LSTM for Periodic Broadcasting in Green IoT Applications over Energy Harvesting Enabled Wireless Networks: Case Study on ADAPCAST

Khiati Mustapha[1], Djamel Djenouri[2,*], Jianguo Ding[3], Youcef Djenouri[4]
[1] USTHB, Algiers, Algeria
[2*] CSRC, Dep. of Computer Science and Creative Technologies,
University of the West of England, Bristol, UK
[3] Blekinge Institute of Technology
[4] Dep. of Mathematics and Cybernetics, SINTEF Digital, Oslo, Norway
mkhiati@usthb.dz, djamel.djenouri@uwe.ac.uk, Jianguo.Ding@bth.se, youcef.djenouri@sintef.no
Correspandance author (*) address: UWE Bristol, Frenchay Campus, Coldharbour Ln, Bristol BS16 1QY, UK.

***Abstract*—The present paper considers emerging Internet of Things (IoT) applications and proposes a Long Short Term Memory (LSTM) based neural network for predicting the end of the broadcasting period under slotted CSMA (Carrier Sense Multiple Access) based MAC protocol and Energy Harvesting enabled Wireless Networks (EHWNs). The goal is to explore LSTM for minimizing the number of missed nodes and the number of broadcasting time-slots required to reach all the nodes under periodic broadcast operations. The proposed LSTM model predicts the end of the current broadcast period relying on the Root Mean Square Error (RMSE) values generated by its output, which (the RMSE) is used as an indicator for the divergence of the model. As a case study, we enhance our already developed broadcast policy, ADAPCAST by applying the proposed LSTM. This allows to dynamically adjust the end of the broadcast periods, instead of statically fixing it beforehand. An artificial data-set of the historical data is used to feed the proposed LSTM with information about the amounts of incoming, consumed, and effective energy per time-slot, and the radio activity besides the average number of missed nodes per frame. The obtained results prove the efficiency of the proposed LSTM model in terms of minimizing both the number of missed nodes and the number of time-slots required for completing broadcast operations.**

***Keywords*—IoT, wireless networks, energy harvesting, green computing.**

I. Introduction

Green IoT is an emerging concept that envision reducing energy consumption of IoT devices through innovative technologies and solutions for the purpose of achieving sustainable environments for futuristic applications. We abstractly consider in this paper a generic IoT system in which a central station, we call it a base station (BS), is connecting wireless devices (WD) to the Internet. Those devices can be sensor nodes, mobile devices, or any tiny devices featured by limited energy supply and using wirless communications. Considering this generic setting, a particular problem is dealt with in this paper; periodic broadcasting service that allows to disseminate messages from the BS to all WD in its vicinity. The problem resides in computing schedules with the minimum number of broadcast time-slots such that the BS can reach the maximum number of WDs with minimum retransmissions. To preserve energy, WDs are generally duty cycled and alternate between active/sleep modes according to their schedules. They tend to be at sleep mode for a much longer period and wake up only for short periods. Another emerging technique is the energy harvesting (EH) from the environment, in which WDs take advantage from the ambient energy sources (solar, wind, EM waves, etc.) to opportunistically collect energy. The amount of energy that can be collected by WDs varies from a technology to another, but is generally very limited and allows to charge only small capacitors. While EH is potentially promising to face the battery limitations, relying on EH is constrained and not yet effective as the incoming energy is not stable and largely vary over time and space. Duty cycling is then still required in EH environments for low power WDs. The environment considered in this paper consists thus in both EH and duty cycled, which elevates the complexity of the problem. A WD readiness for receiving broadcast messages is not only constrained by its schedules, but also by the availability of ambient energy. This makes optimal broadcasting in EHWNs extremely challenging. Given the variability of the environment, it is inevitable that the scheduling should be dynamic.

