Flexible event-driven measurement technique for electricity metering with filtering

de Castro Tomé Mauricio, Nardelli Pedro, da Silva Luiz Carlos Pereira

Please cite the publication as follows:

M. d. Castro Tomé, P. Nardelli and L. C. Pereira da Silva, "Flexible event-driven measurement technique for electricity metering with filtering," 2019 IEEE 17th International Conference on Industrial Informatics (INDIN), Helsinki, Finland, 2019, pp. 1071-1076. doi: 10.1109/INDIN41052.2019.8972086

Flexible event-driven measurement technique for electricity metering with filtering

Mauricio de Castro Tomé[1], Pedro Nardelli[1,2], Luiz Carlos Pereira da Silva[3]

[1] Centre for Wireless Communications, University of Oulu, Finland

[2] School of Energy Systems, LUT University, Finland

[3] School of Electrical and Computer Engineering, State University of Campinas, Brazil

Abstract—**Non-uniform, event-driven sampling of signals can be advantageous for different applications. In this paper, we focus on event-based sampling strategy for electricity metering purposes. Specifically, we propose an improvement in the enhanced event-driven metering (EDM) technique introduced by Simonov *et al*. Our solution provides additional flexibility on the types of measurements to be sent, by including the option to reduce the sending of consecutive measurements. Numerical results are presented for 4 different open databases of electricity consumption and consistently show that, in relation to the other options, our proposed strategy leads to both: (i) reduction in the amount of measurements sent, and (ii) improvements on the signal reconstruction by decreasing its reconstruction error. These two aspects are extremely useful in a scenario of massive deployment of measurement devices.**

Index Terms—**Event-based measurement, smart grids, electricity metering.**

I. Introduction

The event-based approach in signal processing, communications and control is an established research field that provides advantages to the widespread time-based approaches for specific classes of signals [1]. The core idea behind the concept is clearly stated in [2]: *(...) event-based sampling, in which a sample is produced only when something significant (an "event") occurs in the signal.* In this sense, the sampling is non-uniform, in contrast to the usual periodic sampling; however it is also not random. Events are related to variations; for instance, the signal is sampled whenever it crosses a predetermined threshold. Similarly, it can be sampled when it varies too fast by using a threshold in the derivative of the signal; or it can be also sampled when the cumulative value crosses a threshold (integral of the signal).

This kind of quantization has advantages not only in signal processing as shown in [3], [4] but also in communications [1], [5], [6] and process control [7]–[9]. Detection of spikes, higher-order variation or the cumulative values are important for different classes of applications. They range from cardiac monitoring [10] to multi-agent Internet of Things (IoT) enabled systems [11]. One example, which is the focus of this paper, is the application of event-based techniques in modern energy systems (the so-called smart grid) as the "enhanced event-driven metering", proposed and developed by Simonov in [12]–[14]. This approach is itself a derivation of the *change and transmit* strategy proposed by Li *et al* [15], where the information about the measure of interest is sent only when a given variation threshold is exceeded. If an adequate value

is chosen (for example, above the noise level), high levels of compression are achievable because the information below the threshold is discarded instead of being sent, with low impact on the reproduction of the signal.

This is especially useful for signals with slow variation taken from several remote measurement devices (such as electricity meters). The potential savings in communications infrastructure, bandwidth and storage can be huge, while adding little complexity to the measurement devices [16]–[18]. As those contributions have shown, electricity measurement can benefit greatly from this technique since is mostly constant over time, except for short (but usually sharp) variations.

The compression gains allowed by event-based techniques in signal processing have a positive impact in the so-called massive machine-type communications (mMTC) as part of IoT and 5G wireless systems [19]. For instance, in [17], the authors studied how event-based metering could be implemented in Low-Power Wide-Area Networks (LPWAN) communication systems. Those results have shown the advantage of the event-based technique, while also studying the impact of simulated packet reception outages (communication errors) in the reconstruction error of the signal. However, in that paper and in the original EDM proposal, there was no heuristics to look for solutions that compare the performance of the different techniques considering a similar amount of data points. The authors studied this in [18], demonstrating a consistent advantage of EDM.

Although these existing results are relevant, the EDM can still be improved by, for example, discarding redundant samples before they are sent to the receptor. Our results show how the proposed scheme keeps the amount of measurements close to the desired value, and the filtering of redundant measurements further increases the compression, at the expense of a slightly increased error. On the other hand, this increased error is still better than the equivalent time-based reference values.

