Evaluation of Utility-Privacy Trade-Offs of Data Manipulation Techniques for Smart Metering

Huan Yang*, Liang Cheng† and Mooi Choo Chuah‡
Department of Computer Science and Engineering, Lehigh University, Bethlehem, PA 18015
Email: *huy213@lehigh.edu, †cheng@cse.lehigh.edu, ‡chuah@cse.lehigh.edu

***Abstract*—Nowadays, smart meters are capable of collecting, processing, storing, and reporting users' energy consumption data with high fidelity. While facilitating a wide variety of applications such as demand response and anomaly detection, smart meter data exposes users to the risks of revealing sensitive information including occupancy states and activity patterns. Among privacy-preserving data collection methods for smart metering, data manipulation techniques can easily be implemented in smart meters and do not require installation of batteries. However, the trade-offs between data utility and privacy leakage supported by these techniques are not fully explored. In this work, we identify appliance-level utility as well as privacy metrics and quantitatively evaluate the effectiveness of two data manipulation techniques: down-sampling and noise addition. Furthermore, we combine the two techniques and evaluate the levels of appliance-specific information leakage supported under miscellaneous configurations. Using publicly available data set, we show that data manipulation techniques can provide coarse-grained protection of users' privacy while enabling applications requiring appliance-level information such as fault detection. Our evaluation reveals the limitations of these data manipulation techniques, which can help users determine whether investments on more advanced solutions such as battery-based load hiding are necessary.**

I. Introduction

As hardware technology and the smart grid paradigm continue to evolve, capabilities of smart meters have been enhancing over the last decade. Many existing smart meters are equipped with sensing, data processing, storage, and communication modules to report whole-home energy consumption data to utility service providers. In addition, data disclosure policies, such as the Green Button [1], allow users to download their own data or authorize a trusted third party to access it. In [2], the Green Button framework is enhanced to (a) allow utility customers to control the amount of data disclosed, and (b) enable third-party service providers to verify consumers' data authenticity. When energy usage data is disclosed, users have to face the trade-off between the utility of data and the risk of privacy leakage: On the one hand, smart meter data is leveraged in an ever-growing array of applications that can benefit the users, including theft detection [3], anomaly detection [4], and demand response [5]. For instance, abnormal changes in energy consumption can be identified by computing an anomaly score for an individual user and then adjusting it based on smart meter data from the neighborhood [4]. The user is exempted from scrutinizing all the data and can focus on the snippets containing anomalies. On the other hand, the development of non-intrusive load monitoring (NILM) techniques [6] that disaggregate whole-home energy readings into appliance-specific energy usages has led to privacy concerns. As illustrated in Fig. 1, further inferences on appliance-level energy data can reveal sensitive information such as occupancy states [7], user activity patterns [8], or even multimedia contents being played on a TV set [9]. While providing useful services by applying NILM or more advanced data analytics, a trusted data curator (e.g., the utility company, or a third party granted data access by the user) is able to extract sensitive information that a user does not want to disclose. As an example, suppose that a utility company collects smart meter readings to support time-based rate programs and fault detection. Meanwhile, the algorithm in [9] is applied by the utility company to infer the multimedia contents being played on a user's TV set. Over time, knowledge on users' preferences for TV programs will be accumulated and may be sold to advertisement companies without their consent. To facilitate the development of innovative services leveraging smart meter data while limiting the risks of privacy leakage, it is vitally important to not only establish effective regulations and guidelines [10] but also devise privacy-preserving mechanisms.

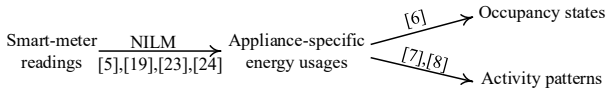


Fig. 1. NILM algorithms and sensitive information that can be extracted from appliance-specific energy usages.

