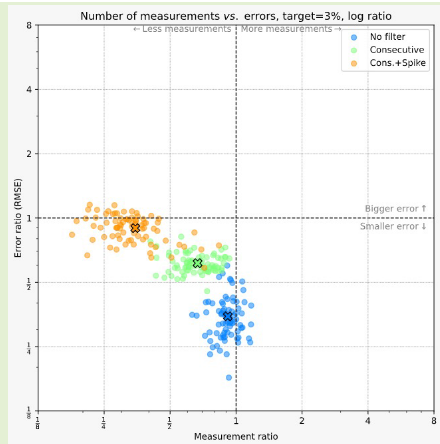
Event-Driven Data Acquisition for Electricity Metering: A Tutorial

Mauricio de Castro Tomé, Daniel Gutierrez-Rojas, Pedro H. J. Nardelli, *Senior Member, IEEE*, Charalampos Kalalas, *Member, IEEE*, Luiz Carlos Pereira da Silva, and Ari Pouttu

***Abstract*—This paper provides a tutorial on the most recent advances of event-driven metering (EDM) while indicating potential extensions to improve its performance. We have revisited the effects on signal reconstruction of (i) a fine-tuned procedure for defining power variation events, (ii) consecutive-measurements filtering that refers to the same event, (iii) spike filtering, and (iv) timeout parameter. We have illustrated via extensive numerical results that EDM can provide high-fidelity signal reconstruction while decreasing the overall number of acquired measurements to be transmitted. Its main advantage is to only store samples that are informative based on predetermined events, avoiding redundancy and decreasing the traffic offered to the underlying communication network. This tutorial highlights the key advantages of EDM and points out promising research directions.**



***Index Terms*— Event-driven data acquisition, smart grids, advanced electricity metering.**

I. Introduction

HOUSEHOLD electrification is now widespread: nine out of ten people worldwide have access to electricity [2]. In developed countries, availability of electricity to the general population approaches 100%, while in developing countries, the percentage is steadily rising. Electricity metering thus becomes fundamental not only for billing purposes but also for monitoring the operational condition of power grid components.

Traditionally, electricity use from consumer units (e.g., households, buildings and industries) has been manually collected: a person is required to move on site and check the electricity meter at the entrance of the place, write down the cumulative consumption, and subtract it from the previous reading to get the consumption for a time period. In this case, the *amount* of energy consumed within that period (usually one month) is what matters, regardless of *how* and *when* that energy was consumed over time. Energy consumption patterns may widely vary among the different consumers. Certain user profiles are associated with nearly constant consumption over time exhibiting only minor variations, whereas for other users energy can be extensively consumed within just a few days of the month or in short periods at specific times.

Recent advancements in communication and computation technologies have rendered the automated measurement acquisition feasible, reducing the need for human intervention. In addition, emerging demand-response programs gradually rely on more frequent measurements to constantly track the consumption at the consumer side in response to supply conditions regulated by utility providers [3]. Nevertheless, the rationale behind this automated data acquisition resembles the manual method, because measurements are generated in a content-agnostic manner without exploiting the inherent attributes of metering information. In principle, these measurements are sent in periodic time intervals and

Manuscript received October 26, 2021; revised January 24, 2022; accepted January 25, 2022. Date of publication January 27, 2022; date of current version March 14, 2022. This work was supported in part by the Academy of Finland (AKA): 6G Flagship, and EnergyNet Fellowship under Grant 321265 and Grant 328869; in part by the FIREMAN Project CHIST-ERA-17-BDSI-003 funded by AKA under Grant 326270; and in part by the Spanish National Foundation under Grant PCI2019-103780. The associate editor coordinating the review of this article and approving it for publication was Prof. Rosario Morello. *(Corresponding author: Pedro H. J. Nardelli.)*

Mauricio de Castro Tomé and Ari Pouttu are with the Centre for Wireless Communications, University of Oulu, 90014 Oulu, Finland.

Daniel Gutierrez-Rojas is with the Department of Electrical Engineering, Lappeenranta–Lahti University of Technology, 53850 Lappeenranta, Finland.

Pedro H. J. Nardelli is with the Centre for Wireless Communications, University of Oulu, 90014 Oulu, Finland, and also with the Department of Electrical Engineering, Lappeenranta–Lahti University of Technology, 53850 Lappeenranta, Finland (e-mail: pedro.nardelli@lut.fi).

Charalampos Kalalas is with the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), 08860 Barcelona, Spain.

Luiz Carlos Pereira da Silva is with the School of Electrical and Computer Engineering, University of Campinas, Campinas 13083-970, Brazil.

Digital Object Identifier 10.1109/JSEN.2022.3147016

are henceforth defined as *time-based* measurements. Another option is to transmit new measurements whenever a fixed amount of energy has been consumed, i.e., often referred to as *energy-based* measurements, without any additional specification of the energy consumption profile within that period of time [4]–[6]. With these strategies, the only way to increase the fidelity of the signal reconstruction is to send more frequent measurements, either by increasing the temporal resolution, or, in the second case, by reducing the size of the reported "energy block". If we extend this line of reasoning, it would be required to have infinitesimal time slices or energy blocks to achieve highly accurate representation of the energy consumption. This would result in significant overload levels for the underlying communication infrastructure (e.g., [7]) since each household would act as a source of tens of thousands of measurements per day; a challenge even for communication systems designed for machine-type communication, including 5G.

By inspecting the metering signal characteristics in the time domain, one can observe that it has essentially a flat shape for most of the time, only with abrupt changes that are caused by the switching of electrical appliances. Conventional time-based measurements may fall short of representing the real energy consumption in a meaningful way. Too few measurements representing the average power over a long period of time, may have almost no similarity with the actual consumption pattern. On the other hand, too many measurements to capture the power variations, may result in several consecutive measurements having very similar – or even identical – power values. An alternative to the aforementioned methods is the so-called event-driven data acquisition [8], where *(…) a sample is produced only when something significant (an "event") occurs in the signal*. That is, the information about the state of the signal is sent only when it exceeds a given predetermined threshold. In stark contrast with conventional periodic-sampling strategies, this approach capitalizes on the largely unexploited intrinsic features of the conveyed information, which influence its relevance and usefulness. Event-driven data acquisition has been widely used in various disciplines, ranging from signal processing [9], [10] to process control [11], [12], and communications [13]–[15]. In the electricity metering domain, event-driven data acquisition, often referred to as event-driven metering (EDM), has been recently employed in [16]–[23]. In addition to these key developments, we can cite the following papers that use event-driven processing for smart meter data analysis [24]–[26].