The remainder of the paper is organized as follows: Section II presents a brief explanation of the event-driven metering technique, as well as our proposed modifications, Section III gives an overview of the datasets used in this study. Section IV shows simulations and compare the results of the proposed modifications against both the original technique and reference values from time-based measurements for selected houses. Section V summarize the results and shows improvements that can be made in subsequent studies.

II. Event-based metering technique

In this section, we will first review the original work in event-driven electricity metering done by Simonov *et al.* We will also present our proposed extension, while discussing how each one of the options work.

A. Simonov's contribution

Let us first revisit the parameters utilized by Simonov [12]–[14] about the event-driven metering (EDM). The proposed EDM implementation is composed of three criteria (*events*):

1) A sudden variation (increase or decrease) in power demand, which is bigger than a set limit ΔP, in Watts;
2) A cumulative power variation over time, denoted here as ΔP_{slow}, which can be set independently of ΔP;
3) A time-based measurement, to be sent every time interval T (starting from 00:00:00 of a given day), independent of the other two measurements (T_{fixed}).

The first event is straightforward: every time the instant power variation goes over a given threshold ΔP, a measurement (sample) is acquired and sent, which is useful for tracking large changes in power. The ΔP_{slow} parameter monitors the power variation from the present moment in relation to the last measurement sent. In this way, the authors dealt with the small cumulative power variations (such as in fuzzy-controlled loads, or the switching of small electric devices over time) that, when combined, might exceed the ΔP threshold,. Lastly, the time-based measurement is used to keep backwards compatibility with legacy metering and billing.

B. Proposed extension

In this section, we describe how we extend Simonov's EDM and our rationale behind it. Specifically, we implemented the EDM algorithm as described in [12]–[14] with the following modifications:

1) We included a *timeout* parameter in addition to the original fixed-interval timed measurements;
2) We implemented a filter to prevent sending consecutive measurements which both exceed the ΔP threshold;
3) We also included a filter for removing "spikes" from the sent measurements.

The *timeout* parameter guarantees that we have at least one measurement every T units of time after the last measurement. This helps tracking periods of extremely low power consumption that might not trigger ΔP or ΔP_{slow}. The timeout parameter also helps with detecting problems: if no measure arrives after the time T, it means that there was either a communication error, meter fault, or energy interruption.

The *consecutive measurements* filter aims to reduce the amount of measurements to be sent. The idea is to deal with consecutive measurements that point in the same direction (upwards or downwards). In almost every case, two consecutive measurements that both exceed the ΔP threshold belong to the same event (for example, the switching of a high-powered appliance such as a heater). This is also due to the effect of the sampling frequency, since the power measurements in use in this study are taken every second, while the appliances usually aren't switched exactly after the beginning of the sampling interval. Figure 1 shows its behavior.

The *spike* filter deals with sudden power variations that occur briefly and then return to their steady-state value (e.g., the starting of a fridge's compressor). This filter reduces the amount of measurements sent while having slow impact on the reconstruction quality – since the spike, although having a high value, is so short that it does not consume a significant amount of energy – and also prevents one possible situation in which the ΔP threshold is set to a value high enough that the positive part of the spike triggers an event (and sends an update in the demand to the utility), while its negative part stays below the detection threshold and is not registered, causing a considerable error in the power estimation. Figure 2 shows the result of said filtering.
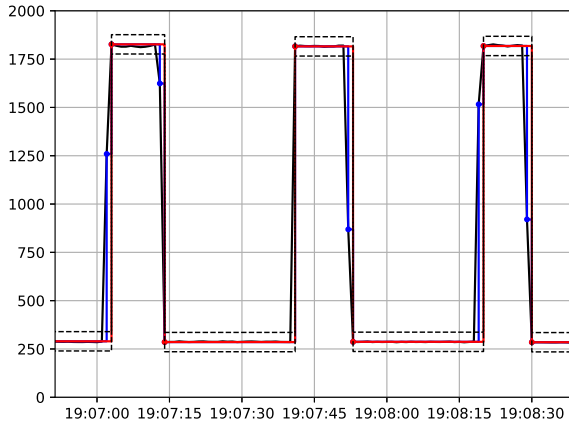