Several categories of methods have been proposed to protect user privacy in smart metering settings [11]. Encryption-based techniques (e.g., [12]) can prevent the extraction of sensitive information by untrusted parties and external attackers. However, these approaches cannot prevent trusted data curators from performing unauthorized data analytics. Battery-based load hiding (BLH) techniques (e.g., [13]) require the installation of batteries to hide or flatten load curves that may reveal sensitive information. Many BLH schemes formulate and solve optimization problems to minimize privacy revelation while achieving objectives of the user and/or the utility company (e.g., [13], [14]). As data is collected and stored in smart meters before being reported to the utility company, it is possible to protect users' privacy through data manipulation. For instance, adding noise to smart meter readings [15] can protect privacy while still support useful applications such as billing. In contrast to privacy-preserving methods from other categories, data manipulation techniques can be implemented

in smart meters through firmware upgrades and do not require extra investments on batteries. However, the utility-privacy trade-offs offered by data manipulation techniques under different settings have not been fully explored. Specifically, we are interested in finding answers to the following questions that may be raised by users and utility companies:

- How well can data manipulation techniques prevent leakage of appliance-level energy consumption information?
- When are investments on BLH techniques necessary to protect privacy?

In this work, we quantitatively explore the trade-offs between privacy leakage and utility of smart meter data enabled by two data manipulation techniques: down-sampling and noise addition. The contributions of this work are as follows: We identify quantitative metrics to measure per-appliance data utility and privacy leakage, facilitating the comparisons both between data manipulation techniques and with BLH schemes. In addition, we quantitatively explore the trade-offs enabled by down-sampling and noise addition, respectively. Furthermore, we evaluate the effectiveness of the combination of these two techniques in privacy protection. Our experiments show that evaluating the utility-privacy trade-offs at appliance level allows users to find out whether sensitive information they do not want to disclose can be protected by implementing data manipulation techniques in smart meters. Our evaluation framework can also help users determine whether or not they need to invest on BLH solutions.

II. Related Work

In addition to securing the communication channel between smart meters and a trusted data curator [12], a category of techniques mask load curves that may reveal sensitive information through load scheduling. If batteries are installed on household premises, users can hide private information, such as appliance usages, occupancy states, habits, and lifestyles, by controlling when and how batteries should be charged and discharged (e.g., [13], [14]). Alternatively, it is also possible to schedule deferrable loads to prevent privacy leakage. For instance, it is shown in [16] that occupancy detection can be thwarted by modulating the power usage of a water heater. Privacy-preserving techniques based on load scheduling rely on the existence of energy sources or deferrable loads on a user's premise. In addition, they also require the deployment of control and communication infrastructure. Users may not want to invest on these load-scheduling-based solutions if methods with lower costs are available.

Privacy-preserving data manipulation techniques allow users to release their data to a trusted data curator while limiting the level of privacy leakage. In [15], noise is added to smart meter readings before they are reported. To ensure data utility, the range of the noise signal is controlled based on the maximum allowed error tolerance of the application to be implemented (e.g., a maximum error of 1% for monthly billing). It should be noted that once noise is added, applications requiring high-fidelity meter readings (e.g., load forecasting and time-based rates) can no longer be supported [15]. Noise addition can also be integrated into data aggregation framework to provide a trade-off between data precision and level of privacy [17]. In essence, smart meter readings are time-series data sampled at regular intervals. To protect temporal privacy from being learned, it has been proposed that smart meter readings can be buffered and disordered before transmission [18]. Another property of time-series data is the sampling rate. In [19], the authors study the impacts of sampling rate on the accuracy of edge detection, which is the first step of many NILM algorithms such as the first one proposed in [20].

In [21], a stochastic process, referred to as privacy mapping, is designed to distort smart meter readings while balancing between privacy and data utility. Such a privacy mapping can be implemented by load-scheduling-based techniques. For data manipulation techniques, the trade-off between hiding sensitive information and maintaining usefulness of smart meter data is not fully explored. In [15], the effectiveness of noise addition in privacy-protection is studied, whereas data utility is only discussed by listing the set of applications that can still be implemented after noise is added. A general privacy-utility framework has been proposed [22] to find optimal privacy-preserving strategies for the mapping from measured to released data. However, the usefulness of down-sampling and noise addition in balancing between data utility and privacy revelation remains unexplored. Finding a way to address this problem can help users choose between low-cost data manipulation techniques and load-scheduling-based solutions.