The events under consideration in the aforementioned papers refer to either instantaneous power transitions or energy variations, both triggered whenever the relevant values cross predetermined thresholds. The EDM method has the potential to significantly reduce the metering data volume generated and transmitted, as far as the values falling under the threshold are discarded instead of being stored and/or sent, causing minimum impact on the quality of signal reconstruction. An important limitation of EDM, though, is that in its original implementation, the thresholds are determined beforehand (and hence, are fixed) from a group of samples that are collected during a time interval. Consequently, any changes in the composition of consumer units (such as the addition or removal of electrical appliances, or the change of the seasons, to name a few) can largely influence the number of measurements sent by the consumer. To the best of the authors' knowledge, this limitation has not been adequately addressed in existing related literature.

This contribution is a tutorial on recent improvements of EDM based on a semantics-aware approach that harnesses intrinsic contextual attributes of metering information. The main objective is to demonstrate how the number of measurements generated by each customer is kept within a given range, by using their own consumption to determine the power variation thresholds. By exploiting this knowledge in semantics-empowered EDM, we identify ways to reduce the acquired data by filtering those measurements that trigger events with low impact on the signal reconstruction process.

The remainder of this paper is organized as follows. Section II provides an overview of the EDM technique and highlights our proposed modifications. Section III offers a comprehensive and in-depth performance assessment of the different EDM enhancements based on energy consumption data provided by the Pecan Street project (https://www.pecanstreet.org/dataport/). Section IV is reserved for concluding remarks related to the improvements of the EDM method and potential future research directions.

II. Event-Driven Metering: Overview

The EDM technique, mainly developed by Simonov and his collaborators in a series of papers [16]–[18], [22], [23], is built upon the three following events which are relevant to the signal reconstruction:

- *Instant power variations,* ΔP, that capture abrupt differences in power consumption levels, with a corresponding threshold δ_P;
- *Energy deviation,* ΔE, defined as the energy evolution that diverges from the expected trend, with an associated threshold δ_E;
- *Fixed time measurements*, at regular time intervals of length T, for compatibility with legacy metering.

It is worth noting that *absence* of data is also informative in EDM. If no data is acquired, it can be inferred that consumption is to some extent constant during the respective window defined by T, as it never exceeds the thresholds determined by ΔP and ΔE.

Fig. 1 illustrates how the EDM mechanism works for an arbitrary scenario where δ_P is configured such that it never triggers an event and there are no fixed time measurements. The original signal is presented in the first plot. In this example, the power threshold is never crossed, as shown by the grey spikes in the second plot. In the third plot, the cumulative energy deviations that trigger events (i.e., the sudden changes after a linear increase or decrease) can be observed. The fourth plot depicts the reconstructed signal using EDM. Finally, the quality of reconstruction is evaluated in the fifth plot, where the reconstruction error is presented as a comparison between the first and the fourth plot.

In this example, we mainly focused on the behavior of the cumulative energy deviation. This quantity represents the accumulated energy that exceeds the *expected trend* given by

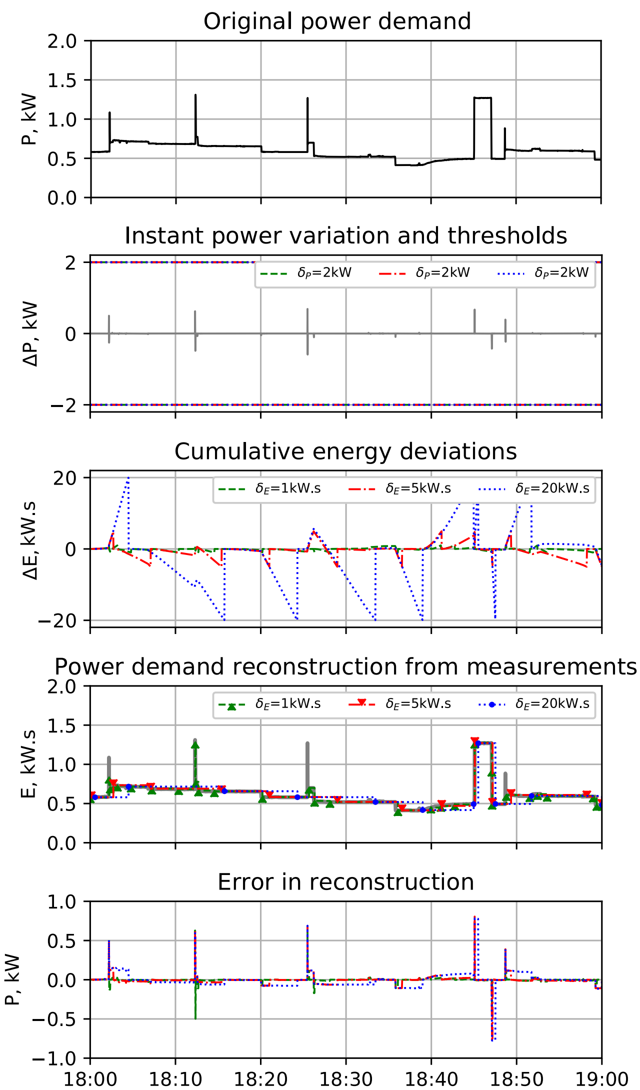


Fig. 1. Example of the EDM mechanism for arbitrary threshold values δ_P = (2kW, 2kW, 2kW) and δ_E = (1kWs, 5kWs, 20kWs). First top: original signal; second top: power variation; middle: cumulative energy deviations; second bottom: reconstructed signal from EDM; first bottom: error in the reconstruction. Adapted from [1].

the transmitted power value in the previous time instant. That is, given that a certain event $P(t)$ triggered a measurement at time t, the value of the energy deviation can be calculated as

$$\Delta E = \left|\sum_{k+1}^{\tau} P(t+k) - P(t)\right|. \tag{1}$$

In case the sum in Eq. (1) exceeds the threshold value given by δ_E, an event is generated. Accordingly, whenever any of the δ_P, δ_E, or T thresholds is reached, a measurement is sent, the value of the accumulated energy deviation is reset, and the current power value $P(t)$ is updated.

