Privacy Preserving Smart Meter Data

Swapna Thorve
Department of Computer Science
Virginia Tech
Blacksburg, Virginia
swapna6@vt.edu

Lindah Kotut
Department of Computer Science
Virginia Tech
Blacksburg, Virginia
lkotut@vt.edu

Mary Semaan
Department of Building Construction
Virginia Tech
Blacksburg, Virginia
maryss@vt.edu

ABSTRACT

Smart meters track fine-grained usage of utilities for billing purposes, offer incentives for the adoption of renewable energy, provide detailed customer feedback, monitor current demand and forecast future energy consumption/needs. A key impediment to the widespread smart meter acceptance and adoption is the privacy and security concerns surrounding the identification of users' household activities from the frequent logging of utility usage over time, to the lax privacy preservation practices leading to exposure of the data and misuse by authorized and/or unauthorized entities.

To address these concerns, we apply *differential privacy* on real-world smart metering data and contemplate the practicality and challenges of this approach. To that end, we consider utility companies as trusted entities to collect real data and that they can share differentially private data with third parties.

We design a two-step approach to generate representative private time series using a dataset of 2 million time series with 48 time stamps and offer recommendations on parameter settings in *differential privacy* while considering the feasibility of this approach given its complexity and the size of the data.

CCS CONCEPTS

• **Security and privacy** → *Data anonymization and sanitization*;

KEYWORDS

Smart Meters, Time Series, Clustering, Privacy Preservation, Differential Privacy

ACM Reference Format:
Swapna Thorve, Lindah Kotut, and Mary Semaan. 2018. Privacy Preserving Smart Meter Data. In *Proceedings of The 7th International Workshop on Urban Computing (UrbComp'18),* Jennifer B. Sartor, Theo D'Hondt, and Wolfgang De Meuter (Eds.). ACM, New York, NY, USA, Article 4, 5 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

Smart meter deployment has become a norm: apart from benefiting utility companies in billing, consumption monitoring and forecasting, it also provides the customer with the granular information they need to monitor energy use, and plan for conservation. The granularity-gains and privacy costs to the consumer are known, and work exist that offer solutions in answering the dilemma on how best to preserve confidentiality of information while at the same time offering the granularity needed to take the most advantage of affordances brought about by smart metering.

Our work, while considering the application of different privacy preserving methods on data, also considers the confounding factors surrounding data release – both from ethical and logical perspectives. From the ethical standpoint, it has been theoretically proven that differential privacy [6] can be used to preserve privacy of sensitive data. Using this guarantee, we attempt the application of *differential privacy* – the state of the art in privacy preservation on real world data.

We utilize smart meter data available at granular scale of half-hourly (48) time intervals, that provide different challenges that may not be as visible when considering non-time-series/discrete/counting data. This is especially evident when applying differential privacy, a factor that is made visible when considering real-world data.

Given previous work and the state of the art in privacy preservations and the challenges with differential privacy parameter tuning for time series data which we discuss subsequently, we offer the following contributions:

(1) New insights on the challenges of applying differential privacy on real-world smart meter data
(2) Application of differential privacy on real-world big data set
(3) A Showcase of challenges in parameter tuning followed by initial recommendations on setting these parameters

2 BACKGROUND

We glance at the evolution of privacy preserving methods, their applications, and the state-of-art. We limit our scope to privacy-preserving methods focusing on time-series data in general – with a specific emphasis on smart meter applications, especially those utilizing real world data.

2.1 Previous Approaches

Data anonymization is a generic and perhaps a default approach of privacy preservation: it is intuitive, requiring a part or whole of personally identifiable information be withheld from third parties, with multiple research considering different facets of personally identifying information that encompasses consumption data. As an approach, it has been proven [19] to be insufficient in preserving the privacy of user data – a minimal subset of shared data could be used to make accurate inference about the omitted data.

Aggregation over a number of smart meters is an improvement over generic anonymization, and offers provably better guarantee on privacy of data. Approaches to smart metering include 'leveling', an approach that serve to preserve granularity of data for

UrbComp'18, August 2018, London, UK

ACM ISBN 123-4567-24-567/08/06.
https://doi.org/10.475/123_4

the customer, while applying masking that serves to generalize appliance-specific data [14], yet preserving the true aggregate for billing purposes.

