Prédiction de données différenciée pour les Smart Grids

Jad Nassar, Karen Miranda, Nicolas Gouvy, Nathalie Mitton

Prédiction de données différenciée pour les Smart Grids †

Jad Nassar[1,2] et Karen Miranda[3] et Nicolas Gouvy[1] et Nathalie Mitton[2]

[1]*HEI, une école d'ingénieur d'Yncréa Hauts-de-France, 13 Rue de Toul, Lille, France. Prénom.nom@yncrea.fr*
[2]*Inria, France. Prénom.nom@inria.fr,* [3]*UAM Lerma, Mexique. kmiranda@correo.ler.uam.mx*

La transformation des réseaux électriques existants en Smart Grids (SGs) ambitionne d'en faciliter l'automatisation pour une meilleure qualité de service tout en y facilitant l'intégration de sources d'énergies renouvelables. Cette évolution vers un réseau électrique plus intelligent nécessite de pouvoir transmettre en temps réel un maximum de données sur l'usage du réseau. Un réseau de capteurs sans fil (WSN) disséminés à travers le réseau électrique est une solution prometteuse vu les coûts réduits et la facilité du déploiement de tels réseaux. Ces avantages se heurtent avec les liens radios instables et les ressources limitées des WSNs. Afin de réduire la quantité de données envoyée sur le réseau, et donc de réduire la consommation énergétique, la prédiction des données est une solution efficace. Cette dernière consiste en une estimation des valeurs mesurées permettant de ne pas envoyer les données brutes lorsque l'estimation s'avère correcte. Ce papier présente un travail en cours qui consiste à utiliser les *time series estimation* avec l'algorithme *Least Mean Square* pour la prédiction des données dans un WSN appliqué aux SGs, tout en considérant les différents types de données et trafics des SGs. Les premiers résultats de simulation numérique présentent une meilleure prédiction de données tout en minimisant l'erreur quadratique moyenne en se comparant avec une solution de l'état de l'art.

Mots-clefs : Smart Grids, WSN, Time Series Estimation, Least Mean Square, Qualité de Service

1 Introduction

Wireless sensor networks controlling and exchanging messages is a promising solution in the shift from the traditional electric grid into a "smarter" one in which the integration of renewable energy resources, energy reduction and auto-adaptation are the main benefits. This is done by equipping the electric grid with wireless sensors located at strategic measuring points to achieve remote monitoring, data collection and control of the grid. In a SG, electricity and energy do exist, but connecting sensors to such high voltage with intermittent and ill-adapted energy levels is sometimes inappropriate. For that, battery-powered sensors must be deployed all over the grid alongside with the main-powered ones. Thus, reporting data measurements at specific intervals has a direct effect on the sensors battery lifetime since the communication task consumes most of their available energy. In such contexts of continuous data reporting, where data changes are limited, will cause redundant information at the destination. To mitigate these energy loss, data reduction strategies aim at reducing the amount of data sent by each sensor by predicting the measured values both at the source and the sink node using specific algorithms, which will require sending the predicted information only if it is shifted from the sensed one by a certain threshold. We focus on a time series forecasting technique, called Least Mean Squares (LMS), which is an adaptive algorithm with very low computational overhead and memory consumption, that despite its simplicity, provides satisfactory performances in terms of computational speed, robustness and precision [HSK96]. In this paper, we propose a modification for the LMS filter adaptation used for prediction in WSN which was introduced in [SR06], in order to adapt it to the different data types. We applied the algorithm to photo-voltaic cells monitoring data set. This is done by tuning the parameters of the filter by training it for a certain time with the real data values of every data set and choosing the values that minimizes the Mean Square Error (MSE).

†This work was partially funded by a grant from the MEL (Métropole Européenne de Lille), the SoMel SoConnected project (Ademe, PIA2), Yncréa Haut-de-France, Collège Doctoral Lille Nord de France and CPER Data. Part of this work was performed when Jad Nassar was a visiting student at the Metropolitan Autonomous University, Mexico.

