A Privacy Preserving Approach to Smart Metering

Merwais Shinwari[1], Amr Youssef[1], and Walaa Hamouda[2]

[1] Concordia Institute for Information Systems Engineering (CIISE)
[2] Electrical and Computer Engineering Department
Concordia University
Montreal, Canada
(m_shinwa,youssef)@ciise.concordia.ca, hamouda@ece.concordia.ca

Abstract. High frequency power consumption readings produced by smart meters introduce a major privacy threat to residential consumers as they reveal details that could be used to infer information about the activities of home occupants. In this paper, we question the need to disclose high frequency readings produced at the home's level. Instead, we propose equipping smart meters with sufficient processing power enabling them to provide the utility company with a set of well-defined services based on these readings. For demand side management, we propose the collection of high frequency readings at a higher level in the distribution network, such as local step-down transformers, as this readily provides the accumulated demand of all homes within a branch. Furthermore, we study the effect of the proposed approach on consumers' privacy, using correlation and relative entropy as measures. We also study the effect of load balancing on consumers' privacy when using the proposed approach. Finally, we assess the detection of different appliances using high frequency readings collected for demand side management purposes.

Keywords: Smart Grid, Smart Meter, Privacy, Advanced Metering Infrastructure (AMI).

1 Introduction

The electric grid in use today is undergoing a transformation to improve its efficiency and reliability through the use of computation and communication technologies. A major part of this is the enhancement of the distribution network through the introduction of smart meters. These meters collect power consumption readings and transmit them to the utility company in an automated way. Unlike traditional meters, smart meters produce detailed chronological high frequency readings that reveal both the time of consumption and the amount of power consumed. Utility companies argue that high frequency readings form the basis for time of use billing schemes, which are expected to cause consumers to shift part of their consumption to off-peak hours resulting in a flat demand profile (i.e., one with a small peak-to-average ratio) and improving energy production and consumption efficiencies [1]. Furthermore, it is argued that a clearer vision of the distribution network helps in improving service quality and reliability. For example, a utility company could centrally detect and respond to

A.I. Awad, A.E. Hassanien, and K. Baba (Eds.): SecNet 2013, CCIS 381, pp. 164–176, 2013.

blackouts and brownouts more effectively in comparison to user initiated notifications [2]. Because of this, smart metering is viewed as the fundamental platform that facilitates service enhancement, reliability, and efficiency and is considered an enabler for technologies such as proactive energy consumption management and load balancing techniques [3] [5].

By their nature, electrical appliances consume power in specific patterns which produce detectable signatures. For example, a standard incandescent lamp constantly consumes a fixed amount of power during its operation period, whereas a refrigerator consumes most of its power during its cooling cycles, when the compressor is running, and significantly less power during its idle cycles. Such patterns can be used to produce signature libraries which can be used for appliance detection and identification [6] [7].

Given a library of power consumption signatures of appliances, and the detailed power consumption of a home, this home's consumption can be decomposed and individual appliances can be detected using Nonintrusive Appliance Load Monitoring technologies (NALM) [8]. Fig. 1 shows an example of appliance detection using power signatures. As shown, many appliances can be identified through their distinct power consumption patterns. That is, with NALM technologies, high frequency readings produced by smart meters offer a window into the activities of homes' occupants. This includes the identification of appliances and any other information possibly inferable from the appliances used. Furthermore, by observing the real time power consumption of a given home, an intruder can identify when the occupants are awake/asleep or whether the home is occupied or not [9].

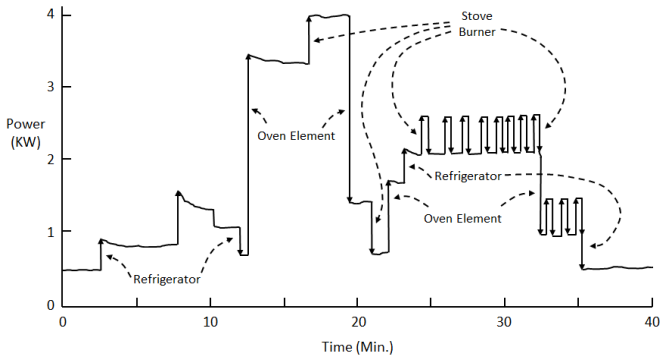

Fig. 1. Appliance identification through power consumption patterns [8]

Although high frequency readings produced by smart meters enable improving the efficiency of the electric grid, they introduce a privacy threat that was not present in the classical grid. In this paper, we present an alternative approach to smart metering with the objective of maintaining its advantageous functionality while preserving consumers' privacy. In the following section, we briefly review some related work in this area. Our approach to smart metering is explained in section 3. In section 4, we assess the privacy gained from the proposed approach. Finally, in section 5, we present our conclusion.