Other privacy-preserving approaches include work that consider utility companies or collecting entities as untrusted parties, and propose provisions for alternate 3rd party escrow mechanism to be used to both anonymize data collected from individual smart meter data for transmission to the utility companies, and to authenticate anonymous meter readings for billing purposes. Different models have been proposed to achieve this: a smart-grid specific privacy-preserving database anonymizer was presented by [16], with distributed incremental data aggregation advanced by [22] – who presented a model where all smart meters were involved both in the routing of data from source to collection and in aggregation by applying symmetric homomorphic encryption.

Considering *differential privacy* (DP) as a state of art in privacy preservation, there have been works that utilize this both as a concept and as an application. Saleheen et. al [18] used DP on physiological time series data (e.g., heart rate, activity data) obtained from smart phones to protect private behavior of individuals but also retain reasonable utility to perform research. In a study by Fan et. al [10], a novel framework, FAST, is developed to release aggregate statistics of time series data based on DP by using adaptive sampling. Differentially private smart metering data is obtained in [1] by applying a novel distributed Laplacian Perturbation Algorithm (DLPA) to grouped smart meters. Homomorphic encryption is further applied to transfer this data to the utility company. Wang et. al [21] considered individual privacy involved in trajectory publishing scenario: DP is employed by designing a novel concept of series-indistinguishability and a correlated Laplace mechanism for time series.

Given the complexity of smart meters and the larger network, different proposals have been made as to where to apply differential private techniques. We group these approaches into two broad categories depending on whether the utility/collecting service is considered a trusted party or not. In the case of the collection company considered an untrusted party, proposals focus on applying differential privacy at the granular smart meter level [1, 23] in an effort to ensure that no personally identifiable data is transmitted. Engel and Eibl [9] worked with DP and smart metering data and generated a differentially private aggregate time series to represent the entire dataset followed by smoothing the curve using moving average filter, which they argued was to improve the utility of the curve. However, for the case where the collecting entity is trusted, approaches to apply differential privacy are not strongly presented in literature.

2.2 Our Approach

From the work considered above, we note the dearth of those that deal directly with data. Most proposed methods provide theoretical foundations and mathematical proofs of approaches of preserving data, but do not offer what effects the application of said approaches on real-world datasets would look like, which - as we have found, present unique challenges that are not accounted for in the theoretical recommendations. Our fundamental assumption is that utility companies have the granular data already, a premise that is held up by the type of data we use in our analysis, and so we consider them a trusted entity.

We consider *Differential Privacy* [6] as the state of art in the privacy preservation domain. This approach has been proven to guarantee that the risk to privacy would not increase should the data be shared and other external data be used to probabilistically attempt to determine personal information.

We employ clustering on the time series as the first stage of the two-step process, allowing for a clearer understanding of variations, peak time consumption and to learn overall patterns in energy consumption data. Clustering also emphasizes the importance of preserving the diversity in the energy consumption patterns. Subsequently, we apply differential privacy to the output of each cluster using *mean query*.

2.3 Assumptions and Limitations

We do not consider methods that perturb data by adding or swapping data as they have an end result of making the resultant data wrongfully noisy – a fact that has ethical complications. We thus consider only those methods that preserve the underlying truthfulness of the data.

Our approach can be compared to [9], who applied differential privacy on a small sample of real-world data. We differ broadly in the size of our dataset – how closely it mirrors the real world, and how we set and justify the parameters set for applying differential privacy. Our main focus is on how differential privacy can be applied on time series data for purposes of sharing with untrusted parties. We consider utility companies as trusted entities to collect real data and that they can share differentially private data with untrusted third parties. Importantly, we consider challenges with applying differential privacy on time series and discuss the challenges of parameter tuning and the implications on the fidelity of real-world data.

3 DATASET

For our analysis, we use UK Power Network's[1] smart metering data for 5567 households in London. The download was free. Each entry describes energy consumption readings that were taken at 30 minute intervals at household level as part of the UK Power Network–led *Low Carbon London* [2], spanning the time between November 2011 and February 2014.

