Protecting Customer Privacy Through Distributed Energy Resource Anonymization

Nicole Henderson*, Midrar Adham*, Robert B. Bass* and Tylor Slay†
*Department of Electrical and Computer Engineering
Portland State University, Portland, OR 97123
Email: nhend2@pdx.edu
†Distributed Systems
Pacific Northwest National Laboratory, Richland, WA 99354

2023 IEEE Power & Energy Society General Meeting (PESGM) | 978-1-6654-6441-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/PESGM52003.2023.10253065

***Abstract*—Due to their stochastic nature, the increase of Renewable Energy Resources as primary sources of energy for power grids creates challenges regarding the reliability and resilience of the system. In order to combat these obstacles, expansion of Distributed Energy Resources (DERs) and their participation in grid services is necessary. Widespread participation requires prioritizing customer privacy and addressing concerns that may arise regarding communication between DERs and Grid Service Providers. Obtaining detailed information about customers' power consumption can lead to privacy risks that may prevent users from willingly participating in services. Anonymization of individual data is one method of privacy protection that should be explored. This paper discusses the use of the IEEE 2030.5 [1] flow reservation resources to split the operating cycles of DER load profiles into unique phases. The splitting of phases increases anonymization of DERs by making it more difficult to determine the individual characteristics of each device. We discuss the results of applying this form of anonymization to a set of simulated DER load profiles and examine the effectiveness of the anonymization through the use of a linear Support Vector Machine classifier.**

***Index Terms*—Distributed Energy Resource, Anonymization, Renewable Energy Resource, Flow Reservation, Support Vector Machine**

I. INTRODUCTION

As Renewable Energy Resources (RERs) are integrated into the bulk power system in an attempt to reduce dependence on fossil fuels, new challenges will arise that must be overcome in order to maintain reliability and resilience of electric power systems. The stochastic nature of RERs reduces flexibility and adds scheduling complexities to power generation. Ensuring a balance between the amount of energy being consumed and supplied by the grid is a complicated and expensive endeavor.

One method of maintaining an equilibrium in the supply and demand of a power grid is through the use of Demand Response (DR) services. DR services use customers' DERs to adjust system load in response to forecasted load imbalance. DR control is traditionally performed through Direct Load Control (DLC) [2]. Service-Oriented Load Control (SOLC) is a new method to achieve the same ends, based on Service-Oriented Architecture (SOA) [3]. According to the Grid Modernization Laboratory Consortium (GMLC), DERs are grid-enabled customer-owned generation, storage, and load assets [4]. Examples of these devices include refrigerators, clothes washers and dryers, and air conditioners.

Aggregating hundreds of thousands of DERs is an affordable option that increases the flexibility of power systems by offering essential reliability grid services. In order to get many devices participating in these services, customers must willingly sign up for participation. Adham [5] explains that the DLC approach gives utilities direct control over customers' devices, while an SOLC approach allows the customer to have ultimate control over their DERs. SOLC DR programs provide customers with the ability to opt out of services at any time, and all transactions must originate from the customer. Customers are more likely to join programs where they feel their needs and concerns are prioritized [6].

Additionally, consumers are more likely to participate in these services when their privacy and security are actively protected. Obtaining detailed information about customers' power consumption in order to manage energy supply comes with privacy concerns. The data collected could be used to predict sensitive information about the customer, such as when they use particular devices, sleep, or are away from home [7]. Due to these concerns, finding ways to anonymize and protect individual customer data should be a top priority.

The goal of this paper is to expand on the content presented in our previous work regarding incentivization of DER participation [8]. As DER programs are adopted by utilities, it is essential that an SOA approach is prioritized and trusting relationships are developed with Service Provisioning Customers (SPCs) to ensure enrollment and widespread participation. This paper focuses on ensuring customer privacy through DER anonymization. We explore anonymizing individual data by splitting DER operating cycles into individual phases of operation through the IEEE 2030.5 flow reservation request resource [1].

II. ENERGY SERVICE INTERFACE

The goal of the Energy Service Interface (ESI) is to allow bi-directional information exchange between a Grid Service Provider (GSP) and customers' DERs while maintaining privacy, security, and trust. In our previous paper, we explored the history of ESI and described our particular implementation, which emphasizes a SOA [8].

This work was supported by US DOE OE0000922.

The ESI is a set of rules that enable communication while setting boundaries between the GSP and DERs. The ESI provides clear definitions of the functions and responsibilities for both parties. While the ESI rules encompass privacy, security, and trustworthiness, this paper centers around the privacy aspect.