Fig. 1. Consecutive measurements filter. It filters out consecutive measurements before sending. In this time window, the number of measurements went from 11 to 6, with almost negligible impact on the signal reconstruction.
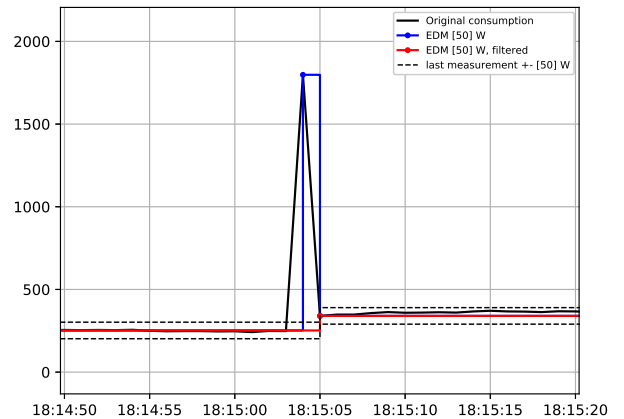

Fig. 2. Example of spike caused by a fridge's compressor, and result of the filtering being applied.

TABLE I
OVERVIEW OF THE DATASETS

DB/house		No. Days	No. empty	avg. meas	P_{min} (kW)	P_{max} (kW)	P_{avg} (kW)
	1	643	0	86196	0.002	11.564	0.371
UKDALE	2	110	0	85540	0.006	6.747	0.337
	5	137	0	85182	0.257	8.554	0.73
	A	92	0	85471	0.256	10.792	1.368
UMASS	B	81	0	86400	0.138	10.406	0.684
	C	81	0	86398	0	21.72	0.685
RAE	1	62	0	86343	0.225	17.206	1.104
	2	59	0	85020	0.001	7.753	0.344
	1	36	14	67898	0.091	12.333	0.384
	2	35	19	70502	0.032	3.253	0.23
REDD	3	45	20	54896	0.001	8.06	0.446
	4	48	23	64609	0	4.105	0.351
	5	44	36	33569	0.157	12.129	0.623
	6	24	3	40339	0.157	11.159	0.476

III. OVERVIEW OF DATABASES

In this paper, we analyzed residential consumption data from several freely available databases, namely:

- UK Domestic Appliance-Level Electricity dataset (UK-DALE) [20]
- UMass Smart* Data Set for Sustainability [21]
- The Rainforest Automation Energy Dataset (RAE) [22]
- Reference Energy Disaggregation Dataset (REDD) [23]

The use of different houses/databases is desirable to test the performance of the proposed method since each house has a unique power profile due to different appliances, number of occupants, and locations (the RAE database is from Canada, UK-DALE is from Great Britain, and both REDD and Smart* datasets are from USA). In this way, we believe we can show the generality and consistency of our contribution.

Table I presents an overview of the houses in the databases. We can see that some houses have lots of missing measurements, and completely missing days. This is especially true for the REDD houses 5 and 6.
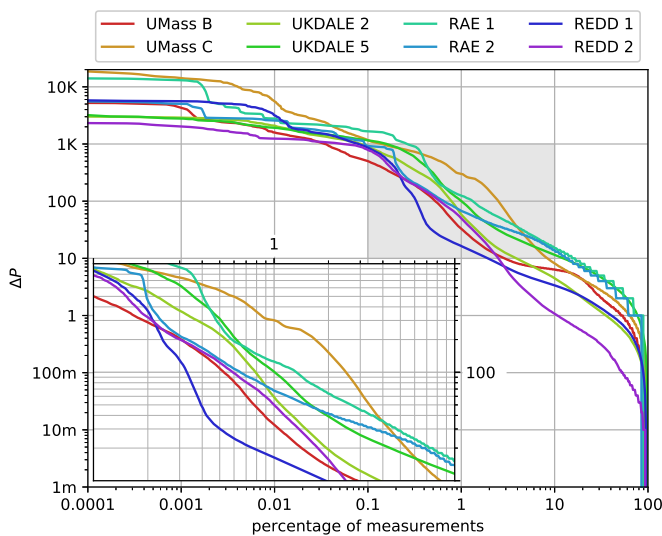

Fig. 3. Ordered ΔP percentage (main plot), and detail from $[0.1, 10] * [1, 1000](\%, W)$ (lower-left corner).

IV. CONFIGURATION AND SIMULATIONS

Before the simulations are run, we need to define some parameters, namely, the ΔP and timeout parameters, which give us the baseline for the number of measurements.