2 Related Work

The privacy impact of collecting power consumption readings in high frequency is well known and widely studied. Reports such as [10], [11], and [12] indicate that the privacy concerns of smart metering must be taken into consideration and addressed at the design stage rather than as a later addition. Furthermore, many researchers have proposed various approaches to address this problem. In this section, we present a selection of the main contributions in this area highlighting the core ideas proposed.

Kalogridis et al. [13] propose masking the power signature of appliances using a rechargeable battery. In particular, the authors propose the use of an energy routing device that controls power flow from the grid to the home and from/to a rechargeable battery following a water-filling based algorithm. This either charges or discharges the battery in a way that masks some details of the home's power demand. Additional work presented in [14] attempts to quantify the privacy offered by the battery solution, and concludes that privacy preservation increases as the battery size gets larger. Although this approach does mask part of the consumption profile, introducing a large rechargeable battery and a power routing device presents a hindrance to consumers. Furthermore, this solution does not offer as much privacy as consumers had before the introduction of smart meters. For example, information deducible from a home's general consumption pattern such as "when did the occupants wake up" or "is the house occupied" can still be attained even with the deployment of this solution.

Efthymiou and Kalogridis [15] argue that although high frequency readings may be needed for operational purposes, there is no need to attribute them to specific consumers. Consequently, the authors propose the use of two sets of readings: one in high frequency, and the other in low frequency. The high frequency readings are to be collected anonymously with the help of an escrow service and are provided to the utility company. Since the consumers' identities are not associated with these records, the consumption and usage characteristics cannot be traced to a specific consumer. The lower frequency readings are to be bounded to their respective consumers and used for billing purposes. Since these do not capture detailed power consumption information, they are not a threat to consumers' privacy. Although this method may seem effective, the use of an escrow service simply transfers the trust problem from the utility company to the escrow service provider, and therefore, does not provide a fundamental solution to the original problem.

In [16], Tomosada and Sinohara propose that smart meters transmit synthetically produced data that shares the same statistical properties of the real readings instead of transmitting the readings themselves. The authors argue that since this virtual demand shares the same statistical properties with the real demand, it can be used for demand side management when averaged over multiple users. In their work, the authors propose a methodology for producing virtual demand from the real demand and conclude that this approach preserves the consumer's privacy. Although this method produces correct statistics, other characteristics critical to demand side management could be lost, for example, the peak value and the time at which this peak value occurs.

3 Proposed Metering Approach

Smart meters are typically used as distributed data acquisition devices. That is, the meters only produce and transmit high frequency readings to the utility company. The utility company, in turn, centrally processes this data producing bills for its subscribers based on the time power was consumed. Fig. 2 illustrates this view of smart meters functionality.

With this approach, high frequency readings are present at the meter, in transmission and in storage at the utility company's processing facility. Having this data at all these points maximizes the potential attack surface for an attacker. This way, the attacker needs to identify some vulnerability in any of these points to be able to access the detailed consumption records. Furthermore, if an attacker is able to identify and exploit some vulnerability at the central processing facility, the impact would be devastating as hundreds of thousands of records could be compromised in a single breach. Verifiably securing large distribution networks, communication networks and processing facilities is practically infeasible. Furthermore, an attempt to secure such interconnected systems would be a tedious task that is almost impossible to implement flawlessly.