Each household was allocated into a CACI Acorn group (2010 revision[3]) classification index – which measures economic circumstances and demographic information for each sector corresponding to a geographical sector. This classification can then be used to make an estimation about the demographic and economics of any given house. The UK Power Networks claimed that these households selected represented a balanced sample of the Greater London population. The dataset size is 10GB with approximately 167 million entries – consisting of half-hourly readings.

We selected 2013 data – as this was the year with the most households participating, and selected a time-series record consists of

[1] https://data.london.gov.uk/dataset/smartmeter-energy-use-data-in-london-households

[2] http://innovation.ukpowernetworks.co.uk/innovation/en/Projects/tier-2-projects/Low-Carbon-London-(LCL)/

[3] https://acorn.caci.co.uk/downloads/Acorn-User-guide.pdf

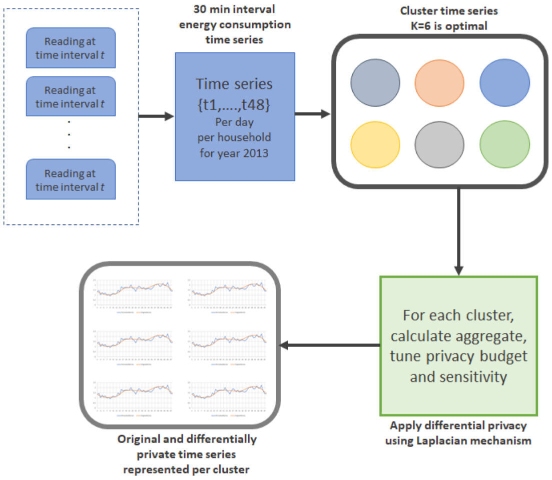
Reading at time interval t
Reading at time interval t
Reading at time interval t
30 min interval energy consumption time series
Time series {t1,....,t48} Per day per household for year 2013
Cluster time series K=6 is optimal
For each cluster, calculate aggregate, tune privacy budget and sensitivity
Apply differential privacy using Laplacian mechanism
Original and differentially private time series represented per cluster

Figure 1: 2-step procedure for generating differentially private time series: after pre-processing, time series are clustered using K-Means (step 1) – k=6 the optimal number of clusters. Differential privacy is then applied to each k using mean query functions and Laplacian mechanism (step 2).

30-minute interval readings throughout the day for a given household. Missing or incomplete data accounting for 0.00026% of the data was discarded.

4 PRIVACY PRESERVATION

We design a two-step approach to generate differentially private time series for the given smart meter dataset as represented in Figure 1. We utilize time-series clustering as a first step and the application of differential privacy to each resulting cluster as the second step. We formalize the problem as follows:

Given a dataset D of n time series, $D = \{T_1, T_2, ..., T_n\}$ where each T_i is composed of 48 points, $T_i = \{t_1, t_2, ..., t_{48}\}$. We group D into k partitions, $C = \{C_1, C_2, ..., C_k\}$ such that

$$D = \bigcup_{j=1}^{k} C_j \text{ and } C_j \cap C_n = \emptyset, j \neq n \, .$$

For each cluster C_j, a differentially private time series is generated such that $T_j^{'} = f_j + Lap_\lambda$, where $T_j^{'}$ = resultant differentially private time series ($T_j^{'} = \{t_1^{'}, t_2^{'}, ..., t_{48}^{'}\}$) for a group of time series in cluster C_j, f_j = aggregate query function for cluster C_j, Lap_λ = Laplacian mechanism M .

4.1 Clustering time series

We perform clustering in order to find prominent energy consumption patterns in the time-series dataset [3, 11] and to produce a meaningful representative. We then utilize this step as a subroutine for producing differentially private representative time series. We considered previous work on general time series clustering as well as those specializing in smart meter data clustering. Work done by [13] for instance, implemented 11 clustering methods to residential metering dataset consisting of 10^5 time series, and applied 6 internal validation metrics considering the suitability of the individual clustering method type applied. Approach by [5] considered previous literature and applied content analysis and categorization to derive energy saving behavior clusters considering 9 attributes (energy savings, cost, frequency of performance, required skill level, observability, locus of decision, household function, home topography, and appliance topography). Other approaches [2, 4, 12, 15] also consider and adopt variants of partition-based algorithms in the process of privacy preservation.

