware-in-the-loop simulation."

Johan Driesen (S93 - M97 SM12) received the M.Sc. and Ph.D. degrees in Electrical Engineering from the Katholieke Universiteit Leuven (KU Leuven), Leuven, Belgium, in 1996 and 2000, respectively. Currently, he is a Professor with the KU Leuven and teaches power electronics and electric drives. In 2000, he was with the Imperial College of Science, Technology and Medicine, London, U.K. In 2002, he was with the University of California, Berkeley. Currently, he conducts research on distributed generation, power electronics, and its applications