Privacy-utility trade-off for time-series with application to smart-meter data

Murat A. Erdogdu
Department of Statistics
Stanford University, CA 94305
erdogdu@stanford.edu

Nadia Fawaz
Technicolor
Los Altos, CA 94301
nadia.fawaz@technicolor.com

Andrea Montanari
Department of Electrical Engineering
Department of Statistics
Stanford University, CA 94305
montanari@stanford.edu

Abstract

We consider the online setting where a user would like to continuously release a time-series of data that is correlated with his private data, to a service provider in the hope of deriving some utility. Due to correlations, the continual observation of the released time-series puts the user at risk of inference of his private data by an adversary. To protect the user from inference attacks on his private data, the time-series is randomized prior to its release according to a probabilistic privacy mapping. The privacy mapping should be designed in a way that balances privacy and utility requirements over time. Our contributions are threefold. First, we formalize the framework for the design of utility-aware privacy mappings for time-series data, under both online and batch models. We provide a sequential scheme that allows to design online privacy mappings at scale, that account for privacy risk from the history of released data and future releases to come. Second, we prove the equivalence of the optimal mappings under the batch and the online models, in the case where the time-series samples are independent across time. We further show that there exists a gap between optimal batch and online privacy mappings when certain conditions are not satisfied. Finally, we evaluate the performance of the framework over synthetic and real-world time-series data. In particular, we show that smart-meter data can be randomized for privacy purposes to prevent disaggregation of per-device energy consumption, while preserving the utility.

1 Introduction

In the era of the Internet of Things, more and more devices collect and report fine-grained time-series data. Examples include sensors or monitoring devices in homes or business offices such as smart meters, HVAC systems (NEST), temperature, light, or motion sensors; as well as health-monitoring devices (fitbit, jawbone); and sensors on handheld devices such smartphones, tablets, game controllers. The collection of time-series data raises privacy concerns (Group and others 2010), as such data is often highly correlated with information that the user may deem sensitive and wants to keep private. For instance, an analyst having access to some of the aforementioned time-series data could infer private information including household composition, user behavior and lifestyle (appliance use, eating and sleeping patterns, presence, household activities) (Lisovich, Mulligan, and Wicker 2010; Robertson 2014), heath status (Hernandez et al. 2014), mobility patterns...

The collection of time-series data may happen with or without user consent, and potentially without the possibility for user to opt-out . The entity collecting the data may also make this data available to third parties (NEST 2014) with or without user knowledge. A natural question arises as to where the trust boundary lies. On one hand, the user may trust the entity aggregating the data but not the third-party with whom the aggregator may share data. For instance, the US department of Energy expressed concerns regarding the control over third-party access to consumer energy usage data (US Department of Energy 2010). On the other hand, the user may not entirely trust the aggregating entity in the first place, and may want to limit the amount of private information leaked by the released data. Data distortion has been proposed as a countermeasure to protect user privacy in both cases: either locally at the user side by randomized response of the user data prior to its release, or in a centralized manner at the aggregator side by randomization of the answer to a query over a database. In either case, the design of the distortion mechanism should satisfy formal privacy guarantees, but also maintain utility of the distorted data. Initially, data distortion approaches to privacy were devised for the static case, and when they were subsequently extended to the dynamic case of time-series data, scalability challenges arose. First, as the sequence of distorted releases carries correlation across time, the amount of distortion introduced by the randomization procedure may grow with the sequence length, thus maintaining utility often becomes challenging. Second, the distortion mechanism balancing the privacy-utility trade-off over time is often obtained through optimizations whose complexity may scale with the length of the sequence.

In this work, we consider the online setting where a user continuously releases a time-series of data that is correlated with his private data, to a service provider in the hope of deriving some utility from this release. Due to correlations, the continual observation of the released time-series puts the user at risk of inference of his private data by an adversary. To protect the user from inference attacks on his private data, samples from the time-series are sequentially randomized prior to their release according to a stochastic process, called the privacy mapping. The privacy mapping should be designed in a way that balances privacy and utility requirements over time. Our contributions are threefold. First, we formalize the framework for design of utility-aware

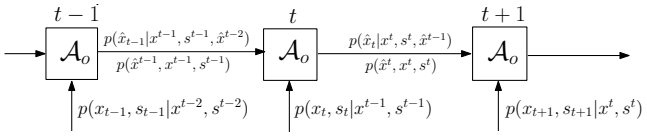

Figure 1: Sequential structure of the online scheme.