The ESI limits information exchange and the use of data, thereby prioritizing DER and customer privacy. Some examples of rules that are upheld to ensure privacy include [8]:

- The GSP only engages DERs on an opt-in basis.
- DERs initiate all communication with the GSP.
- DERs are able to decommit from a resource service at any time without penalty.
- The GSP does not record or discern DER usage patterns.
- Exchange of DER information is limited to the minimum required to implement a particular service.

Along with these rules, DER anonymization can be used to protect privacy by allowing DERs to make service requests that are similar to a variety of other DERs. This increases the similarity of the data, making it more difficult to determine individual characteristics.

III. DISTRIBUTED ENERGY RESOURCES

As mentioned, DERs are grid-enabled customer-owned generation, storage, and load assets. These devices, which include household appliances, battery energy storage systems, electric vehicles, and micro-generators, are owned by customers and are not traditionally controlled by a utility. DERs can be aggregated and have potential to provide improved flexibility and reliability to the bulk power system.

A. Flow Reservation Requests

The IEEE 2030.5 Smart Energy Profile outlines the flow reservation resource function set, which provides an interface for the exchange of "charge" or "discharge" events between DERs and the GSP [1]. This function set allows scheduling DER activity in order to spread out demand during peak periods by giving DERs a means to share the available energy they can produce or consume from the grid for a given time interval.

A flow reservation request takes information from each DER and determines four specific data points:

- **Power**: the maximum power capability.
- **Energy**: the projected integral of power consumed over the duration of the request.
- **Duration**: the time required for the device to serve the requested energy.
- **Interval**: the time window during which the duration must be completed [8].

The flow reservation resource inherently protects customer privacy by only sharing these four attributes, which do not give away Personally-Identifiable Information (PII) such as customer location or what type of DER the customer owns. Although this is helpful in protecting privacy, there are other concerns that require further action to ensure privacy.

The flow reservation request model also gives the option to split DER operating cycles into individual phases of operation to increase flexibility or anonymize requests by creating similar requests to other DERs. Figure 1 demonstrates how a flow reservation request can be broken up into distinct phases.

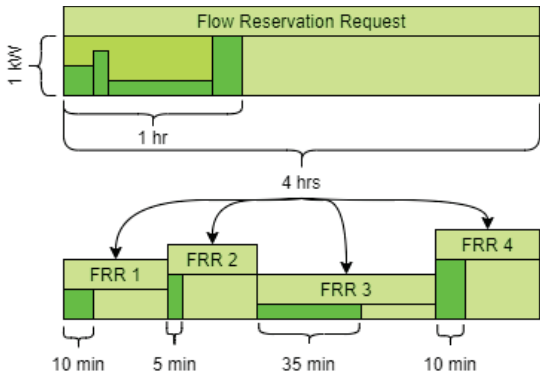

Fig. 1: Example of flow reservation requests for independent phases. Each phase is assigned an individual flow reservation request to increase flexibility [8].

B. Privacy Concerns

Customers may have apprehensions about participating in a DR program that requires their DERs to communicate data with an outside entity. According to Nissenbaum, concerns regarding privacy can be broken down into three categories [9]:

- Limiting surveillance and use of information.
- Restricting access to sensitive, personal, or private information (PII).
- Preventing intrusions in private or personal places.

Given that DERs are located in customers' homes, all three of these categories are necessary points of consideration.

One example of a privacy concern for customers is preventing outside sources from knowing when the customer is away from home. Ibrahem *et al.* propose transmitting redundant readings using an algorithm in order to protect privacy [7]. Kement *et al.* also discuss methods to mitigate this privacy problem by adding noise to metered data, using encryption techniques, and shaping metered loads [10]. Algorithms that allow the shaping of power consumed in a household would ensure that activity present in the data provided by the DERs does not show signs of a customer's absence.

Another privacy concern discussed by Zeifman *et al.* is the ability to identify the type of device and its manufacturer from the data provided by the DERs [11]. One way to prevent this release of PII is by having many points of data that look indistinguishable using a similarity algorithm. This would create groupings in the data, making it a challenge to discern the type of DER.

In our work, rather than using an algorithm to bring similarities, we induced anonymization by splitting the full load profiles of DERs into distinct phases of their operating cycles.

978-1-6654-6441-3/23/$31.00 ©2023 IEEE

C. DER Load Profiles

Load profiles describe the characteristics of a DER related to the flow reservation request components. The power, energy, duration, and interval values vary depending on the type of DER, the manufacturer, and the particular settings of the device. In this paper, the focus will be on dishwashers, air conditioners, water heaters, refrigerators, clothes dryers, and clothes washers.