An extension to the EDM method was recently proposed by Tomé *et al.* in [21], aiming at modifying the acquisition procedure to facilitate signal reconstruction. In addition, the authors have introduced alternative ways to reduce the data volume generated, and remove redundancy during acquisition, i.e., by eliminating measurements referring to the same event

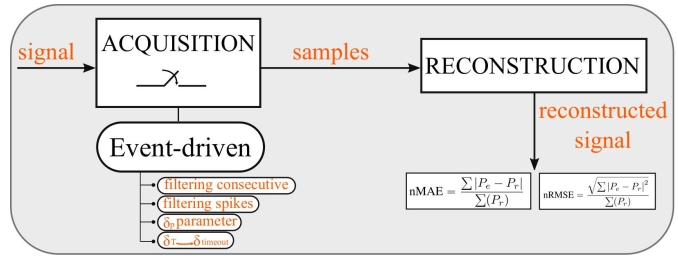


Fig. 2. Framework and functional blocks of EDM.

and adding a timeout event. In that paper, though, the proposed enhancements were not investigated in detail. Here, we further extend the initial results and derived insights by conducting an in-depth assessment of four key operational techniques in EDM, namely: (i) estimation of the power variation threshold, (ii) filtering of consecutive measurements, (iii) filtering of spikes, and (iv) timeout parameter. Moving beyond existing contributions, we have utilized the Pecan Street database as input to this work. It constitutes a well-known data source in the research community and, to the best of our knowledge, it provides an extensive set of measurements to accurately evaluate the proposed methodology.

III. Event-Driven Metering Enhancements: A Comprehensive Assessment

In this section, we will sequentially evaluate the impact of four different enhancements in EDM method: (i) estimation of the power variation threshold δ_P, (ii) filtering of consecutive measurements, (iii) spike filtering, and (iv) timeout parameter. The numerical results are based on long-term simulations of Pecan Street households following the approach depicted in Fig. 2. The employed dataset consists of 1-second (1s) measurements and the original signal (i.e., input of the EDM) is already sampled with a frequency of 1Hz; thus, the proposed data acquisition step is based on such a frequency. Note that this time granularity is not mandatory for the EDM, and thus, it could also be deployed even in minute timescales (although this would preclude its key benefits of identifying sudden variations). A detailed discussion about the physical implementation of EDM is given in [17], [22], [23]. Here, our objective is to provide a comprehensive tutorial of improvements in EDM based on the (algorithmic) definition of the events in a semantic manner, considering also different types of filters and individual definition of thresholds.

As shown in Fig. 2, the raw measurements (signals) captured by the metering devices at each consumer point, pass through the data acquisition component where EDM is applied. Then, the samples obtained via EDM advance to the reconstruction component, where the signal is rebuilt based on the received samples through linear interpolation. To evaluate the quality of the signal, we compare its reconstructed version with the original signal (whose time granularity is 1s) utilizing the *normalized* Mean Average Error (nMAE) and Root-Mean-Square Error (nRMSE) metrics, defined as

$$\text{nMAE} = \frac{\sum |P_e - P_r|}{\sum (P_r)}, \tag{2}$$

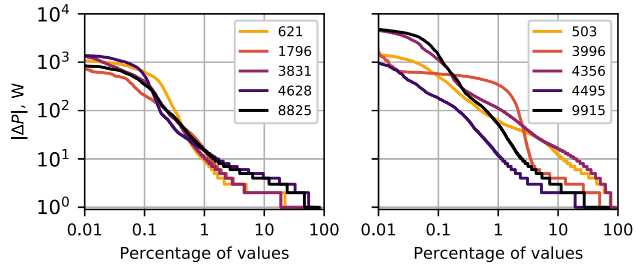


Fig. 3. Instant power differences for a set of houses in the Pecan dataset. Left: houses with similar profile. Right: houses with different profiles. Source [1].

$$\text{nRMSE} = \frac{\sqrt{\sum |P_e - P_r|^2}}{\sum (P_r)}, \tag{3}$$

respectively. In Eqs. (2)-(3), P_r denotes the real power consumption and P_e is the estimation from the measurements.

A. Estimation of Threshold δ_P

A proper estimation of δ_P allows the meter to adapt to different consumption patterns by exploiting the availability of historical consumption data. An example of instant power differences during a week for a set of houses in the dataset is presented in Fig. 3. It can be observed that for houses with similar consumption behavior, as the ones selected in the left plot, setting a fixed trigger would yield a similar number of measurements. For example, setting $\delta_P = 10$W would result in about 1%-2% of the data being transmitted. For houses with entirely different patterns, as the ones selected in the right plot, setting the same (10W) threshold would capture 1% to 20% of the measurements, depending on the house.

It is important to note that this proposal constitutes an improvement compared to the methods described in [19], [21]. The idea here is to develop an adaptive method that defines the thresholds δ_P based on the individual historical profile of each household. In this case, the proposed algorithm first defines such a value considering a specific training dataset and, afterwards, it dynamically updates it by taking into account the actual realizations for a given set of parameters. Such parameters will determine the event under specific policies which differentiate daily profiles.

In this particular case, instead of calculating the triggers in a single week and applying the same value in all days of the following week, it is also possible to perform the computation for each day separately during an extended period. In this context, we have carried out the reading of households' consumption and the corresponding simulation in several ways:

- A fixed-time simulation (considered as the reference), denoted by **T** in the labels;
- A "genie" simulation, labeled **G**, in which the own consumption of the day is used to determine the optimal value of δ_P that would result in the desired amount of measurements;
- A setup taking into account the consumption of the previous day to determine δ_P for the following one (e.g., using the data from Monday to estimate the threshold for Tuesday), labeled **1**;