2 LMS Algorithm Overview

A sensor node generates a data stream $u[n]$ which consists of the previous N readings, which will form the input of the filter. The predicted values at the output of the filter $y[n] = w^T[n]u[n]$, is a linear combination of the previous N samples of the data stream weighted by a weight vector $w[n]$; where $w(n) = [w_1, w_2, ..., w_M]^T$ and $u(n) = [u(n-1), u(n-2), ..., u(n-M)]^T$. We note that both w and u are of length M, where M is an integer corresponding to the memory of the filter (how many previous samples it will use). The error between the output and the desired signal $d[n]$ that the filter tries to adapt to is computed by : $e[n] = y[n] - d[n]$. This error is given as an input for the adaptation algorithm, which will update the weight coefficients at the next instant $n+1$ by : $w[n+1] = w[n] + \mu u[n]e[n]$, where μ is the step size parameter. The weight vector is modified at each step in order to minimize the MSE. With a simple modification for the LMS filter structure, the LMS algorithm can be used for prediction by delaying the input signal by one step, using it as a the reference desired signal $d[n]$. The filter computes the estimated value $\hat{u}$[n] of the input signal at time instance n, as a linear combination of the M previous readings. The step-size μ and the filter length M are two important parameters that need to be defined in order to ensure the convergence and robustness of the algorithm. The former will tune the convergence of the algorithm and the latter impacts directly on the computational load and memory consumption by considering more or less samples. A detailed explanation of the LMS filter can be found in [HW03]. The implementation of the LMS algorithm for data prediction in WSN is first presented in [SR06]. Here, identical filters are introduced at both the source and the sink referred as LMS-DPS (dual prediction scheme). The algorithm consists of three modes of operation : Initialization, normal and stand-alone mode. In the initialization mode the data samples are collected and reported to the sink without prediction. In this phase the step-size μ must be determined. Both the node and the sink compute the value of μ. It must satisfy the following condition[But95] : $0 \leq \mu \leq \frac{2}{\lambda_{max}}$; where λ_{max} is the greatest eigenvalue of the auto-correlation matrix $\mathbf{R}$. In normal mode, both the sink and the node use the last M samples to compute the prediction for the upcoming measurement, and update the filter coefficients. When the error drops below e_{max} for M consecutive iterations, the node switches to stand-alone mode. We note that the default start values for the filter weights are assumed to be zero. In the stand alone mode, the node still collects data and makes predictions, but as long as the error is below e_{max}, the filter is fed with the prediction $y[n]$ instead of the reading value $u[n]$, and the sink receiving no reading from the node assumes that the predicted readings are below the error threshold. If the error exceeds e_{max}, the filter switches back to normal mode and reports the readings.

3 Related Work

In literature, we may find extensive work on time series forecasting techniques for WSNs [MRR13] [BBJK17]. In [SR06], an implementation of LMS algorithm for prediction in WSN is presented. The LMS algorithm uses a dual prediction scheme by running the instance of the filter on both the sink and the node. In [SSD11], a variable step-size is proposed to improve the initial adaptation to the data by switching to a new step-size stable value after μ has sufficiently learned what kind of data the filter receives. Many other works have addressed the variable step-size of LMS [BCO16]. However, these proposals mostly require many adjustments of several parameters in order to optimize μ which is not suitable for a WSN with limited computation capabilities.