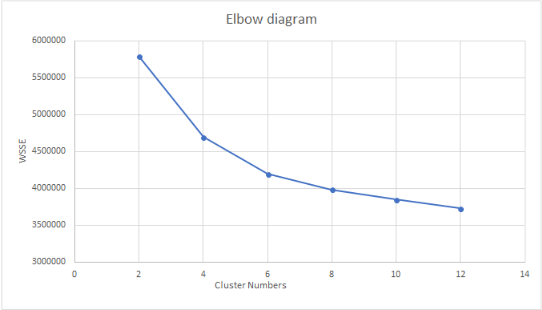

Figure 2: Elbow plot to determine the optimal number of clusters (k) by employing K-Means. Y-axis shows within-cluster sum of squares error (WSSE) and X-axis depicts number of clusters. We choose k=6 for optimal clustering of our data.

We apply the K-Means clustering method on the dataset of 2 million time series using the Apache Spark framework. The Euclidean distance metric is applied which performs a point-to-point distance calculation between any two time series. The setup is run for 100 iterations while varying the number of clusters, then using the *elbow method* we find that the data can be best grouped into six clusters (k=6) – as shown in Figure 2 and summarized in Algorithm 1. We use the output of this procedure as the basis for privacy preserving technique we describe below.

Algorithm 1 Simple K-Means clustering procedure

```
Input : $D = \{T_1, T_2, ..., T_n\}$ where each $T_i = \{t_1, t_2, ..., t_{48}\}$.
Output : Optimal size of clusters $k$ .
procedure ClusterTimeSeries
    $arbitaryClusterSizes \leftarrow \{2, 4, 6, 8, 10, 12\}$
    for each $v$ in $arbitaryClusterSizes$ do
        Perform k-means clustering
        $centroids[v] \leftarrow$ centroids for $v$
        $wsse[v] \leftarrow$ within-sum-of-square-error for $v$
    end for
    $k \leftarrow$ cluster size for $min\{wsse\}$
    return $k$
end procedure
```

4.2 Differential Privacy

Differential privacy (DP) uses random noise from a mechanism to ensure that the adversary fails to guess whether a record is present in the dataset even if he knows all the records except the one [6]. This is achieved by masking the difference between neighboring datasets to the original query f. The maximal difference on the results of query f is defined as the *sensitivity* (Δf), which determines how much perturbation is required for preserving the original answer. The *privacy budget* ϵ parameter controls the privacy guarantee of the *mechanism* M in use [7, 8]. Utility can be evaluated by considering the difference between the non-private and the private output: a smaller distance implying higher utility [20].

A formal definition of differential privacy is given below: a randomized mechanism M gives (ϵ, Δ)-differential privacy for every set of outputs S, and for any neighboring datasets of D and $D^{'}$, if M satisfies:

$$Pr[M(D) \in S] \leq exp(\epsilon).Pr[M(D\prime) \in S] + \Delta \quad (1)$$

Algorithm 2 Differential privacy using Laplacian mechanism

```
Input : $C = \{C_1, C_2, ..., C_k\}$ and time series in each cluster.
Output : $k$ differentially private time series representing each
  cluster $[T^{'}_1,....,T^{'}_k]$ where $T^{'}_j = \{t^{'}_1, t^{'}_2, ..., t^{'}_{48}\}$
procedure GenerateDifferentiallyPrivateTimeseries
    for each cluster $x$ in optimal $k$ do
        $z \longleftarrow$ number of time series in $x$
        for every $p$ interval do
            Aggregate at $p^{th}$ interval, $f_{p,x} = \sum_{n=1}^{z} t_{p,n}$
            Let $target()$ be the target aggregate function
            to produce $f_{p,x}$
            Let $\Delta$f = $\frac{1}{z}$ be the global sensitivity for $x$
            Let $\epsilon$ be the privacy budget
            Let $M$ be Laplace mechanism $Lap_\lambda$ for adding noise
            Target aggregate, $T^{'}_{p,x} = f_{p,x} + Lap_\lambda$
        end for
        $T^{'}_x = \sum_{p=1}^{48} Y_{p,x}$
    end for
    return $[T^{'}_1,....,T^{'}_k]$
end procedure
```