privacy mappings for time-series data, under both the online and batch models. Our framework for time-series data builds on and generalizes the static framework for privacy against statistical inference (du Pin Calmon and Fawaz 2012) to account for temporal correlations in the time-series, and for multiple sequential data releases. We provide a sequential scheme that allows to design online privacy mappings at scale, that accounts for privacy risks from the history of released data and future releases to come. Second, we prove the equivalence of the optimal mappings under the batch and the online models, in the case where the time-series samples are independent across time. We further show that there exists a gap between the optimal batch and online privacy mappings when the time-series samples do not satisfy certain conditions. Finally, we evaluate the performance of the framework over real-world time-series data. We show that smart-meter data can be randomized for privacy purposes to prevent disaggregation of per-device energy consumption, while maintaining the utility over the randomized series.

2 Related Work

The problem of preserving *differential privacy* (Dwork et al. 2006) when an analyst continually tracks statistics over a time-series was studied in (Dwork et al. 2010; Chan, Shi, and Song 2010) for running sum of bits and in (Bolot et al. 2013) for decayed sums of predicates, while (Shi et al. 2011; Rastogi and Nath 2010) considered differential privacy for aggregate-sum queries over the time-series data of multiple users. However, these approaches do not account for temporal correlations between samples of the time-series.

Approaches to protect user privacy for the specific case of smart-meter data (Jawurek, Kerschbaum, and Danezis 2012) include battery-based solutions (Kalogridis et al. 2010; McLaughlin, McDaniel, and Aiello 2011), data distortion (Rajagopalan et al. 2011; Sankar et al. 2013), and cryptographic protocols (Garcia and Jacobs 2011; Erkin and Tsudik 2012; Lin et al. 2012; Danezis et al. 2013). Battery-based solutions (Kalogridis et al. 2010; McLaughlin, McDaniel, and Aiello 2011) consist of off-loading some of the power consumption to batteries to hide some of the load. These approaches do not rely on formal guarantees, and require that the user purchases and installs batteries at home. Privacy-utility tradeoffs for smart-meter data were studied in (Rajagopalan et al. 2011; Sankar et al. 2013) under an information-theoretic framework. Assuming a stationary Gaussian Markov model for the energy load measurements, the authors show that the privacy-utility tradeoff can be optimized through water-filling. The privacy mechanism that distorts the time-series data is designed and applied offline once over the whole sequence prior to the release, and the privacy guarantees hold in the asymptotic regime of a large sequence. In contrast, our approach considers the online setting where distorted data is released sequentially, and it is applicable to any stochastic model for the time-series.

3 Privacy-utility framework for time-series

Notation: The set of integers $\{1, 2, ..., T\}$ will be denoted by $[T]$. $X \in \mathcal{X}$ denotes a random variable which takes values from the set $\mathcal{X}$. $X^T = \{X_1, X_2, \ldots, X_T\}$ denotes a sequence of T random variables.

We consider the dynamic setting where at every time $t \in [T]$, a privacy-conscious user generates samples from two time-series: a sample $S_t \in \mathcal{S}$ of sensitive data that the user would like to keep private, and a sample of data $X_t \in \mathcal{X}$ that the user is willing to release to a service provider, in the prospect of receiving some utility. Assuming that the time-series S^T and X^T are correlated, the sequential observation of samples from X^T by the service provider may allow him to adversarially perform inference attacks on the private S^T.

As a countermeasure to protect the user's privacy, the time-series X^T is not released as such, but is distorted according to a stochastic process called the *privacy mapping*, to generate a new time-series $\hat{X}^T$, from which the user will sequentially release samples to the adversarial service provider. The privacy mapping should be designed in a way that balances privacy and utility requirements over time: the time-series should be altered dynamically in a way that renders inference attacks against the private information S^T harder at any instant, but not so much that the alteration hinders extracting some utility from the released data.

The privacy mapping can be designed according to an online or a batch scheme. The *online* scheme refers to an algorithm that generates a distribution for $\hat{X}_t$ based on all available information up to time t, whereas the *batch* scheme refers to an algorithm that generates the joint distribution for the vector $\hat{X}^T$ based on the information available until time T (after observing all T samples). The difference in performance between batch and online schemes is called *regret*. Online schemes can be further categorized as *interactive* or *non-interactive*. In the interactive setting, at time t, $\hat{X}_t$ is generated based on $(\hat{X}^{t-1}, X^t, S^t)$, whereas in the non-interactive setting, the distorted data is generated based only on the current sample (X_t, S_t).