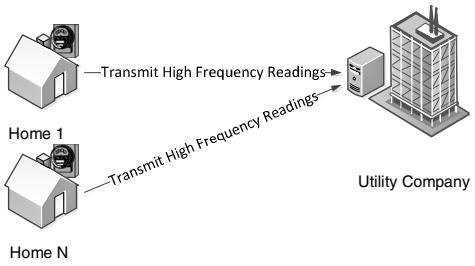

Fig. 2. Smart meters as data acquisition devices. Meters produce and transmit high frequency readings to the utility company

For the purpose of our work, we categorize consumer-oriented data collected by the utility company into two basic types, namely subject data and community data. We define subject data as data collected from and identifiable to a single consumer. We assume that actions taken based on this type of data will only affect its associated consumer. On the other hand, we define community data as data identifiable to a group rather than a single consumer. Furthermore, we assume that actions taken on a large scale, i.e., on the community as a whole, are based on this type of data.

We propose the deployment of smart meters in a way that segregates these two types of data offering each an appropriate level of protection. Therefore, from a privacy perspective, subject data would have a higher level of protection in comparison to community data. To do so, we propose the use of two sets of meters positioned at different locations in the distribution network, namely home meters and zone meters.

3.1 Home Meters

We propose that home meters function as service providing modules rather than data acquisition devices. That is, assuming that each meter is equipped with sufficient

processing power, the meters are to offer the utility company a set of well-defined services computed over the consumer's high frequency readings. In addition, home meters are not to disclose the collected high frequency readings to any party. Furthermore, the services offered by the meter must be developed on a need to know basis.

With this approach, the meters become the entities that perform all required processing on their respective consumers' data and only the outcomes of the processing, i.e., the final results, are made available to the utility company. This allows home meters to provide the desired functionality while eliminating the need to disclose users' high frequency readings, consequently, preserving the consumers' privacy. Furthermore, this introduces a point of control on the type of information the utility company gains access to.

The services provided by a home meter would depend on the protocols/functions it implements. For example, for billing purposes, meters would implement a billing protocol that starts by receiving an authenticated request from the utility company to produce the consumer's bill for a given period. The meter, in turn, uses the high frequency readings from its internal storage to compute the amount owed in dollars based on a pre-agreed upon pricing scheme. The final result of the process would be encrypted and digitally signed by the meter and transmitted to the utility company. This would provide the utility company with the desired information ensuring that it was produced by the meter.

Besides billing, other useful functionalities can be easily implemented. For example, meters could report their operation status or fault codes by periodically transmitting a status message that can be protected using cryptographic techniques. Fig. 3 illustrates the use of home meters as service providing modules.

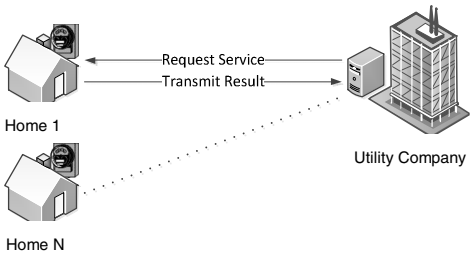

Fig. 3. Home meters as service providing modules. Utilizing the high frequency readings, home meters offer the utility company a set of well-defined services on a need to know basis.

It should be noted that, in our proposed approach, home meters must be trusted as they are the only components that hold and directly process their respective consumer's high frequency power consumption readings. Since a meter is a relatively small device that implements a limited set of protocols, it is feasible to build verifiably secure meters. This can be achieved through the use of trusted computing platforms [17] [18], and the use of formally verifiable designs.

With this approach, since each consumer's data is only present at a single point, the attack surface is significantly reduced. That is, if an attacker is successful in penetrating

a given meter, only the records belonging to the associated customer would be compromised. This is a significant security advantage in comparison to the use of meters as distributed collection nodes and centrally processing the data collected.

3.2 Zone Meters

As stated above, home meters do not disclose unprocessed measurements of power consumption; rather they provide specific information computed from this data. Even though it is possible to produce accurate time of use based consumption bills, information on consumption trends and the time of power consumption would be masked. This introduces a hindrance to the operations of utility companies as such information is aggregated for demand side management. To address this, we propose the use of an additional small set of meters placed at a higher level in the distribution network; typically at local step-down transformers.

Instead of securely producing accumulative readings through an escrow service or using cryptographic approaches, we take advantage of the already existing topology of the distribution network. By observing that the demand at a step down station is the accumulation of the demand of all homes supported by this station, measuring the power consumption at this level is equivalent to aggregating the power demand of individual homes in this zone. Therefore, by collecting measurements at this location using zone meters, we readily obtain the accumulative demand of all homes within a given branch. Fig. 4 illustrates the use of two sets of meters at the home and the zone level.