We apply differential privacy to the time series data using *Laplace mechanism* M, while utilizing clustering as a preceding step (Algorithm 1 and k=6) in order to generate a representative perturbed time series for each cluster by using a mean query f. For M, each half-hourly interval reading is considered independent. The generated noisy time series are then compared to the average time series for each cluster in order to understand the nuances of tuning ϵ and the utility of the generated time series curves (Algorithm 2).

We experiment with different values of ϵ: those recommended in previous work (ln2, ln3, 1, 0.1) [9, 18, 20, 21, 23] and others not considered in previous work (0.01, 0.001, 0.005, 0.0001, 0.0005), and observe that ϵ is greatly influenced by the number of data points in the cluster (Figure 3). We proceed using a global sensitivity ($\frac{1}{N}$) where N is the number of time series (data points) in a cluster.

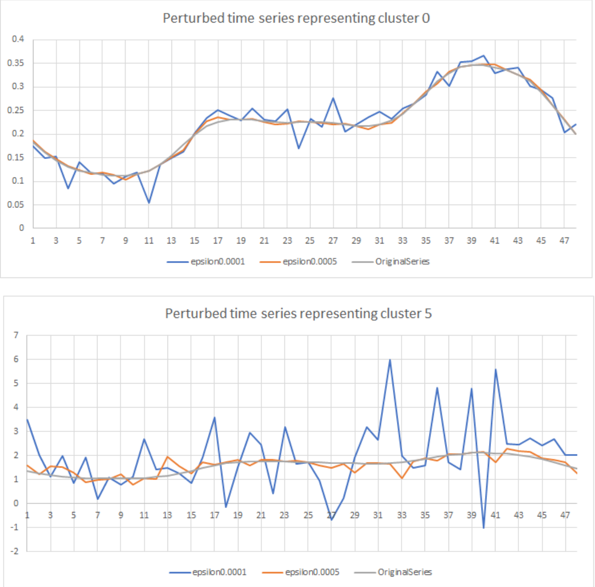

Figure 3: Effect of ϵ on the generated private time series. Demonstration is shown for cluster 0 and cluster 5. Cluster 0 has 659986 time series whereas cluster 5 has 9587 time series. ϵ=0.0005 works better for cluster 5 and ϵ=0.0001 works better for cluster 0. Global sensitivity is 1/N for each cluster.

5 RESULTS AND DISCUSSION

We use the *elbow method* to determine the best k value. Figure 2 plots the within-cluster sum of squares error (WSSE) – a measure of the variability of the observations within each cluster and an internal measure for validating cluster cohesion with respect to cluster numbers. The cluster centroids present a set of time series that mimics the important characteristics such as peak time consumption and overall daily demand profile thereby grouping segments of households with similar demand profiles and preserving prominent consumption patterns in the dataset.

Next, we apply *differential privacy* to the output of each cluster using mean query. We try to understand the effect of number of data points in a cluster with regards to ϵ and Δf. Our observation is that, as the number of data points increases, the ϵ value required to generate a differentially private time series curve decreases. The respective value obtained for ϵ for different clusters also describes the utility of the curve. Due to space constraints, we only showcase two of the six clusters. We find that the often suggested parameter values for ϵ did not generate sufficient level of privacy for our dataset (Figure 3).

To the best of our knowledge, we could not identify literature offering concrete suggestion about tuning Δf when considering large and/or time series datasets. We reason that it is because this calculation is generally complex and expensive. We experimented with local sensitivity [9] but it was computationally expensive and did not generate satisfactory results. In 2017, Rubinstein developed a sensitivity sampling method [17] which we employed using 'diffpriv' package in R. The values returned by this function and the global sensitivity value we consider were almost in the same range – thus our preference for using global sensitivity in this work.