III. Privacy-Utility Framework for Evaluating Data Manipulation Techniques

A user's smart meter collects whole-home energy consumption data at regular time intervals and transmits the data to a trusted entity. We assume that the smart meter has a modular design (e.g., [23]) that allows the utility company to configure the sampling rate of the sensing module and/or implement noise addition [15] through firmware upgrades. We are interested in the following scenario that requires balancing between data utility and privacy leakage: Applications implemented by the trusted data curator rely on the energy usage information of a subset of a user's appliances. The user does not want to disclose information of appliances (or a subset) that are not required by the services he/she subscribes. In this section, we describe our framework for evaluating the utility-privacy trade-offs enabled by low-cost data manipulation techniques.

A. Adversary Model

The communication channel between user's smart meter and the trusted entity is secure, so the data collection system is protected against external attackers. We assume that any entity (e.g., the utility company) entrusted with a user's smart meter data is honest but curious: It correctly implements the services requested by the users but may extract extra information without their consent. We also assume that services offered by the trusted entity rely on the energy usages of a subset of

a user's appliances. Examples of such services include cost-saving recommendations and appliance fault detection. In cost-saving recommendations, a user participating in time-based rate program identifies a set of appliances (e.g., washer/drayer and dishwasher) and seeks recommendations on when to operate them to reduce the energy bills. In appliance fault detection, the user tells the trusted entity the set of potentially faulty appliances he/she owns and requests help to identify the ones to be replaced (e.g., a faulty oven that has been used for years). To enjoy such services, the user authorizes the trusted entity to disaggregate his/her smart meter data and analyze the energy consumptions of a subset of appliances. Appliance-level information that the user does not want to disclose is regarded as private. For example, a user receiving cost-saving recommendations on his/her dishwasher does not want the curious data curator to learn about how and when the kitchen outlets and bathroom ground-fault interrupter (GFI) are operated, which may allow the curious attacker to derive the user's activity patterns or lifestyle (see Fig. 1). In our evaluation, we assume that the trusted entity uses factorial hidden Markov models (FHMM) to disaggregate smart meter readings into per-appliance usage information. This algorithm has been used as a benchmark algorithm in NILM research (e.g., [24], [25]).

B. Data Utility Model

Smart meter data facilitates a wide variety of smart-grid services that provide utility to the users. The distortion introduced by a privacy mapping is used as a metric for data utility in [21]. Since we are interested in applications relying on appliance-specific energy usages (e.g., appliance fault detection), we define data utility for individual appliances. Suppose that a smart meter monitors a user's home during a time period T. Readings generated by the smart meter at regular time intervals is denoted by time series X^T. To implement the services requested by the user, a trusted entity applies an NILM algorithm to extract appliance-level energy usages. We denote the energy readings disaggregated from X^T for appliance i by $\hat{X_i^T}$. The actual energy readings for i during T is represented by X_i^T. Appliance-specific data utility is defined as follows:

Definition 1 (Data Utility Metric): Given two time series X_i^T and $\hat{X_i^T}$ for appliance i, the distortion between $\hat{X_i^T}$ and X_i^T can be measured by their distance $d(X_i^T, \hat{X_i^T})$. Suppose that there are N samples in X_i^T (and $\hat{X_i^T}$), we use the average distortion $\bar{d} = \frac{d(X_i^T, \hat{X_i^T})}{N}$ as the utility metric for i.

In this work, we use the Euclidean distance. Consider the appliance fault detection application. If the average distortion introduced by data manipulation exceeds some threshold (e.g., 10% of the power consumption range for the appliance of interest), the chance of false positives (i.e., normal operations mistaken for faults) increases. The greater the distortion, the less the utility. To measure data utility, we need ground-truth data on per-appliance energy consumption during T, which is available in public data sets such as the REDD [26] data set.