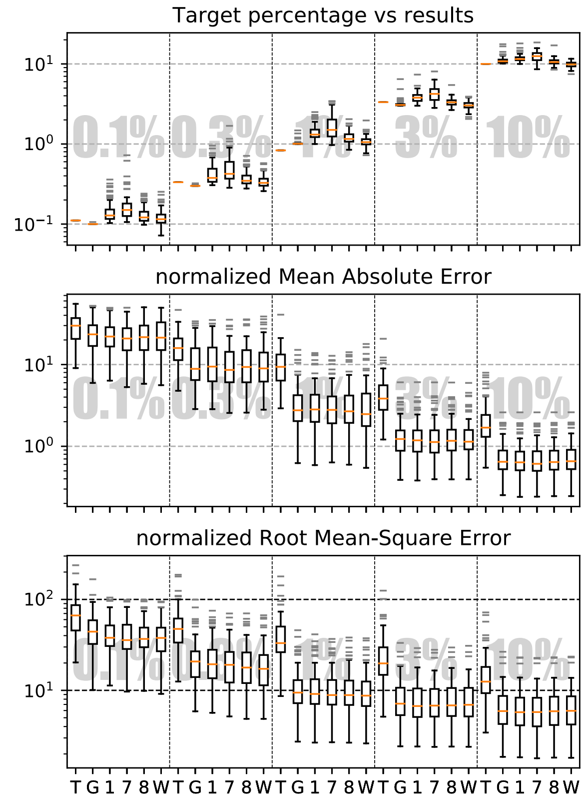


Fig. 4. Distribution of the average number of measurements (top), and normalized values of MAE (middle) and RMSE (bottom) for the 84 houses simulated from the Pecan database. Legend: T: fixed time, G: genie, 1: previous day, 7: same weekday, 8: both previous day and weekday, W: average of the last 7 days. The gray dashes (–) represent outliers. Source [1].

- A setup considering the consumption of the same weekday from the previous week (e.g., using the data from last Sunday in order to calculate the trigger for this Sunday), labeled **7**;
- Using the average of the previous day (1) and the same weekday from the previous week (7), labeled **8**;
- Using the average of the triggers from the previous 7 days, labeled **W**.

The simulation was conducted for all the houses for a period of 5 weeks. In particular, the first week was used for data collection to obtain the initial δ_P triggers (i.e., in this simulation it is the only parameter being tracked), which were sequentially applied in the following weeks. The measurements from the first week were then discarded, and the remaining 4 weeks were used to calculate the average amount of measurements for each house, as well as the corresponding reconstruction errors. The target number of measurements simulated were 0.1%, 0.3%, 1%, 3%, and 10% (or, for the fixed-time baseline, 900s, 300s, 120s, 30s, and 10s, respectively, which would result in a similar number of measurements). The reconstruction performance is evaluated with the aid of nMAE and nRMSE metrics.

In Fig. 4, we can observe that, while all event-driven strategies result in similar error quantities, only strategy **T** (i.e., the reference one, consisting of a fixed time threshold) exhibits a visually worse performance. It is also worth noting that almost all the compression targets in the event-driven strategy

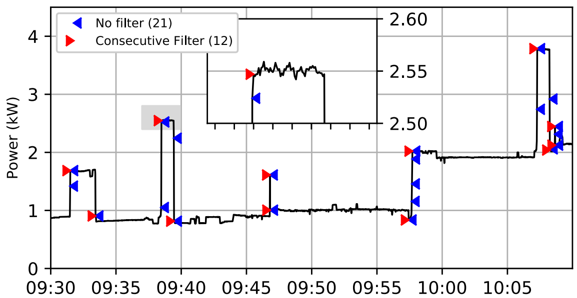


Fig. 5. Example of the consecutive-measurements filter. Source [1].

result in errors within the same range as the subsequent fixed-time target. For instance, the nMAE value for the target of 0.3% in the EDM strategy is similar to the nMAE value of the time-based target of 1%. This becomes even more pronounced when examining the nRMSE metric. While it is intricate to fully interpret its value, it provides a measure of how well the strategies follow large variations in power demand. In any case, what is worth to mention is that the definition of thresholds is essential to improve (optimize) EDM performance; this can be realized by the individualized dynamic method that considers different daily profiles, as demonstrated by the comparison with the fixed time policy **T**.

In the following sections, we will only simulate a single threshold target, i.e., 1% of the original resolution using the average of the last 7 days' triggers.

B. Filtering of Consecutive Measurements

An important limitation of the original EDM implementation is that consecutive power variations that exceed δ_P threshold would generate multiple samples, which greatly increase the amount of measurements while adding negligible information usefulness/value to the data. To deal with this issue, we have implemented a consecutive-measurements filter. It works by storing the measurement that triggered a power event and waiting for the following measurement to check whether it qualifies as a *consecutive* measurement, adding memory in the data acquisition procedure. A pair of subsequent measurements is considered consecutive when

1) $|\Delta P_{t,t-1}| > \delta_P$, and
2) $\text{sgn}(\Delta P_{t+1,t}) = \text{sgn}(\Delta P_{t,t-1})$,

are both satisfied. If the measurements fulfill the aforementioned criteria, the previous measurement is discarded and the current one is stored. Enabling the filter introduces a small delay of one time unit in the transmission of the measurement (i.e., 1s for this database), since it requires the next reading to be acquired and processed before deciding whether to send the current measurement or update it.

An example of applying the filter is shown in Fig. 5, where we can observe a small sample of house 59 from the Pecan database, with the δ_P threshold configured to 1% of the daily measurements (which corresponds to 126W for this house and day). Besides reducing the amount of measurements, the filtered case reaches the top of the peak offering a better representation of the signal with respect to the normal case. Fig. 6 shows the percentage of consecutive measurements that

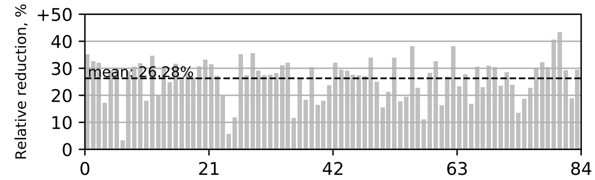


Fig. 6. Percentage reduction when the consecutive-measurements filter is ON. Source [1].