We believe that it is better to generate representative perturbed time series for smaller representative groups of a big dataset, than generating only one single aggregate perturbed time series. These smaller groups tend to capture the trends of the dataset, while retaining utility value. Hence, the need for clustering time series in Step 1.

Researchers as well as utility companies should have a certain amount of knowledge about the dataset to conduct research, provide feedback to groups of customers, design incentives to shift peak demand, etc. The output of such an approach will be valuable to these entities while maintaining privacy at an individual and household level. The application of differential privacy in time series is also very challenging due to the temporal nature (and hence dependence and correlation of values). We strongly believe that the sequential and parallel compositions of differential privacy mentioned in [6, 20] need to be further examined for tuning ϵ.

6 CONCLUSION AND FUTURE WORK

This work offers a look at how we can assure privacy of time series data containing sensitive information. We present a two-stage approach for privacy preservation of smart metering data: clustering followed by application of differential privacy. This method generates differentially private demand curves for representative consumption patterns in the dataset, thus, retaining its value for further usage by researchers.

We briefly discuss the difficulty of evaluating *privacy budget* ϵ and *sensitivity* Δf and their effect on time series. We find that ϵ is dependent on number of data points, hence, obtaining utility from such a mechanism is a challenging task. Future avenues for research include: better attribute selection, Laplace mechanism for correlated time series data, and design of Δf for large time series datasets.

Currently, we are considering offering both the privacy cost/benefit and the ethical implications to utility companies and negotiating a means of releasing such data to the research community. Given the preliminary nature and narrow scope of this work in focusing on the application of differential privacy on real-world data and addressing initial challenges, we will consider a comprehensive comparison of performance of various methods discussed in this work and others in the future.

REFERENCES

[1] Gergely Ács and Claude Castelluccia. 2011. I have a dream!(differentially private smart metering). In *International Workshop on Information Hiding*. Springer, 118–132.

[2] Gagan Aggarwal, Tomás Feder, Krishnaram Kenthapadi, Samir Khuller, Rina Panigrahy, Dilys Thomas, and An Zhu. 2006. Achieving anonymity via clustering. In *Proceedings of the twenty-fifth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM, 153–162.

[3] Saeed Aghabozorgi, Ali Seyed Shirkhorshidi, and Teh Ying Wah. 2015. Time-series clustering âĂŞ A decade review. *Information Systems* 53 (2015), 16 – 38. https://doi.org/10.1016/j.is.2015.04.007

[4] Maria-Florina Balcan, Travis Dick, Yingyu Liang, Wenlong Mou, and Hongyang Zhang. 2017. Differentially Private Clustering in High-Dimensional Euclidean Spaces. In *Proceedings of the 34th International Conference on Machine Learning (Proceedings of Machine Learning Research)*, Doina Precup and Yee Whye Teh (Eds.), Vol. 70. PMLR, International Convention Centre, Sydney, Australia, 322–331. http://proceedings.mlr.press/v70/balcan17a.html

[5] Hilary S. Boudet, June A. Flora, and K. Carrie Armel. 2016. Clustering household energy-saving behaviours by behavioural attribute. *Energy Policy* 92, C (2016), 444–454. https://doi.org/10.1016/j.enpol.2016.02.0

[6] Cynthia Dwork. 2006. Differential Privacy. In *Proceedings of the 33rd International Conference on Automata, Languages and Programming - Volume Part II (ICALP'06)*. Springer-Verlag, Berlin, Heidelberg, 1–12. https://doi.org/10.1007/11787006_1

[7] Cynthia Dwork. 2011. A Firm Foundation for Private Data Analysis. *Commun. ACM* 54, 1 (Jan. 2011), 86–95. https://doi.org/10.1145/1866739.1866758

[8] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating Noise to Sensitivity in Private Data Analysis. In *Theory of Cryptography*, Shai Halevi and Tal Rabin (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 265–284.