4 Problem Statement and Proposed Solution

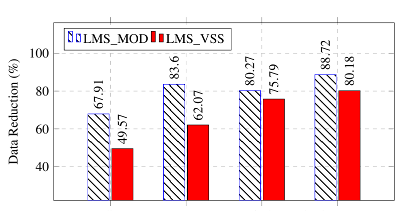
LMS adaptive algorithm is proved to be robust and accurate with a very low computation [HSK96]. However, the choice of the step-size and the filter length are essential in the convergence of the algorithm. Starting with a large step size gives a fast convergence of the filter but results in a larger MSE, and a too small step size degrades the capabilities of the algorithm. Varying the step size to a smaller value after a certain number of iterations is then beneficial. Concerning the filter length, the choice will indicate the computation load of the algorithm (how many samples we will consider on every iteration). We note that increasing the filter length does not necessarily improve the performance of the filter. Choosing the right parameters is then crucial. Many propositions to adapt and adjust these variables were proposed in literature, but having a direct mathematical analysis of the stability and steady-state performance is a very complicated task in LMS [HW03]. These adaptations may seem adequate for one application and kind of data-set, but less efficient to other ones. Our contribution, that we will denote by LMS_MOD (for modified) for the rest of the paper, consists of adding another step to the initialization phase for the LMS prediction algorithm in [SR06] by training the filter with enough data. We vary the step-size and filter length in order to optimize these values for every specific application. We start by the upper bound of μ ($\frac{2}{\lambda_{max}}$). In order to minimize the MSE, we compute : the optimal time to switch to a smaller value, the new value of μ, and the appropriate filter length M. After that, we run our prediction algorithm with these parameters for the rest of the data. Concerning the energy load resulting from this adaptation, we run our adaptation script offload using numerical simulations in order to obtain the coefficients before running it on a WSN.

5 Simulation Setup

In order to validate our proposition, we use real value traces from the solar technology acceleration center ‡ for voltaic cells. We considered the irradiance, current, average wind speed and air temperature between 01/07/2018 and 31/08/2018 with data collected every minute. A description of the traces characteristics is presented in Table 1. It is worth to mention that each data type has different characteristics and range which will make the prediction task more challenging. We considered a fixed threshold for each data type (can be adjusted for specific needs). We considered a one hop communication environment with one source and one destination and no loss in order to prove its efficiency in an optimal case scenario as a first step. We test our algorithm by means of numeric simulation on Matlab. For the adaptation in the initialization phase we execute a Matlab script for one day of collected data chosen randomly. We varied three parameters i, j and k corresponding to the filter length, the factor by which we will divide the old μ and after how many iterations simultaneously (the time we will switch to the new computed μ value). We vary i between 1 and 10, and j, k between 1 and 100 with a step of 5, and we chose the value that minimizes the MSE. We note that the choice of these intervals can be changed, but we realized after several tests that the optimal values fall within these ranges. The obtained values are shown in Table 1 are then fed for the filter in order to predict the data for the whole previously mentioned duration. We compare LMS_MOD to LMS_VSS proposed in [SSD11]. LMS_VSS respects the prediction phases as in [SR06] (Initialization, normal and stand-alone) but with a variable step-size like our proposition. μ starts with the value of μ_{old}=$2\lambda_{max}^{-1}.10^{-2}$, and switches to a stable value μ_{new}= μ_{old}/M after n iterations, where M is the filter length and n is the number of consecutive readings in stand-alone mode. They chose $n = M^{3/2}$. We note that we chose the same filter length used for LMS_MOD for every data set.

Data Type	Max Value	Min value	Std dev	Threshold	Filter Length	μ dividing factor	Nbr of Iter
Irradiance (w/m^2)	1431	1.57	350.07	5	5	31	11
Current (mA)	0.6	$2.5*10^{-5}$	0.22	10^{-3}	5	6	61
Avg Wind Speed (MPH)	43.28	0	4.47	2	2	11	86
Air Temperature $(°C)$	37.58	10.69	5.54	0.5	4	1	1

TABLE 1: Data Traces Description Parameters



(a) Data reduction percentage

Data Type	LMS_MOD	LMS_VSS
Irradiance	44.06	37.05
Current	0.031	0.043
Avg Wind Speed	1.71	2.01
Air Temp.	0.46	1.18