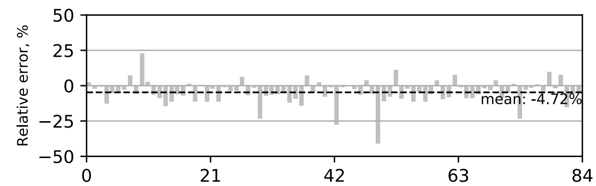


Fig. 7. Relative variation in error when the filter is ON. Source [1].

occur when $\delta_P = 1\%$ is set for each one of the houses. A set of houses demonstrate short reduction in the number of measurements when the filter is activated; however, the majority of the houses show reduction in the order of 20% or more, in the number of measurements. Fig. 7 illustrates the impact of the consecutive-measurements filter on the average relative error performance. Although some houses exhibit an increased error due to the lower number of measurements, the dominant trend reveals a reduction of the average error. This can be intuitively explained by the fact that the filter discards some transient measurements while improving the tracking of steady-state values, as shown in Fig. 5.

C. Filtering of "Spikes"

Certain types of loads, such as electric motors and fridges, are characterized by sharp transients when switched on, but such values quickly converge to their steady-state values. These quick transients are commonly known as *spikes* and, due to their characteristics, may cause an excessive number of measurements when conventional EDM is applied. In our own EDM implementation, we have introduced a *spike filter*; when this filtering is active, such sharp transients are ignored until a steady-state value is reached. A pair of measurements is considered to be a spike when all the following criteria are satisfied:

1) $\text{sgn}(\Delta P_{t+1,t}) = -\text{sgn}(\Delta P_{t,t-1})$,
2) $|\Delta P_{t+1,t}| > \delta_P$ and $|\Delta P_{t,t-1}| > \delta_P$,
3) $|\Delta P_{t+1,t-1}| < \delta_P$.

Fig. 8 depicts the effect of the spike filtering technique in comparison to the consecutive-measurements filter, using the same time interval as in Fig. 5. In particular, as illustrated in the detail of Fig. 8, we have one spike and the resulting measurements from the consecutive-measurements filter. The spike filter detected such type of event and, since the power difference before and after that particular spike is smaller than the δ_P threshold (i.e., 126W as in the previous case), no measurement is transmitted. Otherwise, if the final value exceeded the detection threshold, a single measurement would have been sent (the rightmost one), instead of the two measurements in the case of the consecutive-measurements filter.

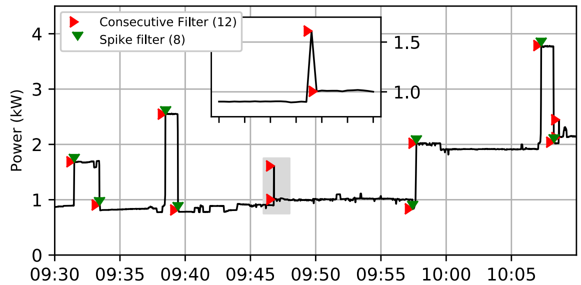


Fig. 8. Example of spike filtering. Source [1].

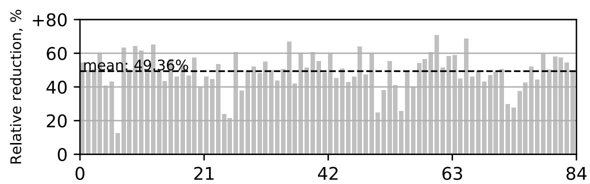


Fig. 9. Relative reduction in measurements when both filters are ON. Source [1].

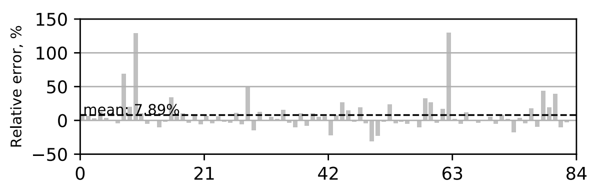


Fig. 10. Relative nMAE variation when both filters are ON. Source [1].

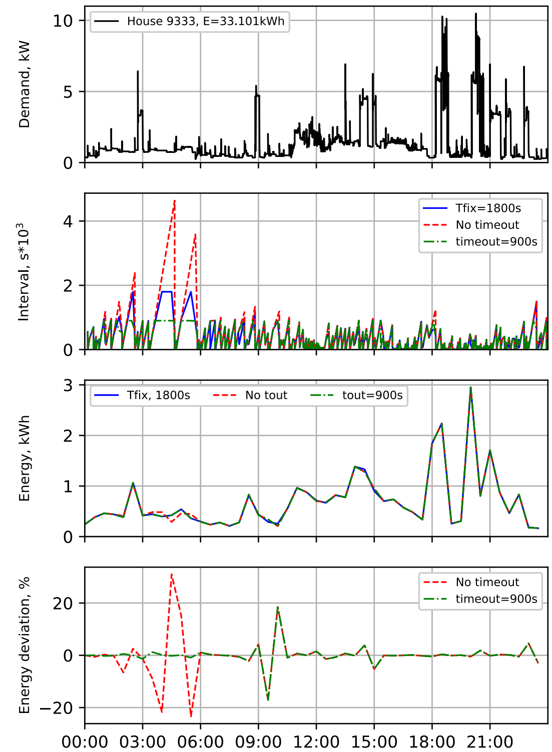


Fig. 11. Deviation in energy report due to *extrapolation* of the previous measurement. From top to bottom: Power demand; Time interval between samples; Energy consumption for the period; Deviation from the actual measurement. Source [1].

Figs. 9 and 10 show the relative reduction in measurements with respect to the baseline case and the nMAE variation due to the reduction in measurements, respectively. It can be observed that the average number of measurements is reduced from approximately 25% when only the consecutive filter is applied, to approximately 50% when both filters are activated, but the average error greatly increases. One way to circumvent this would be to increase the desired amount of measurements (that is, lowering δ_P) to compensate for the sharp reduction in the number of measurements.

D. Replacing δ_T With $\delta_{timeout}$

In the original EDM implementation, a fixed time measurement δ_T was included for compatibility with legacy metering and billing. Although useful for that purpose, we argue that these measurements insert unnecessary redundancy in the data; such information can be easily derived from other measurements, provided that there are adequate points to represent the demand. In addition, using a fixed time measurement would result in *all* houses periodically sending measurements at exactly the same time. Instead, we propose to replace δ_T with a *timeout* parameter, $\delta_{timeout}$, which is sent when no other events have been detected for a long time. This approach is beneficial in two ways: (i) it reduces the probability of transmitting several measurements simultaneously; (ii) it guarantees that we will have a periodic update in the demand of the houses even in the absence of other events.