The big problem that is dealt with in this paper is to define a policy for deciding on re-schedule as a function of the environment changes, i.e., detect at the right moment that the current schedule becomes inefficient and should be updated. Broadcast protocols in wireless networks, including our previous solution (ADAPCAST) [1], aim at constructing schedules with a minimum number of time-slots and thus minimum latency. However, the schedules are formed for static long periods without updates. When time-slots dedicated for broadcasting becomes no more appropriate for most nodes, the schedule should be updated. Therefore, a generic mechanism that allows to plug such a dynamic rescheduling policy to broadcast protocols is required. The policy should adapt the length of the period to the changing parameters such as the effective energy and the random variable (r.v.) related to the incoming amount of energy. For this purpose, we explore machine learning (ML) approaches in this paper, particularly the advanced type of neural networks, the Long Short Term Memory (LSTM) [2]. LSTM allows to face many of the forecasting problems by analyzing sequences as time series where data can be defined as a chronological sequence of observations for a selected variable [2].

Motivated by all this, a generic LSTM is proposed for periodic broadcasting policies in EHWN that dynamically predicts the end of the broadcast period. In the present work, the MAC layer is supposed to be ruled by a slotted CSMA (Carrier Sense Multiple Access) mechanism [3], [4], which is used by most IoT enabled technology such as Zigbee and WiFi. The proposed LSTM allows for the BS to decide at the end of each frame about the end of the current period according to the Root Mean Square Error (RMSE) [5] value generated by the output layer. As proof of concept and without loss of generality, the proposed LSTM is applied to ADAPCAST. However, the proposed solution is generic and remains applicable to any time-slotted broadcast policy over CSMA that needs forecasting to decide about scheduling updates. The LSTM enforced ADAPCAST evaluated by simulation and the obtained results show a clear enhancement in terms of minimizing the number of missed nodes and the number of time-slots devoted to broadcasting during each frame. The reminder of this paper is organized as follows: Section II gives an overview about the area of research and the related work. The proposed solution is detailed in Section III, while Section IV presents the simulation tests. Finally, Section V draws the conclusion.

II. Related Work snd Background

A. LSTM for Energy Forecasting: short literature review

Many complex predictive models have been proposed including decision trees, K-means clustering [6], Bayesian inference [7], and advanced neural networks such as LSTM, which are appropriate for time series problems. Cheng et al. [8] proposed PowerLSTM, an LSTM-based power demand forecasting model for a smart grid system. It considers weather conditions besides time features. Selvin et al. [9] proposed a deep learning model that applies a sliding window approach for predicting future price values on a short term based on Nash-Sutcliffe Efficiency, then it lists listed companies and compares their performance based on percentage error. Many works considered EH for a particular of WN, the wireless sensor networks (WSN). Cui et al. in[10] considered solar EH enabled WSN and addressed uncertainty of solar energy. An LSTM neural network was proposed to predict solar energy for three days based on historical solar energy collection data and environmental data. Based on energy prediction results, a predictive task scheduling strategy is put forward to improve the performance of the WSN. Wang et al. [11] introduced the concept of multi-energy interaction characteristic of regional energy system. They proposed an energy prediction approach based on LSTM in which a multi-task learning model is used to achieve interaction among multi-energy system. LSTM model was also introduced by [12] for building electrical energy forecasting based on prediction. LSTM and ML based methods have also been used for other related applications such as the prediction of the locations, e.g., deep Fuzzy-LSTM in [13], the prediction of optimal deployment for wireless sensors [14], and many smart city related IoT applications such as energy management in smart buildings [15], urban traffic management [16], [17], pedestrian collective behavior analysis, [18], etc.

B. ADAPCAST

A short description of ADAPCAST is provided here as the protocol used later for the case study. More details are available in [1]. ADAPCAST is a centralized solar EH based broadcast policy proposed for WSN. It runs at the BS level and the node level. First, the BS maintains and updates the required information to run a Hidden Markovian Model (HMM) [19] corresponding to every sensor node (SN). The BS level runs in three phases. i) The first phase generates observation sequences by running the HMMs. ii) The second phase consists in applying an optimal set selection algorithm as a heuristic of the covering set problem to determine the optimal broadcast time-slots. Finally, iii) the third phase uses the Baum-Welch learning algorithm to adjust the initial HMMs of SN's to increase the likelihood of meeting the most appropriate sequences that produce the broadcast time-slots obtained in phase (ii). The BS communicates the derived HMM models to theirs corresponding nodes. This process is executed at the initialization of the network. It was supposed abstractly that the obtained scheme is used for a long period until a "significant change" on the HMMs parameters takes place. However, no details has been provided on how to infer or estimate such a change. This problem addressed in the current paper and an LSTM model is proposed for this purpose.