Table I gives an overview of the primary characteristics of each DER during its operating cycle, including the approximate interval between use[1]. These values vary between device types, manufactures, and the settings of the device being used.

TABLE I: DER primary characteristics of operation for a Clothes Washer (CW), Clothes Dryer (CD), Air Conditioner (AC), Water Heater (WH), Dish Washer (DW), and Refrigerator (Rfg) [12].

Device	Power (W)	Energy (Wh)	Duration (min)	Interval (hr)
CW	278	99	60	12
CD	2951	1440	60	4
AC	1107	186	10	1
WH	4531	506	7	1
DW	1172	438	60	12
Rfg	523	111	13	1

For example, Figure 2 demonstrates the variations in load profiles for a clothes dryer in two models operating in two different cycle modes. The clothes dryer DER type was chosen for its wide range in both power and operating time. The LG clothes dryer demonstrates a considerably higher power draw than the GE counterpart, with a much higher volatility during its operation. The GE *auto-perm* and *auto-regular* modes of operation appear to be nearly identical in power consumption with the primary difference being the time between each phase of operation.

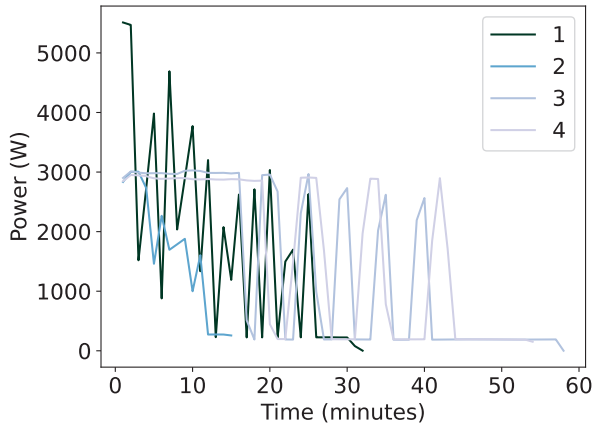

Fig. 2: Load profiles of two clothes dryers in two different modes. 1 and 2 are the LG dryer in auto-regular and delicates respectively. 3 and 4 are the GE dryer in auto-perm and auto-regular respectively. [12].

[1]Data set for [12]: https://ari.vt.edu/research-data.html

D. Anonymization and Classification

Flow reservation requests provide the same range of values for all DERs. Anonymization occurs when flow reservation requests are made to a variety of DERs. As more flow reservation requests are made, there is an increase in overlap between the load profiles of DERs, thereby protecting individual details.

In order to determine the level and effectiveness of anonymization, a classification must be defined. For this research, a Support Vector Machine (SVM) was used to categorize our simulated data. SVM is a machine learning algorithm that can be used in classification problems with large amounts of data [13]. This classification method allows the creation of visual aids to examine groupings of data and how closely the values correlate.

E. Data Simulation

We were unable to obtain a large set of real-time data for various DERs, including their respective manufacturers and operating modes. To validate our proposed solution, we have simulated 10 additional sets of data for each DER type. We made the following assumptions:

- The randomized data fall within the range of properties for existing data.
- If there was only data for one DER of a given type, there would be an approximate 5% variance for each property.

F. Classification

The base set of data with no changes made to it resulted in the linear SVM scatter plot with classification regions shown in Figure 3. The device categories are clearly separated, showing that they are easy to differentiate between.

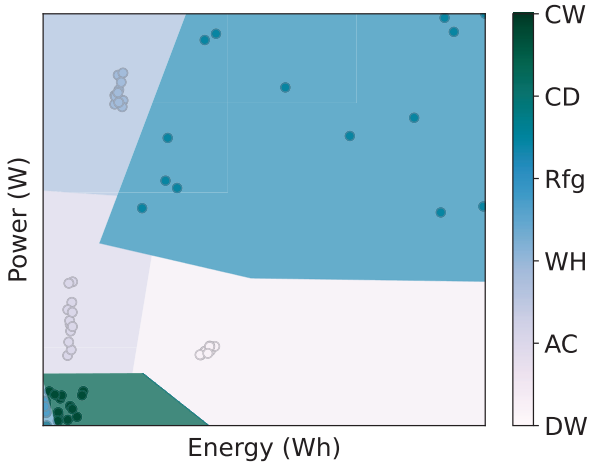

Fig. 3: Scatter plot with linear SVM classification applied to the base set of data before phases have been split. Color groupings include data from the following devices: Clothes Washers (CW), Clothes Dryers (CD), Air Conditioners (AC), Water Heaters (WH), Dish Washers (DW), and Refrigerators (Rfg).