Fig. 11 illustrates the impact of these parameters on the periodic energy consumption report. We have simulated a series of measurements using $\delta_P = 1\%$. The timed measurements were set to either none at all, $\delta_T = 1800$s, or $\delta_{timeout} = 900$s, and were compared with the periodic reporting every 1800s. The periodic reporting, when there are no fixed time measurements, is extrapolated from the last measurement that occurred *before* the reporting period. We can observe that the absence of any type of timed measurement causes a deviation in the energy reporting, especially during times with low activity (for example, between 03:00 and 06:00); instead, when $\delta_{timeout}$ is applied, the deviation is greatly reduced.

When there is no need for immediate reporting of the energy consumption, as it is usually the case in energy billing, we can rather interpolate the measurements which occurred both before and after the reporting period, as shown in Fig. 12. This, in turn, results in a significant reduction of the deviation between the actual and inferred measurements, even when $\delta_{timeout}$ is not set. Table I shows the deviation in the energy report and the error in signal reconstruction for each one of the previous cases. Replacing δ_T with $\delta_{timeout}$ results in a slight decrease in the number of measurements and has negligible impact on both energy report and reconstruction error.

When applying this modification to a population of houses, a remarkable reduction in the peak amount of measurements can be achieved, since the fixed time reporting is excluded. Fig. 13 shows a simulation of one day applying the same parameters, i.e., $\delta_P = 1\%$ and setting either $\delta_T = 1800$s or

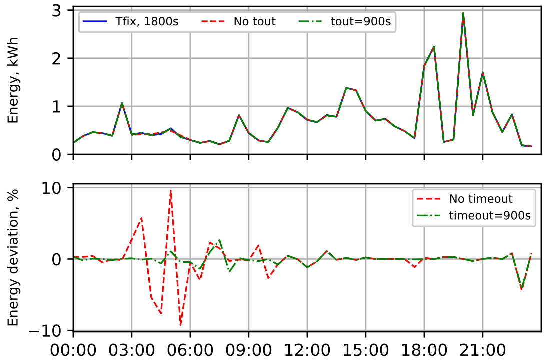


Fig. 12. Deviation in energy report due to *interpolation* of measurements. Top: Energy consumption for the period; Bottom: Deviation from the actual measurement. Source [1].

TABLE I

COMPARISON OF TIMED SAMPLES FOR HOUSE 9333 WITH $\delta_P = 1\%$ AND 30MIN REPORTING. SOURCE [1]

Name	# of samples Power	Time	E deviation, kWh Max	mean	Reconstruction Error MAE	RMSE
TDM, 1800s	–	49	–	–	–	–
EDM, δ_T=1800s	862	49	0.0000	0.0000	0.0444	0.1096
Using extrapolation						
EDM, no timeout	864	–	0.1313	0.0100	0.0513	0.1162
EDM, $\delta_{timeout}$ = 900s	864	27	0.0499	0.0041	0.0432	0.1116
Using interpolation						
EDM, no timeout	864	–	0.0368	0.0040	0.0513	0.1162
EDM, $\delta_{timeout}$ = 900s	864	27	0.0106	0.0013	0.0432	0.1116

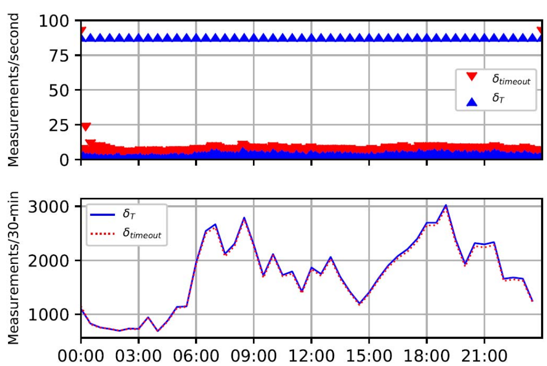


Fig. 13. Measurements per second (top) and per 30min interval (bottom) for the Pecan dataset. Source [1].

$\delta_{timeout}$ = 900s, for the simulated neighborhood of the Pecan database. It can be observed that the peak of simultaneous measurements went from approximately 90 (corresponding to the number of houses) to less than 20 measurements,[1] while a smaller number of measurements, i.e., from 83425 to 81989, was attained. Although a reduction of about 2% in the total number of measurements may not sound significant, the 3- to 6-fold reduction of simultaneous measurements is of utmost importance for the underlying communication infrastructure. As the generated traffic becomes akin to a continuous flow of few messages, instead of sporadic time instants with high bursts of messages, congestion is alleviated and the reliability of transmitted data can be improved. More details of the relation between communication and EDM can be found in [20]. Note that, although EDM cannot provide significant advantage in terms of computation directly, a properly designed EDM can decrease the traffic offered to the communication network and improve the performance of data transfer from metering devices to a fusion center for further processing.

[1]The first and last values also add up to the number of houses simulated when using $\delta_{timeout}$, but this is an artifact of simulating a single day, since the first and last samples are necessary to correctly reconstruct the demand curve for a given period.

IV. LESSONS LEARNED AND FUTURE WORK

The work presented in this paper has demonstrated the efficacy of an alternative method of electricity metering based on *informative events*. In particular, Section III offers a systematic and in-depth assessment of the EDM procedure when enhanced with four different functional modifications. In particular, as shown in Section III-A, it is feasible to adjust the power variation threshold from each house in an automated manner so that the average number of measurements per day follows the desired target percentage. In Sections III-B and III-C, performance evaluation reveals that, even with a low number of measurements, there is still a noteworthy degree of redundancy residual in the data, which can be reduced by means of filtering. Applying the consecutive-measurements filter results in considerable reduction in the aggregate data size while having almost no impact on the quality of the signal. Enabling the spike filter yields higher compression in the data; however, in the specific case simulated, it has shown to be too aggressive, giving rise to much higher signal degradation than in the former case. This effect could be mitigated either by decreasing the power variation threshold value to capture the lower power thresholds, or by allowing a *non-strict* mode, which would skip the spikes but would register the following steady-state measurement. In Section III-D, we have shown that replacing the fixed time measurements with a timeout limit, significantly reduces the amount of simultaneous measurements while having an almost negligible impact in the accuracy of the measurements for billing purposes. In addition, the underlying communication infrastructure becomes less susceptible to congestion and relevant capacity-overload issues. The measurements can be slightly decreased as well.