C. Long Short Term Memory (LSTM)

LSTM is is a special kind of recurrent neural network capable of handling long-term dependencies and mitigating the vanishing and exploding gradient problems faced by Recurrent Neural Networks(RNNs) over time [20]. The key component in LSTM models is the cell state. As shown in Fig.1, each cell is considered as a block. It integrates filters called gates to regulate the flow of information where activation functions, e.g., sigmoid and tanh are used. There are three kind of gates, input, forget, and output.

III. Proposed Solution

The LSTM model proposed herein aims at maximizing node coverage efficiency for broadcast operations in EHWNs, i.e., minimizing the number of missed nodes and the required number of timeslots. It applies to any broadcast method operating under time-slotted CSMA MAC protocol. Without loss of generality, we present the proposed model when applied to ADAPCAST [1] and show the enhancement that can be achieved from LSTM. The principal is to dynamically adapt to the weather conditions changes by predicting the end of the current period according to the obtained RMSE value [5],

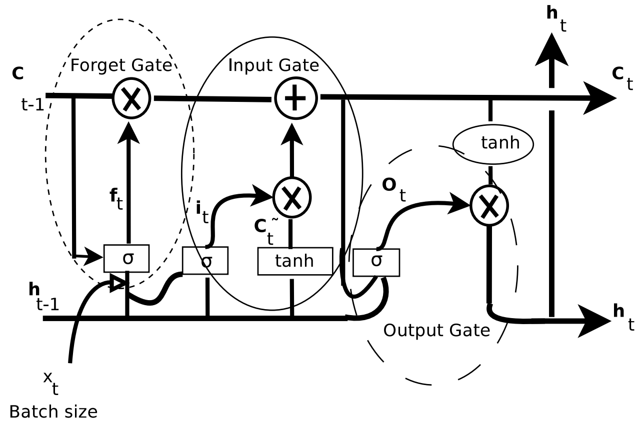

Fig. 1. LSTM cell block

which is the output of the proposed LSTM as illustrated in Fig 2. The obtained RMSE values represent the difference between the predicted values form the testing (validation) set and those from the training set of the proposed LSTM, with respect to the following features, the average number of timeslots required, the average number of missed nodes, the average amount of incoming energy, the average amount of effective energy, and the gap between the optimal (desired) observations and the generated ones with respect to the activity of the radio. The later is expressed as the Hamming distance per frame [21].

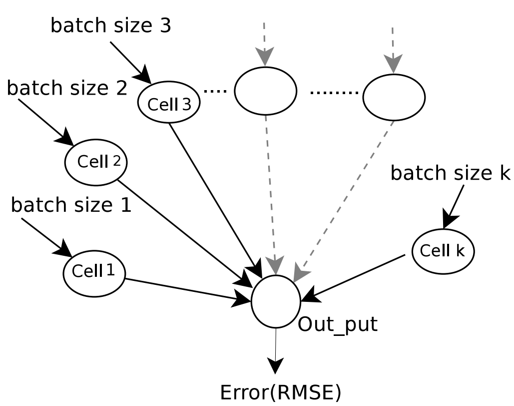

Fig. 2. Proposed LSTM model

Fig. 3 illustrates the LSTM-enhanced ADAPCAST. At the end of each time frame, the BS has to assess the need to ending the current broadcasting period. This is by measuring the gap between the predicted and the current values through RMSE and compare it with a given threshold, say δ. The latter represents the degree of sensitivity of the LSTM model, in which the gap is proportional to the number of missed nodes. Notice that the higher the gap is, the higher the number of missed nodes will be. The LSTM model is divided into three parts as presented in Fig.3, 1) input layer, 2) single hidden layer, 3) output layer. The input layer designates the dataset that is divided into equal-size subsets, while the hidden layer is trained based on the training set. Through the Adam optimizer, the parameters are optimized using the minimum loss value as the measurement principle. Finally, notice the output layer corresponds to the predictions made according to the learned RMSE.