We now introduce the privacy and distortion metrics used to define the privacy-utility trade-off for time-series.

Privacy metric: We first define the general notion of *Information Leakage* as the amount of information that the observation of vector $\hat{X}^t$ leaks about vector S^t.

Definition 1. *The* Information Leakage $\mathcal{J}(\hat{X}^t; S^t)$ *from* $\hat{X}^t$ *to* S^t *quantifies the improvement in the inference of* S^t *after observing* $\hat{X}^t$.

Definition 1 captures a broad class of adversaries performing inference attacks on time-series. $\mathcal{J}(X^T; S^T)$ is used as a privacy metric in the following defintion:

Definition 2. *A sequence* $X^T \in \mathbb{R}^T$ *is* ϵ^T*-private with respect to a sequence* S^T *if* $\forall t \in [T]$*, the information leakage at time* t *is bounded by* ϵ_t*, i.e.,* $\forall t \in [T], \quad \mathcal{J}(X^t; S^t) \leq \epsilon_t$.

In the sequel, as in (du Pin Calmon and Fawaz 2012), we focus on a specific metric for the information leakage,

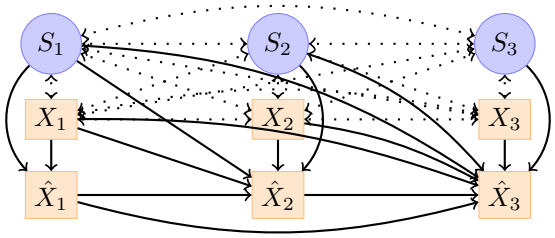

Figure 2: Online-scheme dependency-graph. Dotted lines: general dependency. Solid lines: Markov relation.

namely the mutual information (Shannon 1948) between the vectors of private data and of distorted data up to time T: $\mathcal{J}(\hat{X}^t; S^t) = I(S^T; \hat{X}^T) = H(S^T) - H(S^T|\hat{X}^T)$, where $H(\cdot)$ and $H(\cdot|\cdot)$ denote the *entropy* and the *conditional entropy*, respectively. $I(S^T; \hat{X}^T)$ quantifies the amount of information that vector $\hat{X}^T$ has about S^T.

Distortion metric: The distortion metric $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ describes the proximity of the distorted sequence $\hat{X}^T$ to the original data X^T. We will assume that the distortion metric is separable: $d(X^T, \hat{X}^T) = \frac{1}{T}\sum_{t=1}^{T} d(X_t, \hat{X}_t)$. Separable metrics include Hamming distance, and l_p-norms.

Privacy-Utility trade-off for time-series: The design of the privacy mapping should minimize the expected distortion $d(X^T, \hat{X}^T)$, while enforcing a privacy constraint ϵ_t at each time step t which will be specified by the user. That is, the privacy mapping should generate a distorted version of X^T which is ϵ^T-private and is close to the original data X^T.

The batch scheme is given by:

$$\mathcal{A}_b := \begin{array}{ll} \underset{p(\hat{x}^T|x^T,s^T)}{\text{minimize}} & \mathbb{E}[d(X^T, \hat{X}^T)] \\ \text{subject to} & \mathcal{J}(\hat{X}^t; S^t) \leq \epsilon_t, \ \forall t \in [T]. \end{array} \tag{3.1}$$

The online scheme is given by: $\forall t = [T]$,

$$\mathcal{A}_o := \begin{array}{ll} \underset{p(\hat{x}_t|x^t,s^t,\hat{x}^{t-1})}{\text{minimize}} & \mathbb{E}[d(X_t, \hat{X}_t)] \\ \text{subject to} & \mathcal{J}(\hat{X}^t; S^t) \leq \epsilon_t. \end{array} \tag{3.2}$$

Note that the online scheme minimizes the distortion between X_t and $\hat{X}_t$ over $p(\hat{x}_t|x^t, s^t, \hat{x}^{t-1})$ in a recursive manner that , at step t, it receives $p(\hat{x}^{t-1}, x^t, s^t)$ from the previous step and solves the optimization problem (See Figure 1). The sequential nature of the online algorithm generates a dependency graph that enables a recursive relation. The case for $T = 3$ is shown in Figure 2. The simple idea comes from the conditional independence of $\hat{x}^{t-1}$ and (x_t, s_t) conditioned on (x^{t-1}, s^{t-1}). At step t, the online algorithm requires the joint distribution $p(\hat{x}^{t-1}, x^t, s^t)$ as an input. This can be achieved simply by

$$p(\hat{x}^{t-1}, x^t, s^t) = p(x_t, s_t|x^{t-1}, s^{t-1})p(\hat{x}^{t-1}, x^{t-1}, s^{t-1}).$$