Although high frequency readings produced by zone meters are not directly attributable to a single consumer, they produce readings of the composite demand of all homes supported by their branch which is a function of the consumption of each home. In the following section we analyze the privacy impact of the proposed approach and the visibility of a home's demand through readings produced by zone meters. We also specifically consider the impact of load balancing on consumers' privacy when using the proposed approach.

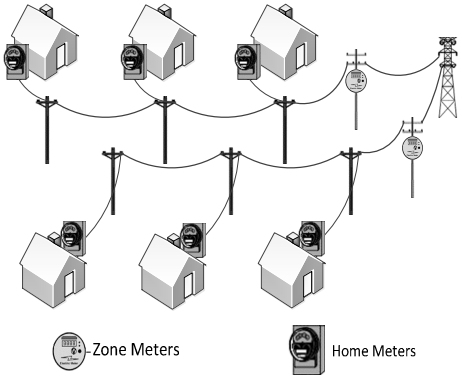

Fig. 4. The use of meters at different levels in the distribution network. Home meters offer a set of well-defined services at the consumer level whereas zone meters produce high frequency readings at the neibghorhood level.

4 Privacy Assessment

In our proposed approach, high frequency readings produced by home meters are not disclosed to the utility company or any other party. This eliminates the direct privacy impact introduced by these readings. Although zone meters introduce indirect threats, the superposition of signatures from different homes introduces an obfuscation effect on individual signatures which lessens the disclosed information. In this section, we assess the obfuscation gained by overlapping home signatures. The number of homes supported by a single zone meter is a main factor in the level of signature overlap; therefore we consider this as a primary factor in our simulation. Furthermore, since time of use billing schemes are expected to flatten the overall demand of a community which affects the overlapping of signatures, we take this into consideration as well.

4.1 Simulation Environment

We produced our simulation environment using a set of appliances with distinct power signatures similar to [8] and using measurements from [19]. Each home is allocated a set of appliances and the operation time of each appliance is selected randomly. The simulation is conducted over a period of 24 hours. Furthermore, to assess the impact of load balancing, two sets of results are produced for each simulation scenario. The first represents the case where power is consumed at will. This type of consumption results in the appearance of peak demand hours as is the case with the classical power grid [20]. The second represents the case where consumers shift part of their consumption to off-peak hours using load balancing techniques such as those described in [3], [4], and [5]. This results in a relatively flat consumption profile for the community as a whole. Fig. 5 shows the simulated power demand for a sample home. As depicted in the figure, the consumption patterns of many appliances can be easily identified.

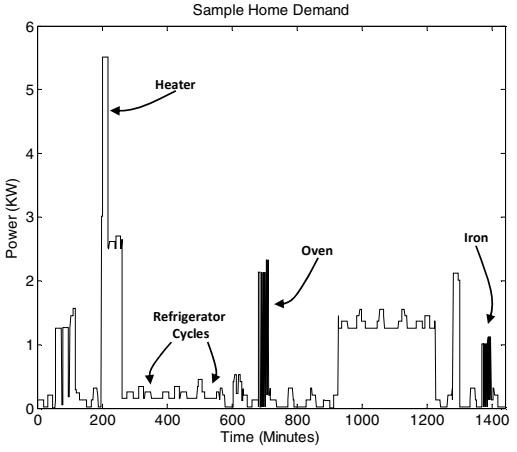

Fig. 5. Simulation of the demand of a single home with a sampling interval of one minute

To simulate the readings produced by zone meters, a variable number of homes are simulated and accumulated. Fig. 6 illustrates the aggregate demand of a community of 50 homes. As depicted in the figure, the overlapping of signatures of different homes distorts the appliance signatures. The figure also reflects the effect of load balancing on the overall consumption of the community. As shown, the use of load balancing results in a flatter overall demand with a lower peak to average ratio. This results in a more uniform level of overlapping between appliance signatures.