TABLE I
LIST OF APPLIANCES IN THE EXPERIMENTS

Index	Appliance	Index	Appliance
1	oven	2	refrigerator
3	dishwasher	4	kitchen-outlets-1
5	lighting-1	6	washer-dryer
7	microwave	8	bathroom-gfi
9	electric-heat	10	stove
11	kitchen-outlets-2	12	lighting-2
13	lighting-3		

C. Privacy Model

A user does not want information on how and when appliances closely related to occupancy or activity patterns (e.g., kitchen outlets, lighting, and electric heat) to be leaked if their information is not required by the services he/she subscribes. We use mutual information to evaluate the information leakage if the trusted data curator applies NILM to manipulated smart meter readings. We define per-appliance privacy metric using mutual information as follows:

Definition 2 (Privacy Metric): Given two time series X_i^T and $\hat{X_i^T}$ over the same time period T for appliance i, the mutual information $I(X_i^T, \hat{X_i^T})$ between the two series is

$$I(X_i^T, \hat{X_i^T}) = \sum_{x \in X_i^T} \sum_{y \in \hat{X_i^T}} \ln \frac{p(x,y)}{p(x)p(y)},$$

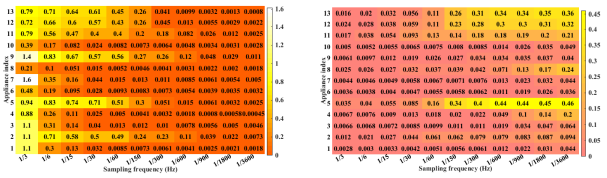
where $p(x)$ and $p(y)$ are the probability density functions of random variables $x \in X_i^T$ and $y \in X_i^T$, and $p(x,y)$ is the joint probability density function.

Mutual information quantifies the amount of common information embedded in both the ground-truth series X_i^T and the series $\hat{X_i^T}$ derived from manipulated smart meter readings. The greater the mutual information, the more the privacy leakage. Mutual information has been widely used in privacy-preserving techniques based on load scheduling (e.g., [27], [13], [22]) as well as data manipulation (e.g., [21]). We note that there are other privacy metrics such as correlation coefficient [15] and relative entropy [14]. We choose mutual information as privacy metric to facilitate comparisons not only among data manipulation techniques but also with load-scheduling-based schemes.

D. Data Manipulation Techniques

Among the data manipulation techniques reviewed in Section II, we choose to explore the utility-privacy trade-offs enabled by down-sampling and noise addition. Both techniques can be implemented in existing smart meters through firmware upgrades.

- *Down-sampling.* For smart meters with a modular design (e.g., [23]), down-sampling can be implemented in two ways. The utility company can directly lower the sampling rate of the front-end sensing module. Alternatively, it is also possible to down-sample stored meter readings X^T by summing up every n samples. The sampling rate of the resulting meter readings $\hat{X^T}$ is $\frac{1}{n}$ of X^T. Note that the sampling rate of $\hat{X^T}$ is constrained by the memory size of smart meters.
- *Noise addition.* The implementation of noise addition in smart meters is discussed in [15]. Noise is added after



(a) Mutual information in nats (b) Normalized distortion metric

Fig. 2. Utility-privacy trade-off enabled by down-sampling.

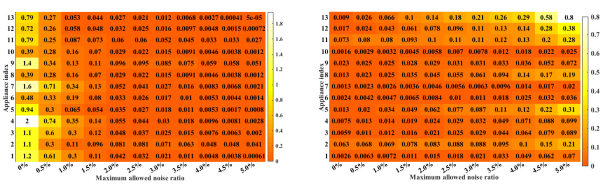
power measurements are taken by the sensing module. In this work, we assume that the random noise X_n^T has a uniform distribution, i.e., $X_n^T \sim U(-R, R)$. If X^T contains L samples and L is large enough, the central limit theorem ensures that the error e_n^T introduced by noise addition has a normal distribution, i.e., $e_n^T \sim N(0, \frac{L \cdot R^2}{3})$. We use the jumping daily window (JDW) approach to determine the maximum allowed error because it is shown to be the best noise addition strategy in [15]. We choose a confidence level of 98% for e_n^T to implement noise addition.

It is possible to combine these two techniques: We can first down-sample X^T and then apply noise addition on the down-sampled time series $\hat{X^T}$.