A. ΔP threshold estimation

To estimate the parameters used on the EDM strategy, some data manipulation is needed. Remembering Section II, the main purpose of the EDM measurements is the tracking of relevant instant power variations (ΔP). To estimate this parameter, we employed the following procedure:

1) Extract a week's worth of data from each house;
2) Calculate the *absolute* instant power variation $|\Delta P|$;
3) Order the resulting data by decreasing ΔP value.

Note that we have proceeded similarly for all houses but some were discarded from a few plots: houses 3 to 6 from REDD due to lack of measurements, house A from UMass due to high noise levels, and house 1 from UK-DALE (arbitrarily chosen so as not to overcrowd the plot). The resulting plot, Figure 3, show us the percentage of measures *smaller* or *equal* than a given value. We can see that most of the power variations are in the range of just a few Watts; for basically all of the houses, about 90% of the instant power variations are in the range of 10W or less, while values greater than 100W are in the range of 3% or less, which is a good indicator that the EDM strategy can achieve good compression. The lower-left corner of the figure gives us a more detailed view of the region $[0.1, 10] * [1, 1000](\%, W)$.

B. Timeout

Choosing an optimal value for the timeout is not straightforward, and also, having different values for each house is not practical. Hence, we decided to fix the value in $T=900s$ (15 minutes). Figure 4 shows an example of the inter-arrival rate for houses for $\Delta P=10W$. With this threshold, two of the houses would never hit the timeout threshold, while the rest of them have less 1% of its measurements exceeding that value.
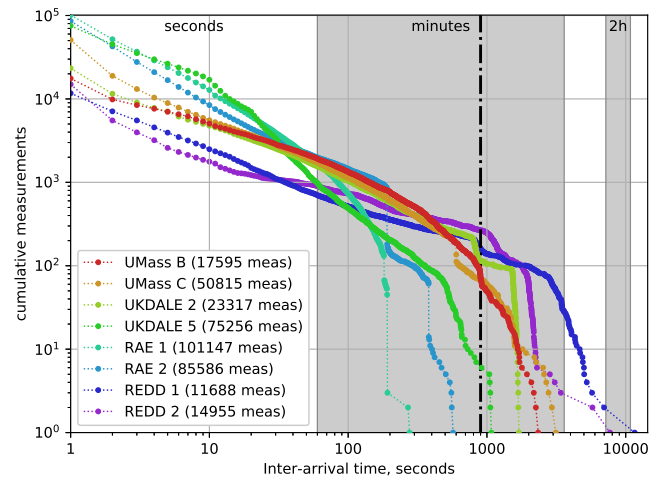

Fig. 4. Inter-arrival times for all houses, $\Delta P = 10W$. Vertical dashed line marks $T=900s$

C. Simulated Scenarios

We selected 2 of the houses and simulated 25 scenarios for each one of them, namely:

- Using only the threshold ΔP;
- Original EDM approach (that is, cumulative filter ON and $T_{fixed} = 900s$);
- A scenario using the cumulative filter, the spike + consecutive filter, no fixed time measurements (that is, $T_{fixed} \to \infty$), and the timeout set to 900s;
- Same as above, but also setting $T_{fixed} = 3600s$;
- A time-based signal with similar amount of measurements as the expected compression rate for a given ΔP, for performance comparison.

Each one of the configurations was repeated for 5 different values of instant power variation, which would result in measurements proportional to 0.1%, 0.3%, 1%, 3%, and 10% of the original data for that week. The results were condensed in tables with the following columns:

- Threshold/target: shows the ΔP parameter obtained from the previous week that would result in that percentage of measurements;
- Options: indicate which of the options in the EDM measurement are enabled in each one of the simulations, namely **(C)**umulative measurements, spike + consecutive **(F)**ilter, fixed **(T)**ime measurements, and **(t)**imeout. The symbol • indicates that the option is enabled, while ○ indicates that it is disabled;
- Final %: percentage in relation to the original measurements that was achieved by applying the parameters;
- Power and Energy MAE/RMS: We calculated the *normalized* values of both Mean Absolute and Root Mean Square errors to assess the quality of the signal.

The best results of each column for each one of the thresholds are marked **bold**, while values worse than the reference (fixed-time) values are marked in *italic*.

The simulations were conducted in the following way:

1) We collected measurements for a week;
2) We determined thresholds for ΔP that would result in the desired percentage for that week;
3) We applied these results on the *following* week's data.