978-1-6654-6441-3/23/$31.00 ©2023 IEEE

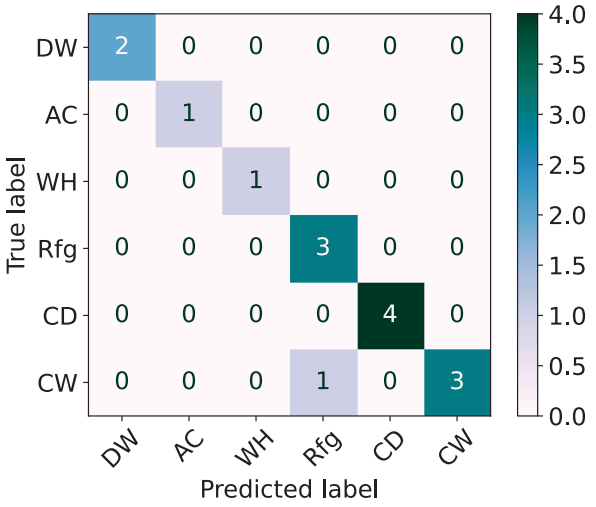

Fig. 4: Confusion matrix for the base set of data before the phases have been split. Includes data from the following devices: Clothes Washers (CW), Clothes Dryers (CD), Air Conditioners (AC), Water Heaters (WH), Dish Washers (DW), and Refrigerators (Rfg).

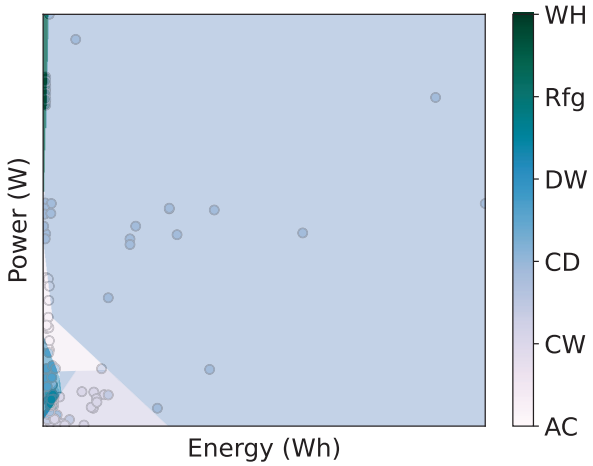

Fig. 5: Scatter plot with linear SVM classification regions for the data set with split phases. Color groupings include data from the following devices: Clothes Washers (CW), Clothes Dryers (CD), Air Conditioners (AC), Water Heaters (WH), Dish Washers (DW), and Refrigerators (Rfg).

Figure 4 shows the corresponding confusion matrix for the base set of data. The purpose of the confusion matrix is to show the number of instances in which the actual device type and predicted device type overlap. This allows the effectiveness of the anonymization of data to be determined. The more predicted types a true device can correlate with, the more difficult it becomes to identify what that device is.

As can be seen in Figure 4, only the true clothes washer correlates with two different predicted categories: refrigerators and clothes washers. All the other devices only correlate with their matching predicted devices, thus there is a low level of anonymization.

When the full load profile values of the DERs operating cycles were split into distinct phases, the data became more complex. The split data resulted in the scatter plot with linear SVM classification regions as shown in Figure 5. This Figure shows that there is intense overlapping of data points. As such, the specific devices are more difficult to differentiate between.

This set of data with the split in phases resulted in the confusion matrix shown in Figure 6. There is far more overlap between true device categories and predicted devices. The clothes washer, dishwasher, and refrigerator all increased their anonymity by splitting their phases of operation into separate flow reservation requests. The most successful was the dishwasher, which was predicted to be a clothes washer the majority of the guesses. By splitting individual device phases of operation into multiple flow reservation requests, the similarity in data points to other devices, manufacturers, and modes will increase.