IV. Experiments on Smart-Meter Dataset

A. Experiment Settings

We take out 13 appliances from House 1 of the REDD [26] data set. Table I lists these appliances and their indexes. The smart meter readings are the aggregation of the energy readings of the 13 appliances, i.e., $X^T(t) = \sum_{1 \le i \le 13} X_i^T(t), \forall t \in T$. The sampling rate of the REDD data set is approximately $1/3$ Hz. We divide the data set into a training set and a testing set. The training set contains per-appliance energy readings of 1 week, and the remaining 11 days of the data forms the testing set. We down-sample the training set to generate a group of training sets at different sampling rates. The FHMM algorithm is trained on these training sets to generate FHMM models for the House 1 at different sampling rates. For down-sampled smart meter readings (with or without noise added), the FHMM model at the corresponding sampling rate is used to perform disaggregation. To evaluate the data manipulation techniques, we generate three groups of testing sets by applying down-sampling, noise addition, and their combination, respectively. Each testing set contains only the whole-home smart meter readings. The curious attacker applies the FHMM algorithm to the testing set to infer the energy consumption of the 13 appliances. For appliance i, we use the ground-truth data X_i^T and the inferred time series $\hat{X_i^T}$ to compute the



(a) Mutual information in nats (b) Normalized distortion metric

Fig. 3. Utility-privacy trade-off enabled by noise addition.

utility and privacy metrics. To compute mutual information, we use a bin size of 50 W. We are interested in how well data manipulation techniques can prevent the leakage of sensitive appliance states. Simultaneously, the manipulated data should retain a certain utility level such that users can still enjoy their desirable services.

B. Evaluation Results

1) Down-Sampling: We down-sample the testing data set from $1/3$ Hz to the following sampling rates: $1/6$ Hz, $1/15$ Hz, $1/30$ Hz, $1/60$ Hz, $1/150$ Hz, $1/300$ Hz, $1/600$ Hz, $1/900$ Hz, $1/1800$ Hz, and $1/3600$ Hz. We can use the data in Fig. 2 to explore the trade-off between data utility and privacy leakage enabled by down-sampling. Suppose that a simple fault detection application uses down-sampled data.The user asks if the microwave is faulty or not, so the utility company performs NILM to obtain appliance-level energy usages. First, we determine the level of privacy protection. According to [13], user's privacy can be protected if a BLH scheme can lower the mutual information below 0.3 nats. Using Fig. 2a, we can find that a sampling rate lower than $1/150$ Hz reaches the same level of privacy protection. Then, we use Fig. 2b to see if the distortion level is tolerable by the fault detection application. Note that we normalize the distortion metrics over the average power consumption of individual appliances as the per-appliance distortion levels in Fig. 2b. We can use a low sampling rate (e.g., $1/900$ Hz) since the distortion level for the microwave is only 2.3% of its average power level of 2260 Watts. The disaggregated energy usage of the microwave does not significantly diverge from the actual usage ranges, so we are still able to detect faults such as a magnetron tube failure, which results in about a 2000-Watt decrease in power consumption. On the other hand, if the target appliances are lights, we cannot significantly reduce the sampling rate since the distortion level is too high (e.g., about 35% at $1/900$ Hz). The information leakage will also be higher and cannot be effectively limited. In this case, a user must choose load-scheduling-based techniques or combine down-sampling with noise addition for privacy protection. For existing smart metering system using low sampling frequencies (e.g., $1/1800$ Hz or $1/3600$ Hz), we can observe from Fig. 2a that user privacy in terms of appliance states is effectively protected.

2) Noise Addition: Using the $\frac{1}{3}$-Hz testing set, we add noise and generate testing sets with the following maximum allowed error ratios: 0.5%, 1.0%, 1.5%, 2.0%, 2.5%, 3.0%, 3.5%, 4.0%, 4.5%, and 5.0%. We can use the data presented in Fig. 3 to explore the utility-privacy trade-offs enabled by noise addition. Again, suppose that the user wants the utility company to help detect faults in his/her microwave. To reach a good privacy protection level (i.e., mutual information lower than 0.3 for all appliances), Fig. 3a suggests that the maximum allowed noise ratio should be at least 1.5%. If the microwave is the only appliance the user wants the utility to learn about through NILM, we can further increase the ratio to 5.0%

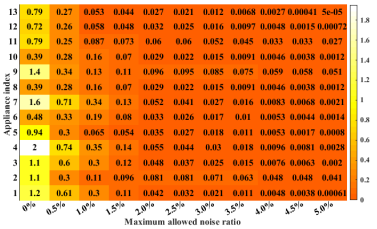


Fig. 4. Average privacy level enabled by the combination of down-sampling and noise addition in terms of mutual information in nats.

because Fig. 3b shows that this ratio only results in a 2% distortion for the microwave.