In this way, we believe that we are closer to a real situation, collecting data and then analyzing it to determine the parameters to be applied for the next period, instead of relying on a "genie" to solve the problem.

For comparison, we also provided fixed time measurements with roughly the same amount of points, rounded to the nearest usual fraction of minute – that is, 15 minutes (900 seconds) for 0.1%, 5 minutes (300s) for 0.3%, 2 minutes (120s) for 1%, 30s for 3%, and 10s for 10%.

The "Power" and "Energy" errors are both shown because each one of them is calculated in a different way: the Power estimation is calculated based on the current power that is sent in the measurement (that is, the instant power value when the event happened), while the the Energy error value is the calculated by dividing the accumulated energy since the last measurement by the time between the samples. One could say that the Power error is the *forecasting* error, while the Energy error is the *smoothing* error.

D. UMass Smart* - house B

The parameters chosen for the simulations as well as the results are presented in Table II.

We can see that using the measurements from the previous week to determine the thresholds for the current week yields a number of measurements similar to the expected values, as shown in the "Final %" column. The baseline for the event-based strategy, marked as ○○○○, refers only to the instant ΔP values and is provided as reference to compare how the data collection week compares with the measuring week. Figure 5 shows the cumulative ΔP plot for the second week of the datasets, which resembles closely the results from the previous week, presented in Figure 3.

The original proposal for EDM (marked as •○•○ in the tables) has fixed time measurements every 900s and cumulative tracking of the power variation, and exceeds the amount of measurements in all the cases. This is expected, since we are adding a fixed amount of timed measurements to the baseline (96 points per day, or 672 points per week), as well as some extra measurements due to the ΔP_{slow} parameter. This effect gets less pronounced as the ΔP threshold is lowered. Applying our propositions (marked as ••x•) – that is, replacing $T_{fixed} = 900s$ for a timeout of 900s, and enabling or disabling the fixed time measurements – consistently outperforms all the other strategies with regards to compression, with the threshold targeting 0.1% of the data being the exception, due to the effect of the timeout/fixed-time measurements dominating in this case.

TABLE II
SIMULATION RESULTS FOR UMASS HOUSE B

| Threshold/ target	Options C F T t	Final %	Power MAE	Power RMS	Energy MAE	Energy RMS
	○ ○ ○ ○	**0.11**	*0.678*	*0.822*	*0.161*	*0.250*
501W	• ○ • ○	*0.25*	0.140	0.279	**0.059**	**0.129**
(0.1%)	• • ○ •	*0.19*	**0.121**	**0.229**	0.071	0.150
	• • • •	*0.20*	**0.121**	0.233	0.071	0.150
900s		0.11	0.242	0.743	0.186	0.499
	○ ○ ○ ○	**0.27**	*0.193*	0.273	0.086	0.132
204W	• ○ • ○	*0.44*	**0.059**	**0.106**	**0.034**	**0.082**
(0.3%)	• • ○ ○	**0.27**	0.088	0.152	0.045	0.115
	• • • •	0.28	0.083	0.147	0.046	0.115
300s		0.33	0.168	0.662	0.121	0.408
	○ ○ ○ ○	0.75	0.045	0.068	0.015	0.055
33W	• ○ • ○	*1.12*	**0.013**	**0.019**	**0.010**	**0.050**
(1%)	• • ○ •	**0.61**	0.019	0.088	0.016	0.111
	• • • •	0.62	0.018	0.088	0.016	0.110
120s		0.83	0.106	0.513	0.077	0.311
	○ ○ ○ ○	2.37	0.018	0.027	0.008	0.047
10W	• ○ • ○	*3.39*	**0.006**	**0.008**	**0.006**	**0.046**
(3%)	• • ○ •	**1.52**	0.011	0.108	0.014	0.119
	• • • •	1.54	0.011	0.108	0.014	0.118
30s		3.33	0.043	0.296	0.031	0.180
	○ ○ ○ ○	9.78	0.007	0.014	0.006	0.043
6W	• ○ • ○	*11.60*	**0.003**	**0.004**	**0.005**	**0.039**
(10%)	• • ○ •	**2.47**	0.010	0.119	*0.014*	0.114
	• • • •	2.50	0.010	0.119	*0.014*	0.114
10s		10.00	0.018	0.175	0.013	0.103
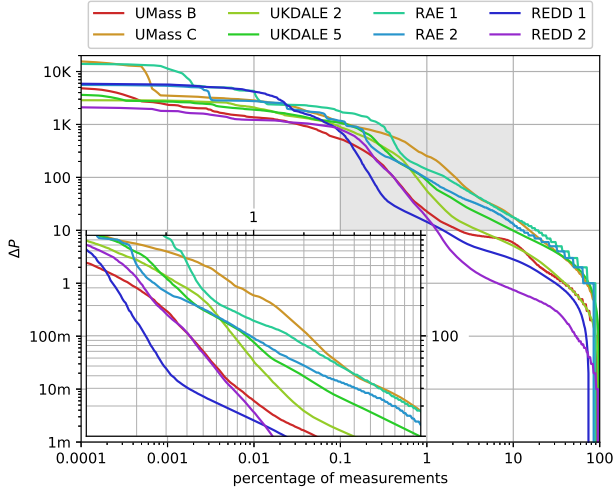
Fig. 5. Ordered ΔP percentage for week 2