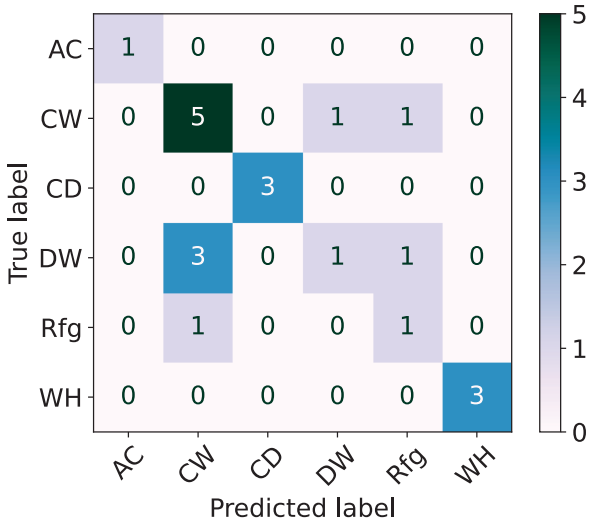

Fig. 6: Confusion matrix for the data with split phases. Includes data from the following devices: Clothes Washers (CW), Clothes Dryers (CD), Air Conditioners (AC), Water Heaters (WH), Dish Washers (DW), and Refrigerators (Rfg).

IV. CONCLUSION

Prioritizing customers' privacy and comfort is essential in order to increase DER participation and ultimately improve the reliability of power grids. This paper builds upon our previous work and explains why DER anonymization would be valuable for protecting customer privacy [8].

Inducing anonymization of DER characteristics can be accomplished by splitting device phases of operation into multiple flow reservation requests. Finding simple methods of anonymizing SPC information will assist in the process of protecting customer PII.

978-1-6654-6441-3/23/$31.00 ©2023 IEEE

We discussed how splitting the operating cycles of DER load profiles into distinct phases is a straightforward approach to anonymizing DER data by creating overlap in the flow reservation request properties between various DERs. The SVM and confusion matrix plots highlight the effectiveness of anonymization in making it difficult to determine individual DER device types. When particular DER devices share characteristic values with other types of devices, customer privacy is better protected.

Future work will focus on obtaining more data on additional DERs, including models and modes of operation to find the upper limit to anonymization. Building on the work presented in this paper and the exploration of additional forms of anonymization for devices that were not impacted by the splitting of operation cycles should be explored.

REFERENCES

[1] "IEEE standard for smart energy profile application protocol," *IEEE Std 2030.5-2018 (Revision of IEEE Std 2030.5-2013)*, 2018.

[2] J. Stitt, "Implementation of a large-scale direct load control system," *IEEE Trans. Power App. & Sys.*, vol. 104, no. 7, pp. 1663–1669, 1985.

[3] S. Jones, "Toward an acceptable definition of service [service-oriented architecture]," *IEEE Software*, vol. 22, no. 3, pp. 87–93, 2005.

[4] Grid Modernization Laboratory Consortium, "Interoperability strategic vision, a GMLC white paper," Pacific Northwest National Laboratory, Tech. Rep., Mar 2018.

[5] M. A. Adham, M. Obi, and R. B. Bass, "A Field Test of Direct Load Control of Water Heaters and its Implications for Consumers," in *2022 IEEE Power & Energy Society General Meeting*, 2022.

[6] S. Widergren, R. Melton, A. Khandekar, B. Nordman, and M. Knight, "The plug-and-play electricity era: Interoperability to integrate anything, anywhere, anytime," *IEEE Power & Energy Mag.*, vol. 17, no. 5, pp. 47–58, 2019.

[7] M. I. Ibrahem, M. Mahmoud, M. M. Fouda, F. Alsolami, W. Alasmary, and X. Shen, "Privacy preserving and efficient data collection scheme for AMI networks using deep learning," *IEEE Internet of Things J.*, vol. 8, no. 23, pp. 17 131–17 146, 2021.

[8] T. Slay, J. M. Acken, and R. B. Bass, "Incentivizing distributed energy resource participation in grid services," 2022.

[9] H. Nissenbaum, "Privacy as contextual integrity," *Wash. L. Rev.*, vol. 79, p. 119, 2004.

[10] C. E. Kement, M. Ilić, H. Gultekin, C. T. Cicek, and B. Tavli, "Privacy protection via joint real and reactive load shaping in smart grids," *Sustainable Energy, Grids and Networks*, vol. 32, p. 100794, 2022.

[11] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," *IEEE Trans. Consumer Electronics*, vol. 57, no. 1, pp. 76–84, 2011.

[12] M. Pipattanasomporn, M. Kuzlu, S. Rahman, and Y. Teklu, "Load profiles of selected major household appliances and their demand response opportunities," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 742–750, 2014.

[13] D. A. Pisner and D. M. Schnyer, "Support vector machine," in *Machine Learning*. Elsevier, 2020, pp. 101–121.

978-1-6654-6441-3/23/$31.00 ©2023 IEEE