An interesting aspect of the timeout parameter is the informative nature of not receiving a new sample. In this case, the time limit imposed by the $\delta_{timeout}$ parameter is used to infer the state of the network in a given neighborhood. If a house has not transmitted any measurement for a time period longer than $\delta_{timeout}$, a fault may have occurred either in the grid or in the communication system. Fig. 14 depicts a representative example for the case of missing measurements. Random errors in individual houses may eventually occur (e.g., due to interference), which are represented by small bars (the left side of the bars indicates when the timeout limit was exceeded while the right side denotes that normal communication was restored) spread all over the plot. On the other hand, groups of houses exceeding the timeout may indicate that a power or communications outage in the neighborhood area has happened. A power outage would look like the top-left part of the plot. After the interruption has occurred, an aggregator node would observe that several houses start exceeding the $\delta_{timeout}$

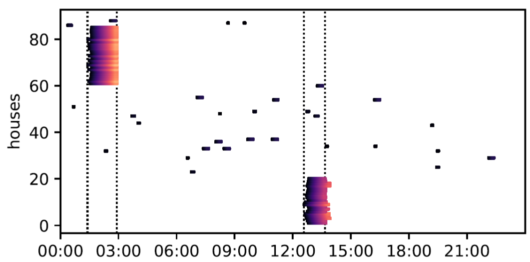


Fig. 14. Transmission errors under EDM with timeout. Top-left: power interruption on a neighborhood. Bottom-center: communication error in the neighborhood's aggregator. The remaining points correspond to single transmission errors. Vertical lines indicate the intervals in which the error events occur for the groups of houses. Source [1].

limit one after another; when the power is restored, all the affected houses would resume operation. The communications failure (i.e., in the bottom part) looks similar to the previous case, with the difference that, upon service restoration, the arrivals of the measurements are not synchronized assuming that no retransmission policy is implemented.

Although the EDM enhancements presented in this paper demonstrate overall good performance, some of the choices made are purely arbitrary, such as determining that the target number of measurements would be 1% of the measurements within a day. Such a choice was mainly motivated in an effort to show that it is indeed possible to have good signal reconstruction with an approximately constant number of measurements; this is in stark contrast with the conventional EDM approach where the signal error would have been constrained, regardless of the final number of measurements. Although the error metrics in conventional EDM demonstrate good results, the lack of flexibility in parameter determination (which is also considered fixed) and the unpredictability of the end result (regarding the number of measurements), make it difficult to accurately estimate the required infrastructure needs for a large deployment of meters. Future work will meticulously address the estimation of EDM parameters, aiming to adjust the power triggers in a more precise way. A possible way in this direction would be to use the event-driven measurements as a pre-processor for non-intrusive load monitoring (NILM) techniques, identifying the appliances by their power events, and, based on historical data, adjust the parameters according to the expected use patterns as soon as they are detected.

An additional research line refers to the joint consideration of EDM with missing data imputation techniques in an effort to regain perspective on the mechanisms which causally induce occlusions in measurement trajectories. Incomplete/missing data in metering applications can be typically attributed to (i) communication impairments, e.g., during transmission over an unreliable wireless link, (ii) sensor hardware failures, or (iii) security attacks, e.g., denial-of-service. Missing values often result in severe data quality degradation, which in turn harms the fidelity of energy consumption estimation and hinders informed system control. Since the foundational principles of EDM aim at reducing data redundancies, which could be instead used to compensate for the missing values, resorting to advanced imputation techniques becomes imperative [27]. Such techniques focus on the exploitation of intrinsic temporal and spatial cross-correlations among ambient measurement streams to extract their governing dynamics and curtail the inadvertent loss of measurements' quality. In the path forward, we plan to explore the potential of EDM in the context of incomplete datasets where imputation mechanisms are jointly employed to make valid inferences for the missing data and improve the signal reconstruction accuracy.

Acknowledgment

This work is based on the doctoral dissertation of Mauricio de Castro Tomé [1].

References

[1] M. de Castro Tomé, "Advanced electricity metering based on event-driven approaches," Ph.D. dissertation, Dept. Inf. Technol. Elect. Eng., Commun. Eng., CWC-Netw. Syst., Univ. Oulu, Oulu, Finland, 2021.

[2] B. Dudley *et al.*, "Bp statistical review of world energy," BP Stat. Rev., London, U.K., Tech. Rep. 6, 2018, pp. 1–64.

[3] R. R. Mohassel, A. Fung, F. Mohammadi, and K. Raahemifar, "A survey on advanced metering infrastructure," *Int. J. Elect. Power Energy Syst.*, vol. 63, pp. 473–484, Dec. 2014.

[4] M. Jachimski, Z. Mikos, G. Wrobel, G. Hayduk, and P. Kwasnowski, "Event-based and time-triggered energy consumption data acquisition in building automation," in *Proc. Int. Conf. Event-Based Control, Commun., Signal Process. (EBCCSP)*, Jun. 2015, pp. 1–6.

[5] J. Lunden and S. Werner, "Real-time smart metering with reduced communication and bounded error," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Nov. 2014, pp. 326–331.

[6] S. Werner and J. Lunden, "Event-triggered real-time metering in smart grids," in *Proc. 23rd Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2015, pp. 2701–2705.

[7] G. C. Madueno, J. J. Nielsen, D. M. Kim, N. K. Pratas, Č. Stefanović, and P. Popovski, "Assessment of LTE wireless access for monitoring of energy distribution in the smart grid," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 675–688, Mar. 2016.

[8] Y. Tsividis, "Event-driven data acquisition and digital signal processing—A tutorial," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 57, no. 8, pp. 577–581, Aug. 2010.

[9] B. Zhao, K. He, L. Stankovic, and V. Stankovic, "Improving event-based non-intrusive load monitoring using graph signal processing," *IEEE Access*, vol. 6, pp. 53944–53959, 2018.

[10] S. Roheda, H. Krim, Z.-Q. Luo, and T. Wu, "Decision level fusion: An event driven approach," in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 2598–2602.

[11] Y. Tang, "Self-triggered control of uncertain networked systems under packet dropouts," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Mar. 2016, pp. 2138–2142.