(b) RMSE for LMS_MOD and LMS_VSS

FIGURE 1: Performance study

‡. Andreas, A. ; Wilcox, S. ; (2011). Solar Technology Acceleration Center (SolarTAC) ; Aurora, Colorado (Data) ; NREL Report No. DA-5500-56491. http ://dx.doi.org/10.5439/1052224

6 Performance Evaluation

Figure 1b shows the Root Mean Square Error (RMSE) which corresponds to the root of the MSE that is given by : $MSE = \frac{1}{n}\sum_{n}^{i=1}(y[n]-u[n])^2$ for LMS_MOD and LMS_VSS. We observe that our proposal reduced the RMSE for the current, average wind speed and air temperature. The only value where LMS_VSS performs better is for the irradiance data set and this is probably due to the high deviation of the collected data as we observe from the high value of the standard deviation, 350.07, in Table 1. For that, training the filter for one day to estimate the parameters may not be enough to minimize the RMSE. On the other hand, as we see in Figure 1a, the percentage of data reduction in stand-alone mode where the error is below the given threshold) for the same data set for LMS_MOD is higher than the percentage for LMS_VSS. Same for the other data sets, LMS_MOD performs better than LMS_VSS.

7 Conclusion and Future Work

In this paper, we presented a work in progress that consists of a modification of the LMS prediction algorithm for WSN in order to adapt it to different applications with different characteristics, by training the filter for one day traces in order to optimize the parameters that minimize the MSE. We tested our approach with real value traces for photo-voltaic cells, and performed simulations considering one hop communication networks. Our first numerical results shows a better performance than LMS_VSS in terms of RMSE and percentage of data economy. As future work, we will continue investigating our approach with more tests on different data sets and considering more metrics. Later on, we will implement our algorithm on a sensor platform [§], to evaluate its performance in a real scenario with interference and losses.

Références

[ACDFP09] Giuseppe Anastasi, Marco Conti, Mario Di Francesco, and Andrea Passarella. Energy conservation in wireless sensor networks : A survey. Ad hoc networks, 7(3) :537–568, 2009.

[BBJK17] Siddhartha Bhandari, Neil Bergmann, Raja Jurdak, and Branislav Kusy. Time series data analysis of wireless sensor network measurements of temperature. Sensors, 17(6) :1221, 2017.

[BCO16] Dariusz Bismor, Krzysztof Czyz, and Zbigniew Ogonowski. Review and comparison of variable step-size lms algorithms. International Journal of Acoustics and Vibration, 21(1) :24–39, 2016.

[But95] Hans-Jürgen Butterweck. A steady-state analysis of the lms adaptive algorithm without use of the independence assumption. In International Conference on Acoustics, Speech, and Signal Processing (ICASSP), volume 2, pages 1404–1407. IEEE, 1995.

[HSK96] Babak Hassibi, Ali H Sayed, and Thomas Kailath. H-infinity optimality of the lms algorithm. IEEE Transactions on Signal Processing, 44(2) :267–280, 1996.

[HW03] Simon Haykin and Bernard Widrow. Least-mean-square adaptive filters, volume 31. John Wiley & Sons, 2003.

[MRR13] Karen Miranda, Tahiry Razafindralambo, and Victor Ramos. Using efficiently autoregressive estimation in wireless sensor networks. In International Conference on Computer, Information and Telecommunication Systems (CITS), pages 1–5. IEEE, 2013.

[SR06] Silvia Santini and Kay Romer. An adaptive strategy for quality-based data reduction in wireless sensor networks. In Proceedings of the 3rd international conference on networked sensing systems (INSS), pages 29–36, 2006.

[SSD11] Biljana Stojkoska, Dimitar Solev, and Danco Davcev. Data prediction in wsn using variable step size lms algorithm. In Proceedings of the 5th International Conference on Sensor Technologies and Applications, 2011.

§. https ://www.iot-lab.info/