Regarding the errors in reconstruction from P and E, the EDM strategies consistently outperform periodic sampling with similar amount of measurements, and sometimes even higher sampling rates. For example, the error values for the 1% target are comparable to the sampling rate of 10s, while using around 10 times less data. Among the different configurations, Simonov's approach usually has smaller error values, but again, that comes at the expense of more data points. Our modifications have slightly higher error rates, but achieve better compression in every scenario.

E. RAE Dataset, house 2

House 2 from the RAE dataset, shown in Table III, have slightly different results with regards to compression, with most of the EDM measurements going over the target amount of points. This is most likely due to increased activity levels in the current week in relation to the previous one, which was used to estimate the ΔP level for the power events. Nevertheless, our implementation managed to be pretty close to the target compression on average, and improving over it when the target measurement percentage increases.

V. CONCLUSION

In this paper, we suggested improvements upon the event-driven (EDM) technique. Our modifications achieve greater compression compared to both the original method and the fixed-time measurements, while having acceptable reconstruction error rates, slightly higher than the EDM strategy, but always smaller than timed measurements with similar number of points. This trade-off might be interesting in a scenario of massive deployment of measurement devices with limited battery, since our modifications result in further reduction in the amount of measurements.

ACKNOWLEDGEMENTS

TABLE III
SIMULATION RESULTS FOR RAE HOUSE 2

Threshold/ target	Options C F T t	Final %	Power MAE	Power RMS	Energy MAE	Energy RMS
923W (0.1%)	○ ○ ○ ○	*0.14*	0.387	0.714	0.225	0.379
	● ○ ● ○	*0.27*	**0.192**	**0.368**	**0.080**	**0.217**
	● ● ○ ○	*0.18*	0.236	0.427	0.149	0.301
	● ● ● ●	*0.19*	0.232	0.424	0.149	0.300
900s		0.11	0.453	1.55	0.346	1.058
217W (0.3%)	○ ○ ○ ○	*0.41*	0.231	0.326	0.160	0.265
	● ○ ● ○	*0.69*	**0.115**	**0.179**	**0.080**	**0.217**
	● ● ○ ○	*0.43*	0.154	0.238	0.106	0.234
	● ● ● ●	*0.44*	0.149	0.235	0.104	0.233
300s		0.33	0.295	1.286	0.236	0.857
68W (1%)	○ ○ ○ ○	*1.41*	0.077	0.145	0.048	0.197
	● ○ ● ○	*1.92*	**0.034**	**0.052**	**0.034**	**0.187**
	● ● ○ ●	**1.00**	0.045	0.134	0.039	0.204
	● ● ● ●	*1.02*	0.044	0.134	0.039	0.204
120s		0.83	0.233	1.13	0.167	0.696
33W (3%)	○ ○ ○ ○	*3.61*	0.042	0.075	0.031	0.180
	● ○ ● ○	*4.77*	**0.022**	**0.031**	**0.027**	**0.171**
	● ● ○ ○	**1.75**	0.036	0.189	0.036	0.220
	● ● ● ●	1.77	0.036	0.189	0.036	0.220
30s		3.33	0.095	0.611	0.069	0.381
14W (10%)	○ ○ ○ ○	9.76	0.024	0.038	0.023	0.145
	● ○ ● ○	*11.52*	**0.012**	**0.016**	**0.021**	**0.138**
	● ● ○ ○	**3.48**	0.032	0.276	0.037	0.237
	● ● ● ●	3.50	0.032	0.276	0.037	0.236
10s		10.00	0.054	0.392	0.039	0.241

REFERENCES

[1] M. Miskowicz, *Event-based control and signal processing*. CRC Press, 2015.