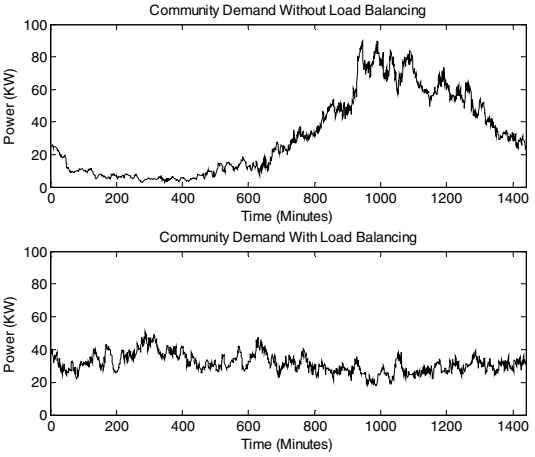

Fig. 6. Community demand without load balancing (top), and with load balancing (bottom)

Fig. 7 shows the decrease in correlation between a home's demand and the overall community consumption as the number of homes in the community is increased. The decreasing trend in the figure can be attributed to the distortion caused to the demand signature from other homes. As the number of homes increase, so does the level of distortion. Furthermore, the figure shows that the use of load balancing techniques further reduces the correlation value resulting in better privacy. This is because load balancing results in even overlapping and eliminates the low overlap between signatures during low demand hours.

Fig. 8 shows the correlation value in the average case for each home when simulating a community of 50 homes. The results indicate that the signatures of all simulated homes were distorted to a similar level.

As an assessment of the difference between the probabilistic distributions in the data sets, we compute the Kullback Leibler divergence (also known as the relative entropy) while increasing the number of homes in the community. This is a well known information theoretic measure that can be used to quantify the relationship between two signals. Given two signals with probability distributions P and Q, the Kullback Leibler divergence can be defined as:

$$D(P||Q) = \int_{-\infty}^{\infty} p(x) \ln \frac{p(x)}{q(x)} \, dx \tag{1}$$

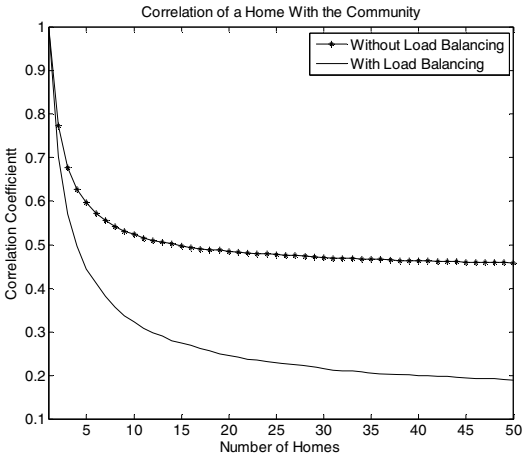

Fig. 7. Correlation between a home's demand and the community demand as produced by zone meters. Results presented are the average case of 100 iterations while increasing the number of homes in the community up to 50.

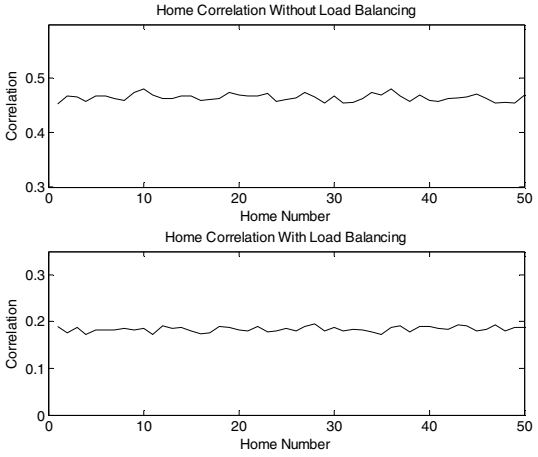

Fig. 8. Correlation of each simulated home with its community. All homes' signatures were distorted to a similar level. Results presented are the average case of 100 iterations for a community of 50 homes.

As shown in Fig. 9, the value of the Kullback Leibler divergence is zero for a single home, indicating identical distributions, and grows rapidly to saturate when accumulating about 10 homes. This indicates that accumulating a relatively small number of homes would have a good effect on masking individual homes' consumption profiles. Furthermore, as depicted in the figure, the use of load balancing helps achieve better privacy protection.