In contrast, the batch version minimizes over the whole joint distribution $p(\hat{x}^T|x^T, s^T)$ in a single run. This imposes a regret between online and batch versions of Privacy-Utility trade-off where the regret is defined as the difference between the optimal distortion achieved by the algorithms. Contrary to existing work, both algorithms allow us to restrict the information leakage at each time point t.

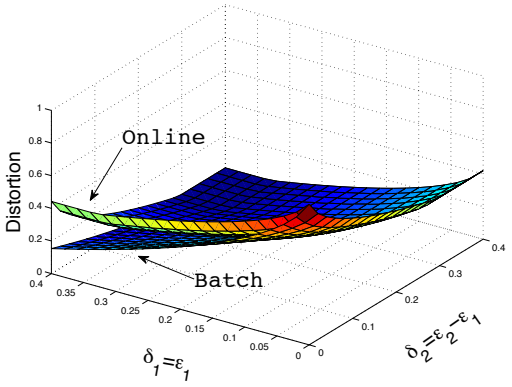

Figure 4: Privacy-Utility trade-off under a dependent model showing that there exists a regret.

4 Analysis of the schemes

Convexity of the optimization

Theorem 4.1. *Assume that the information leakage metric is mutual information and the alphabets $\mathcal{X}, \mathcal{S}$ are finite. Then, problems (3.1) and (3.2) are convex optimizations.*

By Theorem 4.1, Problems (3.1) and (3.2) can be solved using efficient convex optimization techniques. However, without any modeling assumption, the number of variables of the convex programs grows exponentially with T. Instead of full-dependence structure, a simplifying model assumption, such as *HMM* or time-window dependency, allows to decrease the problem size. For instance, for independent samples, the online problem scales linearly with T.

Regret under independence

Theorem 4.2. *Let the information leakage metric $\mathcal{J}$, be the mutual information. For a given batch problem with privacy levels $\{\epsilon_t\}_{t=1}^T$, if the random pairs $\{S_t, X_t\}_{t=1}^T$ are independent from each other, then there exists a choice of privacy levels $\{\epsilon'_t\}_{t=1}^T$ for the online problem resulting in no regret.*

If we further assume that the random pairs are i.i.d. and the increments of privacy levels, $\{\epsilon_t - \epsilon_{t-1}\}_{t=1}^T$, are non-decreasing, then online and batch problems are the same for the same choice of privacy levels, resulting in no regret.

Theorem 4.2 states that the online and the batch problems are the same and there is no regret under certain conditions. Also, the online scheme reduces to a non-interactive one under independence. Simulations in Figure 4 show that a positive regret might occur when the assumptions on the privacy levels are not satisfied or the samples (S_τ, X_τ) are correlated. Figure 4 also illustrates the convexity of the problems.

5 Experiments on smart-meter dataset

We experimented our online scheme on the Reference Energy Disaggregation Dataset (REDD) (Kolter and Johnson 2011; Kolter and Jaakkola 2012), which consists of power consumption of 6 houses. For each house, the power consumption of each appliance in the house is available every