[2] Y. Tsividis, "Event-driven data acquisition and digital signal processing—a tutorial," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 57, no. 8, pp. 577–581, 2010.

[3] 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. 53 944–53 959, 2018.

[4] S. Roheda, H. Krim, Z.-Q. Luo, and T. Wu, "Decision level fusion: An event driven approach," in *2018 26th European Signal Processing Conference (EUSIPCO)*. IEEE, 2018, pp. 2598–2602.

[5] 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 Transactions on automatic control*, vol. 58, no. 4, pp. 1041–1046, 2013.

[6] F. Wang, Z. Wang, J. Liang, and X. Liu, "Event-triggered recursive filtering for shift-varying linear repetitive processes," *IEEE transactions on cybernetics*, 2018.

[7] Y. Tang, "Self-triggered control of uncertain networked systems under packet dropouts," in *2016 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2016, pp. 2138–2142.

[8] A. Goldenshluger and L. Mirkin, "On minimum-variance event-triggered control," *IEEE control systems letters*, vol. 1, no. 1, pp. 32–37, 2017.

[9] J. Xu, D. W. Ho, F. Li, W. Yang, and Y. Tang, "Event-triggered risk-sensitive state estimation for hidden markov models," *IEEE Transactions on Automatic Control*, 2019.

[10] S. M. Qaisar and A. Subasi, "An adaptive rate ecg acquisition and analysis for efficient diagnosis of the cardiovascular diseases," in *2018 IEEE 3rd International Conference on Signal and Image Processing (ICSIP)*. IEEE, 2018, pp. 177–181.

[11] S. Trimpe and D. Baumann, "Resource-aware iot control: Saving communication through predictive triggering," *IEEE Internet of Things Journal*, 2019.

[12] M. Simonov, "Event-driven communication in smart grid," *IEEE Communications Letters*, vol. 17, no. 6, pp. 1061–1064, 2013.

[13] ——, "Hybrid scheme of electricity metering in smart grid," *IEEE Systems Journal*, vol. 8, no. 2, pp. 422–429, 6 2014.

[14] M. Simonov, H. Li, and G. Chicco, "Gathering Process Data in Low-Voltage Systems by Enhanced Event-Driven Metering," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1755–1766, 9 2017.

[15] H. Li, S. Gong, L. Lai, Z. Han, R. C. Qiu, and D. Yang, "Efficient and secure wireless communications for advanced metering infrastructure in smart grids," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1540–1551, 2012.

[16] M. C. Tome, P. H. J. Nardelli, H. Alves, and M. Latva-aho, "Joint sampling-communication strategies for smart-meters to aggregator link as secondary users," in *2016 IEEE International Energy Conference (ENERGYCON)*. IEEE, 4 2016, pp. 1–6.

[17] M. de Castro Tomé, P. H. Nardelli, and H. Alves, "Long-range low-power wireless networks and sampling strategies in electricity metering," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 2, pp. 1629–1637, 2019.

[18] M. d. C. Tomé, P. Nardelli, H. Alves *et al.*, "Event-based electricity metering: An autonomous method to determine transmission thresholds," *arXiv preprint arXiv:1803.02103*, 2018.

[19] L. Thrybom and Á. Kapovits, "5G and Energy," 5G-PPP, Tech. Rep. September, 2015. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-White{_}Paper-on-Energy-Vertical-Sector.pdf

[20] J. Kelly and W. Knottenbelt, "The uk-dale dataset, domestic appliance-level electricity demand and whole-house demand from five uk homes," *Scientific data*, vol. 2, p. 150007, 2015.

[21] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht, "Smart*: An open data set and tools for enabling research in sustainable homes," *SustKDD, August*, vol. 111, no. 112, p. 108, 2012.

[22] S. Makonin, Z. J. Wang, and C. Tumpach, "Rae: The rainforest automation energy dataset for smart grid meter data analysis," *Data*, vol. 3, no. 1, p. 8, 2018.

[23] J. Z. Kolter and M. J. Johnson, "Redd: A public data set for energy disaggregation research," in *Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA*, vol. 25, 2011, pp. 59–62.