[9] Günther Eibl and Dominik Engel. 2017. Differential privacy for real smart metering data. *Computer Science - Research and Development* 32, 1 (01 Mar 2017), 173–182. https://doi.org/10.1007/s00450-016-0310-y

[10] Liyue Fan and Li Xiong. 2012. Adaptively Sharing Time-Series with Differential Privacy. *CoRR* abs/1202.3461 (2012). arXiv:1202.3461

[11] FĂľlix Iglesias and Wolfgang Kastner. 2013. Analysis of Similarity Measures in Times Series Clustering for the Discovery of Building Energy Patterns. *Energies* 6, 2 (January 2013), 1–19. https://ideas.repec.org/a/gam/jeners/v6y2013i2p579-597d23092.html

[12] Somesh Jha, Luis Kruger, and Patrick McDaniel. 2005. Privacy Preserving Clustering. In *Computer Security – ESORICS 2005*, Sabrina de Capitani di Vimercati, Paul Syverson, and Dieter Gollmann (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 397–417.

[13] Annika Todd Ling Jin, Doris Lee, Alex Sim, Sam Borgeson, Kesheng Wu, and C. Anna Spurlock. 2017. Comparison of Clustering Techniques for Residential Energy Behavior using Smart Meter Data. *2nd International Workshop on Artificial Intelligence for Smart Grids and Smart Buildings. In conjunction with AAAI 2017* (2017). https://aaai.org/ocs/index.php/WS/AAAIW17/paper/view/15166

[14] Stephen McLaughlin, Patrick McDaniel, and William Aiello. 2011. Protecting Consumer Privacy from Electric Load Monitoring. In *Proceedings of the 18th ACM Conference on Computer and Communications Security (CCS '11)*. ACM, New York, NY, USA, 87–98. https://doi.org/10.1145/2046707.2046720

[15] J. Ren, J. Xiong, Z. Yao, R. Ma, and M. Lin. 2017. DPLK-Means: A Novel Differential Privacy K-Means Mechanism. In *2017 IEEE Second International Conference on Data Science in Cyberspace (DSC)*. 133–139. https://doi.org/10.1109/DSC.2017.64

[16] C. Rottondi, G. Verticale, and A. Capone. 2012. A security framework for smart metering with multiple data consumers. In *2012 Proceedings IEEE INFOCOM Workshops*. 103–108. https://doi.org/10.1109/INFCOMW.2012.6193469

[17] Benjamin I. P. Rubinstein and Francesco Aldà. 2017. Pain-Free Random Differential Privacy with Sensitivity Sampling. *CoRR* abs/1706.02562 (2017). arXiv:1706.02562 http://arxiv.org/abs/1706.02562

[18] Nazir Saleheen, Supriyo Chakraborty, Nasir Ali, Md Mahbubur Rahman, Syed Monowar Hossain, Rummana Bari, Eugene Buder, Mani Srivastava, and Santosh Kumar. 2016. mSieve: Differential Behavioral Privacy in Time Series of Mobile Sensor Data. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 706–717. https://doi.org/10.1145/2971648.2971753

[19] Latanya Sweeney. 2002. K-anonymity: A Model for Protecting Privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* 10, 5 (Oct. 2002), 557–570. https://doi.org/10.1142/S0218488502001648

[20] Zhu Tianqing, Li Gang, Zhou Wanlei, and Yu Philip S. 2017. *Differential Privacy and Applications* (1st ed.). Springer International Publishing.

[21] Hao Wang and Zhengquan Xu. 2017. CTS-DP: Publishing correlated time-series data via differential privacy. *Knowledge-Based Systems* 122 (2017), 167 – 179. https://doi.org/10.1016/j.knosys.2017.02.004

[22] Xu Zhang, Mi Wen, Kejie Lu, and Jingsheng Lei. 2017. A privacy-aware data dissemination scheme for smart grid with abnormal data traceability. *Computer Networks* 117 (2017), 32–41.

[23] Z. Zhang, Z. Qin, L. Zhu, J. Weng, and K. Ren. 2017. Cost-Friendly Differential Privacy for Smart Meters: Exploiting the Dual Roles of the Noise. *IEEE Transactions on Smart Grid* 8, 2 (March 2017), 619–626. https://doi.org/10.1109/TSG.2016.2585963