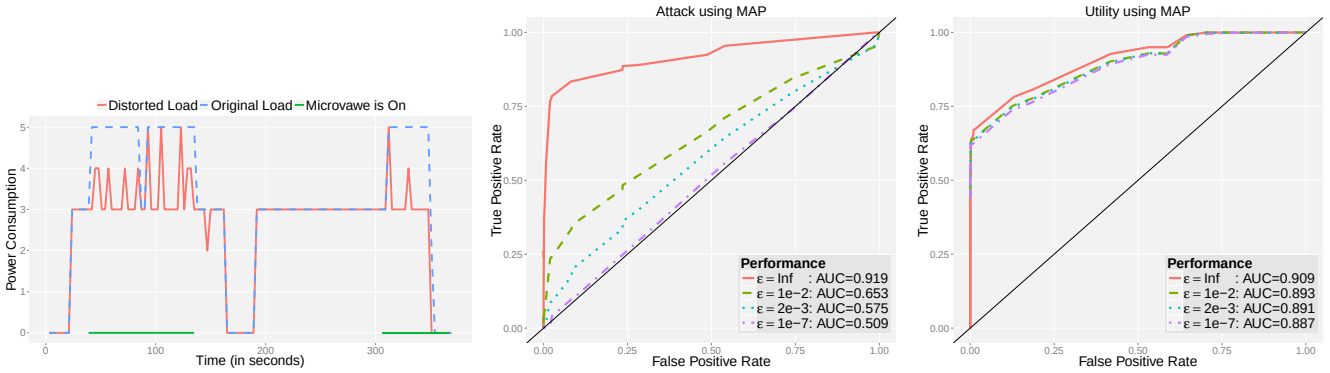

Figure 3: *(Left):*Dashed blue line: aggregate load X^T. Continuous red line: distorted load $\hat{X}^T$. Green segments on time axis: microwave ON. *(Center):* ROC for MAP inference attack on private data (microwave). *(Right):* ROC for MAP utility inference (washer-dryer). Distorted load allows for washer-dryer utility inference but prevents attack on private microwave data.

3 seconds. For a given house, we refer to *aggregate load* of that house at time t as the total power consumption of all the appliances at that time.

The scenario for the experiments involves a household, and a service provider, who may behave adversarially. The service provider offers some utility to the household, which requires inference of the state of washer-dryer in the house from the aggregate load. To provide utility to the user, for instance automated control of the washer/dryer, the service provider needs process the aggregate load data received the user. The utility U_t represents the outcome of the service provider's algorithm that runs on the released data. In this experiment, U_t will be the result of the inference on the state of *washer-dryer* from the load data.

The household is willing to give the aggregate load X_t to the service provider, but wishes to keep the information related to their eating patterns private, in particular the microwave usage which can also be inferred from the aggregate load. In this experiment, we choose S_t to be the state of the *microwave* (ON or OFF). If the user released X_t as is, it can be used adversarially by the service provider to infer information regarding S_t, thus raising privacy concerns. The user will instead release a distorted version $\hat{X}_t$. Examples of adversarial providers include third-parties, such as apps, to whom the company operating the smart-meter may give access to the data it collects (NEST 2014), or a malicious insider such as a curious employee. Our goal is to get the utility related to the washer-dryer, while keeping the sensitive information regarding microwave usage private. The dataset provides ground truth for both the microwave and washer-dryer state, which allows us to verify the performance of our approach. Note that neither the private data nor the utility are limited to components of the aggregate load, and could be any information correlated with the aggregate load.

In the training process, we obtained the empirical distributions for (S_t, X_t) and (U_t, X_t). These distributions are also available to the adversarial provider, and the household. This worst-case setting provides too much information to the adversary but might be the case as the adversary might have collected data from a different but similar household. Using the training set, the adversarial provider trains models for attack and utility, respectively, and they are given a test set to make inference using these models.

We experimented on the above scenario where the inference algorithm is *maximum a posteriori* (MAP) estimation. The household provides the distorted load $\{\hat{X}_t\}_{t=1}^T$ using our non-interactive online scheme where the privacy leakage levels are defined as $\epsilon_t = t\epsilon$. The case where the household does not use the algorithm and sends the aggregate load itself is denoted by $\epsilon = \text{Inf}$.

For each leakage level ϵ, the distorted aggregate load is fed to the trained inference algorithms (See Figure 3(left) for a comparison between X^T and $\hat{X}^T$). ROC curves for MAP inference are shown in Figure 3: the center plot shows the ROC of the adversarial inference on the microwave, and the right plot shows the ROC of the utility inference on the washer-dryer. As ϵ decreases (information leakage), the quality of the adversarial inference attack degrades whereas the utility inference remains unchanged.

6 Discussion

We consider the setting where a user continuously releases a time-series that is correlated with another private time-series, to a service provider in the hope of deriving some utility from this release. We propose general online and batch schemes for the design of utility-aware privacy mappings, which bound the private information leakage while minimizing the distortion of the data generated by the mapping. These general schemes can be adapted to any modeling assumption suitable for a given application. We prove that both schemes can be cast as convex optimizations. Then, under the assumption of independent data samples across time, we show that the solutions of the online and the batch optimizations are the same, thus there is no regret. We then provide an example with correlated samples in which there exists a positive regret between batch and online solutions. Experiments on a smart-meter dataset show that leakage can be bounded over time while maintaining the utility of the released data. Further applications may include privacy for time-series data from health-monitoring devices, sensors in houses, offices, cars or handheld devices.