3) Combining Down-Sampling and Noise Addition: The group of testing sets are generated by first down-sampling the $\frac{1}{3}$-Hz testing set and then adding noise. Fig. 4 shows the mutual information metric averaged over the 13 appliances at each *(sampling rate, maximum allowed noise ratio)*-combination. Due to space constraint, we include per-appliance utility and privacy metrics for this scenario in our technical report [28]. At each sampling rate, the impacts of noise addition on per-appliance data utility exhibits similar patterns to what we observed in Fig. 3b. As shown in Fig. 4, the privacy level (in terms of average mutual information) is high. Suppose that the utility company needs meter readings with a relatively high sampling rate for conventional applications such as billing and that the user still wants the utility to help diagnose his/her microwave. Combining down-sampling and noise addition can help them find the proper trade-off between utility and privacy: At $1/6$ Hz with a 1.0% noise addition, the privacy levels for all appliances are lower than 0.3, whereas the distortion level for the microwave is about 2.9%. With this configuration, the utility is able to generate energy bills with an 1% error and perform fault detection for the microwave. Meanwhile, the level of privacy protection is low enough to protect the user.

V. Conclusion

In this work, we quantitatively evaluate privacy-preserving data manipulation techniques for smart metering. We show that users can balance between privacy revelation and data utility by choosing the proper data manipulation techniques. Although extra investments on batteries and control infrastructure are not required, we note that the granularity of control supported by these techniques is coarse. By evaluating utility-privacy trade-offs quantitatively, a trusted data curator will be able to help users choose between low-cost data manipulation techniques and BLH schemes for privacy protection.

References

[1] D. S. Sayogo and T. A. Pardo, "Understanding Smart Data Disclosure Policy Success: The Case of Green Button," in *14th Annu. Int. Conf. on Digital Government Research*, 2013, pp. 72–81.

[2] G. Lahoti, D. Mashima, and W.-P. Chen, "Customer-Centric Energy Usage Data Management and Sharing in Smart Grid Systems," in *1st ACM Workshop on Smart Energy Grid Security*, 2013, pp. 53–64.

[3] P. Jokar, N. Arianpoo, and V. C. M. Leung, "Electricity Theft Detection in AMI Using Customers' Consumption Patterns," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 216–226, 2016.

[4] P. Arjunan, H. D. Khadilkar, T. Ganu, Z. M. Charbiwala, A. Singh, and P. Singh, "Multi-User Energy Consumption Monitoring and Anomaly Detection with Partial Context Information," in *ACM Conf. on Embedded Systems for Energy-Efficient Built Environments*, 2015, pp. 35–44.

[5] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A Survey on Demand Response Programs in Smart Grids: Pricing Methods and Optimization Algorithms," *Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 152–178, 2015.

[6] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, "Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey," *Sensors*, vol. 12, no. 12, pp. 16 838–16 866, 2012.

[7] M. Jin, R. Jia, Z. Kang, I. C. Konstantakopoulos, and C. J. Spanos, "PresenceSense: Zero-Training Algorithm for Individual Presence Detection Based on Power Monitoring," in *1st ACM Conf. on Embedded Systems for Energy-Efficient Buildings*, 2014, pp. 1–10.

[8] J. Alcalá, J. Ureña, and Á. Hernández, "Activity Supervision Tool Using Non-Intrusive Load Monitoring Systems," in *2015 IEEE Conf. on Emerging Technologies Factory Automation*, 2015, pp. 1–4.