[12] J. Xu, D. W. C. Ho, F. Li, W. Yang, and Y. Tang, "Event-triggered risk-sensitive state estimation for hidden Markov models," *IEEE Trans. Autom. Control*, vol. 64, no. 10, pp. 4276–4283, Oct. 2019.

[13] J. Wu, Q.-S. Jia, K. H. Johansson, and L. Shi, "Event-based sensor data scheduling: Trade-off between communication rate and estimation quality," *IEEE Trans. Autom. Control*, vol. 58, no. 4, pp. 1041–1046, Apr. 2013.

[14] M. Miskowicz, *Event-Based Control and Signal Processing*. Boca Raton, FL, USA: CRC Press, 2015.

[15] F. Wang, Z. Wang, J. Liang, and X. Liu, "Event-triggered recursive filtering for shift-varying linear repetitive processes," *IEEE Trans. Cybern.*, vol. 50, no. 4, pp. 1761–1770, Apr. 2020.

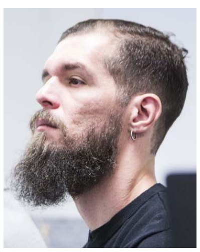
[16] M. Simonov, "Event-driven communication in smart grid," *IEEE Commun. Lett.*, vol. 17, no. 6, pp. 1061–1064, Jun. 2013.

[17] M. Simonov, "Hybrid scheme of electricity metering in smart grid," *IEEE Syst. J.*, vol. 8, no. 2, pp. 422–429, Jun. 2014.

[18] M. Simonov, H. Li, and G. Chicco, "Gathering process data in low-voltage systems by enhanced event-driven metering," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1755–1766, Sep. 2017.

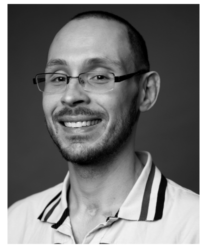
[19] M. D. C. Tome, P. H. J. Nardelli, and H. Alves, "Event-based electricity metering: An autonomous method to determine transmission thresholds," in *Proc. IEEE 87th Veh. Technol. Conf. (VTC Spring)*, Jun. 2018, pp. 1–5.

[20] M. de Castro Tomé, P. H. J. Nardelli, and H. Alves, "Long-range low-power wireless networks and sampling strategies in electricity metering," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1629–1637, Feb. 2019.
[21] M. de Castro Tome, P. Nardelli, and L. C. Pereira da Silva, "Flexible event-driven measurement technique for electricity metering with filtering," in *Proc. IEEE 17th Int. Conf. Ind. Informat. (INDIN)*, Jul. 2019, pp. 1071–1076.
[22] M. Simonov, G. Chicco, and G. Zanetto, "Event-driven energy metering: Principles and applications," *IEEE Trans. Ind. Appl.*, vol. 53, no. 4, pp. 3217–3227, Jul./Aug. 2017.
[23] M. Simonov, G. Chicco, and G. Zanetto, "Real-time event-based energy metering," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 2813–2823, Dec. 2017.
[24] A. Ozadowicz and J. Grela, "An event-driven building energy management system enabling active demand side management," in *Proc. 2nd Int. Conf. Event-Based Control, Commun., Signal Process. (EBCCSP)*, Jun. 2016, pp. 1–8.
[25] S. Mian Qaisar and F. Alsharif, "Signal piloted processing of the smart meter data for effective appliances recognition," *J. Electr. Eng. Technol.*, vol. 15, no. 5, pp. 2279–2285, Sep. 2020.
[26] A. Di Giorgio and L. Pimpinella, "An event driven smart home controller enabling consumer economic saving and automated demand side management," *Appl. Energy*, vol. 96, pp. 92–103, Aug. 2012.
[27] C. Kalalas and J. Alonso-Zarate, "Sensor data reconstruction in industrial environments with cellular connectivity," in *Proc. IEEE 31st Annu. Int. Symp. Pers., Indoor Mobile Radio Commun.*, Aug. 2020, pp. 1–6.



Mauricio de Castro Tomé received the B.S. and M.Sc. degrees in electrical engineering from the State University of Campinas, Brazil, in 2011 and 2014, respectively, and the Ph.D. (Hons.) degree from the University of Oulu, Finland, in 2021, researching power and energy measurements and load estimation in smart grids.

Daniel Gutierrez-Rojas received the B.Sc. degree in electrical engineering from the University of Antioquia, Colombia, in 2016, and the M.Sc. degree in protection of power systems from the University of São Paulo, Brazil, in 2017. He is currently pursuing the Ph.D. degree with LUT University, Finland. From 2017 to 2019, he worked as a Security of Operation and a Fault Analyst for Colombia's National Electrical Operator. His research interests include predictive maintenance and communications and electrical protection systems.



Pedro H. J. Nardelli (Senior Member, IEEE) received the B.S. and M.Sc. degrees in electrical engineering from the State University of Campinas, Brazil, in 2006 and 2008, respectively, and the Ph.D. degree from the University of Oulu, Finland, and State University of Campinas, following a dual degree agreement, in 2013. He is currently an Associate Professor (tenure track) with LUT University, Finland, and an Academy of Finland Research Fellow. His research focuses on wireless communications and energy systems.

Charalampos Kalalas (Member, IEEE) received the B.S. degree in electrical and computer engineering from the National Technical University of Athens in 2011, the M.Sc. degree in wireless systems from the KTH Royal Institute of Technology in 2014, and the Ph.D. *(cum laude)* degree from the Universitat Politècnica de Catalunya in 2018. He is a Researcher with the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA). His research interests include wireless systems with applications to cyber-physical and networked control systems.

Luiz Carlos Pereira da Silva received the Graduate degree in electrical engineering from the Federal University of Goias, Brazil, in 1995, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Campinas, Brazil, in 1997 and 2001, respectively. From 1999 to 2000, he was a Visiting Ph.D. student with the University of Alberta, Canada. He is currently an Associate Professor with the University of Campinas. His research interests are power system transmission and distribution.

Ari Pouttu is a Professor with the Centre of Wireless Communications (CWC), University of Oulu (UO), and is leading a research group targeting dependable wireless solutions, including solutions for 5G and is acting as a collaborative projects leader in national 6Genesis flagship targeting 6G solutions. He has published more than 50 journals or conference papers in the field of wireless communications and holds two patents.