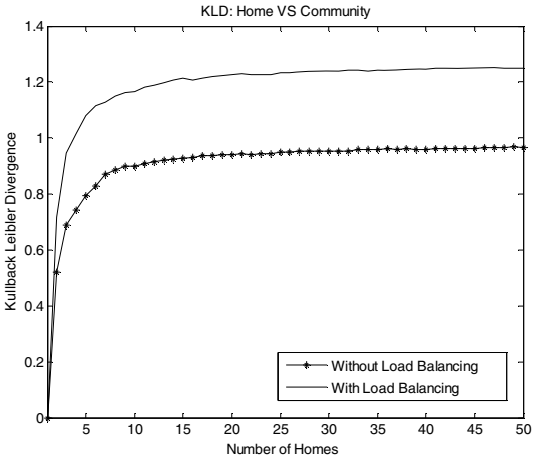

Fig. 9. KLD of a home's signature and the community demand as visible through zone meters. Results presented are the average case of 100 iterations while increasing the number of homes up to 50.

4.2 Appliance Detection

Assuming the availability of a library of appliance signatures [19], we assess the detection of the operation of appliances using cross correlation from readings produced by zone meters while increasing the number of homes in the community. Unlike previous simulations, the appliance to be detected is only run once by one member of the community. This ensures that a false detection is not caused by a duplicate signature of the appliance in question.

This case was simulated allowing the target appliance (i.e., the one to be detected) to be turned on randomly, following a uniform distribution throughout the 24 hour simulation interval with a step size of one minute; i.e., a total of 1440 possible time slots. We define a correct detection as one where the precise time slot was identified.

Fig. 10 shows the percentage of correct detections for a sample appliance as the number of homes in the community is increased. As the figure shows, the use of load balancing techniques causes a more rapid deterioration in detection accuracy, i.e., it achieves better privacy protection. Fig. 11 shows the error in detection, in time slots, as a function of the number of homes in the community. As shown, the detection becomes more distant from the real start time as the number of homes is increased. Furthermore, the use of load balancing causes an increase in the detection error, which implies better preservation of privacy.

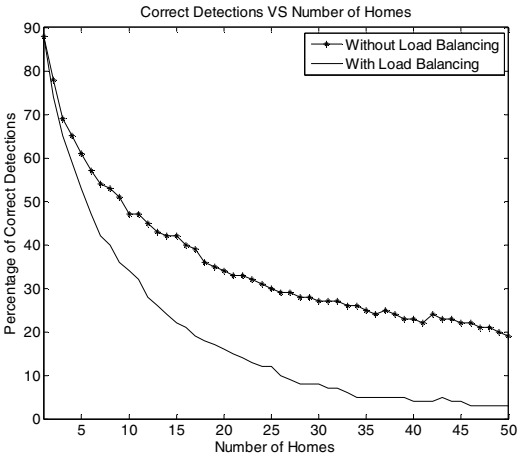

Fig. 10. Detection of a sample appliance using cross correlation deteriorates as the number of homes in the community is increased. Results presented are based on 100 detection attempts while increasing the number of homes up to 50.

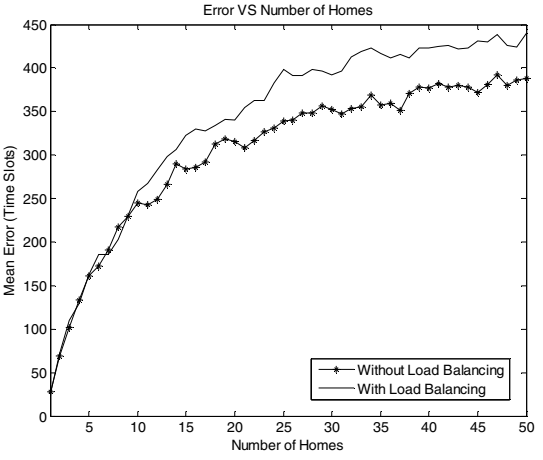

Fig. 11. The mean error in appliance signature detection increases as the number of homes is increased. Results presented are based on 100 detection attempts while increasing the number of homes up to 50.