[9] U. Greveler, P. Glösekötterz, B. Justusy, and D. Loehr, "Multimedia Content Identification through Smart Meter Power Usage Profiles," in *Int. Conf. on Information and Knowledge Engineering*, 2012.

[10] "Guidelines for Smart Grid Cyber Security," National Institute of Standards and Technology (NIST), Tech. Rep. NISTIR 7628, 2010.

[11] S. Finster and I. Baumgart, "Privacy-Aware Smart Metering: A Survey," *Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1732–1745, 2014.

[12] F. Benhamouda, M. Joye, and B. Libert, "A New Framework for Privacy-Preserving Aggregation of Time-Series Data," *ACM Trans. Inf. Syst. Secur.*, vol. 18, no. 3, pp. 10:1–10:21, 2016.

[13] J. Zhao, T. Jung, Y. Wang, and X. Li, "Achieving Differential Privacy of Data Disclosure in the Smart Grid," in *IEEE Conf. on Computer Communications*, 2014, pp. 504–512.

[14] L. Yang, X. Chen, J. Zhang, and H. V. Poor, "Optimal Privacy-Preserving Energy Management for Smart Meters," in *IEEE Conf. on Computer Communications*, 2014, pp. 513–521.

[15] P. Barbosa, A. Brito, and H. Almeida, "Defending Against Load Monitoring in Smart Metering Data Through Noise Addition," in *30th Annu. ACM Symp. on Applied Computing*, 2015, pp. 2218–2224.

[16] D. Chen, S. Kalra, D. Irwin, P. Shenoy, and J. Albrecht, "Preventing Occupancy Detection From Smart Meters," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2426–2434, 2015.

[17] M. Savi, C. Rottondi, and G. Verticale, "Evaluation of the Precision-Privacy Tradeoff of Data Perturbation for Smart Metering," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2409–2416, 2015.

[18] X. Yang, X. Ren, S. Yang, and J. McCann, "A Novel Temporal Perturbation Based Privacy-Preserving Scheme for Real-Time Monitoring Systems," *Computer Networks*, vol. 88, pp. 72–88, 2015.

[19] G. Eibl and D. Engel, "Influence of Data Granularity on Smart Meter Privacy," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 930–939, 2015.

[20] G. W. Hart, "Nonintrusive Appliance Load Monitoring," *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.

[21] M. Erdogdu, N. Fawaz, and A. Montanari, "Privacy-Utility Trade-Off for Time-Series with Application to Smart-Meter Data," in *Workshops at the 29th AAAI Conference on Artificial Intelligence*, 2015.

[22] L. Sankar, S. R. Rajagopalan, S. Mohajer, and S. Mohajer, "Smart Meter Privacy: A Theoretical Framework," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 837–846, 2013.

[23] Low-Cost Single/Dual-Phase Isolated Electricity Measurement Reference Design. Texas Instrument. [Online]. Available: http://www.ti.com/tool/TIDM-TWOPHASEMETER-I2040

[24] J. Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial hmms with application to energy disaggregation," in *15th Int. Conf. on Artificial Intelligence and Statistics*, vol. 22, 2012, pp. 1472–1482.

[25] N. Batra, J. Kelly, O. Parson, H. Dutta, W. Knottenbelt, A. Rogers, A. Singh, and M. Srivastava, "NILMTK: An Open Source Toolkit for Non-Intrusive Load Monitoring," in *5th Int. Conf. on Future Energy Systems*, 2014, pp. 265–276.

[26] J. Z. Kolter and M. J. Johnson, "REDD: A Public Data Set for Energy Disaggregation Research," in *SustKDD workshop on Data Mining Applications in Sustainability*, 2011.

[27] W. Yang, N. Li, Y. Qi, W. Qardaji, S. McLaughlin, and P. McDaniel, "Minimizing Private Data Disclosures in the Smart Grid," in *2012 ACM Conf. on Computer and Communications Security*, 2012, pp. 415–427.

[28] H. Yang, "Utility-Privacy Trade-Offs of Data Manipulation Techniques for Smart Metering: A Quantitative Evaluation," Lehigh Univ., CSE Dept., Tech. Rep., 2016.