References

Bolot, J.; Fawaz, N.; Muthukrishnan, S.; Nikolov, A.; and Taft, N. 2013. Private decayed predicate sums on streams. In *Proceedings of the 16th International Conference on Database Theory*, 284–295. ACM.

Chan, T. H.; Shi, E.; and Song, D. 2010. Private and continual release of statistics. In *Automata, Languages and Programming*. Springer. 405–417.

Danezis, G.; Fournet, C.; Kohlweiss, M.; and Zanella-Béguelin, S. 2013. Smart meter aggregation via secret-sharing. In *Proceedings of the first ACM workshop on Smart energy grid security*, 75–80. ACM.

du Pin Calmon, F., and Fawaz, N. 2012. Privacy against statistical inference. In *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, 1401–1408. IEEE.

Dwork, C.; McSherry, F.; Nissim, K.; and Smith, A. 2006. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography*. Springer.

Dwork, C.; Naor, M.; Pitassi, T.; and Rothblum, G. N. 2010. Differential privacy under continual observation. In *Proceedings of the 42nd ACM symposium on Theory of computing*, 715–724. ACM.

Erkin, Z., and Tsudik, G. 2012. Private computation of spatial and temporal power consumption with smart meters. In *Applied Cryptography and Network Security*, 561–577. Springer.

Garcia, F. D., and Jacobs, B. 2011. Privacy-friendly energy-metering via homomorphic encryption. In *Security and Trust Management*. Springer. 226–238.

Group, S. G. I. P. C. S. W., et al. 2010. Nistir 7628 guidelines for smart grid cyber security. *Privacy and the smart grid* 2.

Hernandez, J.; Li, Y.; Rehg, J. M.; and Picard, R. W. 2014. Bioglass: Physiological parameter estimation using a head-mounted wearable device. In *Proceedings of the 4th International Conference on Wireless Mobile Communication and Healthcare, MobiHealth*.

Jawurek, M.; Kerschbaum, F.; and Danezis, G. 2012. Privacy technologies for smart grids - a survey of options. Technical Report MSR-TR-2012-119.

Kalogridis, G.; Efthymiou, C.; Denic, S. Z.; Lewis, T. A.; and Cepeda, R. 2010. Privacy for smart meters: Towards undetectable appliance load signatures. In *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, 232–237. IEEE.

Kolter, J. Z., and Jaakkola, T. 2012. Approximate inference in additive factorial hmms with application to energy disaggregation. In *International Conference on Artificial Intelligence and Statistics*, 1472–1482.

Kolter, J. Z., and Johnson, M. J. 2011. Redd: A public data set for energy disaggregation research. `http://redd.csail.mit.edu/`.

Lin, H.-Y.; Tzeng, W.-G.; Shen, S.-T.; and Lin, B.-S. P. 2012. A practical smart metering system supporting privacy preserving billing and load monitoring. In *Applied Cryptography and Network Security*, 544–560. Springer.

Lisovich, M. A.; Mulligan, D. K.; and Wicker, S. B. 2010. Inferring personal information from demand-response systems. *Security & Privacy, IEEE* 8(1):11–20.

McLaughlin, S.; McDaniel, P.; and Aiello, W. 2011. Protecting consumer privacy from electric load monitoring. In *Proceedings of the 18th ACM conference on Computer and communications security*, 87–98. ACM.

NEST. 2014. Nest developer program.

Rajagopalan, S. R.; Sankar, L.; Mohajer, S.; and Poor, H. V. 2011. Smart meter privacy: A utility-privacy framework. In *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*, 190–195. IEEE.

Rastogi, V., and Nath, S. 2010. Differentially private aggregation of distributed time-series with transformation and encryption. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*, 735–746. ACM.

Robertson, J. 2014. Your outlet knows: How smart meters can reveal behavior at home, what we watch on tv. *Bloomberg news*.

Sankar, L.; Rajagopalan, S. R.; Mohajer, S.; and Poor, H. V. 2013. Smart meter privacy: A theoretical framework. *Smart Grid, IEEE Transactions on* 4(2):837–846.

Shannon, C. E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27.

Shi, E.; Chan, T.-H. H.; Rieffel, E. G.; Chow, R.; and Song, D. 2011. Privacy-preserving aggregation of time-series data. In *NDSS*, volume 2, 4.

US Department of Energy. 2010. Data access and privacy issues related to smart grid technologies.