5 Conclusion

In this paper, we showed that it is possible to achieve the objectives of smart metering without compromising the privacy of residential consumers. In our approach, home meters function as service providing modules rather than data acquisition devices allowing them to provide the desired functionality, such as time of use billing, while

eliminating the need to disclose users' high frequency readings to the utility company. Collecting high frequency readings is done at a higher level such as at the local step-down transformer which allows utility companies to achieve their operational objectives. While our approach requires home meters to be trusted, this is more feasible than attempting to secure the meters, the transmission network, and all the processing facilities. Furthermore, we also showed that our approach achieves better privacy protection when consumers opt to use load balancing techniques. Our results imply that, with the proposed approach, the more efficient operation of the grid can result in better privacy protection for individual customers.

References

1. The Ontario Smart Metering Initiative, http://www.consumerscouncil.com/site/consumers_council_of_canada/assets/pdf/SM_Report.pdf
2. Collier, S.E.: Ten steps to a smarter grid. In: Proc. IEEE Rural Electric Power Conference, REPC 2009, pp. B2–B7 (April 2009)
3. Caron, S., Kesidis, G.: Incentive-based energy consumption scheduling algorithms for the smart grid. In: Proc. First IEEE International Conference on Smart Grid Communications, SmartGridComm, pp. 391–396 (2010)
4. Chen, C., Kishore, S., Snyder, L.V.: An innovative RTP-based residential power scheduling scheme for smart grids. In: Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, pp. 5956–5959 (2011)
5. Mohsenian-Rad, A.H., Wong, V., Jatskevich, J., Schober, R.: Optimal and autonomous incentive-based energy consumption scheduling algorithm for smart grid. In: Proc. IEEE PES Conference on Innovative Smart Grid Technologies, pp. 1–6 (2010)
6. Liang, J., Ng, S.K.K., Kendall, G., Cheng, J.W.M.: Load signature study—part i: basic concept, structure, and methodology. IEEE Transactions on Power Delivery 25, 551–560 (2010)
7. Liang, J., Ng, S.K.K., Kendall, G., Cheng, J.W.M.: Load signature study—part ii: disaggregation framework, simulation, and applications. IEEE Transactions on Power Delivery 25, 561–569 (2010)
8. Hart, G.W.: Nonintrusive appliance load monitoring. Proceedings of the IEEE 80(12), 1870–1891 (1992)
9. Quinn, E.: Privacy and the new energy infrastructure. Working Paper Series (2009), http://ssrn.com/abstract=1370731
10. Cavoukian, A.: Privacy by design: achieving the gold standard in data protection for the smart grid, http://www.ipc.on.ca/images/resources/achieve-goldstnd.pdf
11. Cavoukian, A.: Smart privacy for the smart grid: embedding privacy into the design of electricity conservation, http://www.ipc.on.ca/images/resources/pbd-smartpriv-smartgrid.pdf
12. Cavoukian, A.: Operationalizing privacy by design: the Ontario smart grid case study, http://www.ipc.on.ca/images/Resources/pbd-ont-smartgrid-casestudy.pdf

13. Kalogridis, G., Efthymiou, C., Denic, S.Z., Lewis, T.A., Cepeda, R.: Privacy for smart meters: towards undetectable appliance load signatures. In: Proc. IEEE Smart Grid Commun. Conf., Gaithersburg, Maryland, pp. 232–237 (2010)
14. Kalogridis, G., Zhong, F., Basutkar, S.: Affordable privacy for home smart meters. In: Proc. IEEE Int. Workshop Smart Grid Security Commun., SGSC, Busan, Korea, May 26-28, pp. 77–84 (2011)
15. Efthymiou, C., Kalogridis, G.: Smart grid privacy via anonymization of smart metering data. In: Proc. IEEE Smart Grid Commun. Conf., Gaithersburg, Maryland, pp. 238–243 (2010)
16. Tomosada, M., Sinohara, Y.: Virtual energy demand data: estimating energy load and protecting consumers' privacy. In: Proc. 2011 IEEE PES Innovative Smart Grid Technologies, ISGT 2011, Medellin, Colombia, pp. 1–8 (2011)
17. Mitchell, C. (ed.): Trusted computing. Institution of Electrical Engineers (2005)
18. Pearson, S.: Trusted computing platforms, the next security solution. HP Labs (2002)
19. Richardson, I., Thomson, M., Infield, D.: A high-resolution domestic building occupancy model for energy demand simulations. Energy and Buildings 40(8), 1560–1566 (2008)
20. Ontario Demand and Market Prices. The Independent Electricity System Operator, http://www.ieso.ca/