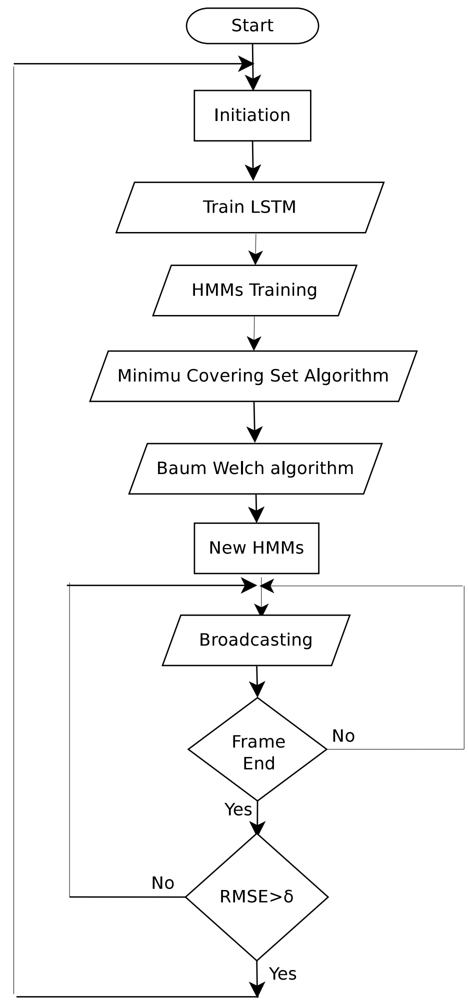

Fig. 3. LSTM based version of ADAPCAST

In order to avoid the rise in the number of missed WDs while waiting for the end of the predetermined period, the proposed LSTM allows to dynamically decide whether the current period has to be stopped or not (i.e., reversely starting a new period). This is by predicting the next RMSE values (with respect to the input vector, y) that is given by,

$$RMSE = \sqrt{\frac{1}{(n)}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}, \quad (1)$$

where $\hat{y}_i$ designates the predicted value, y_i is the real value, and N designates the number of observations of the variable y. Each vector, y_i, includes the following:

$$\begin{cases} y_{i,1}: & Average\ number\ of\ missed\ nodes \\ y_{i,2}: & Activity\ of\ the\ radio\ (Eq.2) \\ y_{i,3}: & Average.\ number\ of\ broadcast\ slots \\ y_{i,4}: & Average\ effective\ energy \\ y_{i,5}: & Average\ incoming\ energy \end{cases}$$

Let us consider a cycle of, n, time slots, and denote the activity by a vector a, i.e., $a_i = 1$ if the radio is active at time slot t_i and 0 otherwise. $y_{i,2}$ is given by the following formula:

$$y_{i,2} = \frac{1}{(n-b)} \sum_{i=1}^{n} (\acute{a}_i \bigoplus \hat{a}_i), \quad (2)$$

where $\hat{a}_i$ is the generated vector of activity, $\acute{a}$ is optimal one, $\bigoplus$ is the "exclusive or" (XOR) operator, and b is the number of time slots required to accomplish the broadcast.

The use of the proposed LSTM allows to dynamically determine the length of the periods. For this purpose, the data set is first generated by cumulating values related to the different parameters including the generated sequences, the effective energy, the radio activity, the number of missed WDs, and the incoming energy. This is by running the HMMs at the BS. The obtained dataset is then split into equal-size subsets that are introduced to the LSTM (Fig.3). The RSME is generated by the LSTM at the output layer. Small gap between the current and the previous RMSE values means that the model is still valid and can be kept, while a high value triggers alarm for rescheduling. Moreover, the gap between the previous and the current RMSE values determines the sensitivity of the proposed model where the changes in the inputs (especially the changes related to the weather conditions) directly affect this gap. We propose a bias denoted by (δ) that designates the degree of the sensitivity of the model. This is a very important parameter and should be set very carefully. The smaller δ is, the more sensitive the model becomes, which means rescheduling needs to be performed more frequently. The impact of this parameter will be investigated empirically in the next section.

IV. Simulation Study & Discussion

This section is devoted to the evaluation by simulation of the proposed LSTM model when applied to ADAPCAST. The proposed LSTM is implemented using python and a "Radean HD8600 M series" GPU. The Datasets used in the test were artificially generated, and each batch is composed of 1000 sets. Due to the limitation in terms of memory and computing resources, we limited the proposed LSTM to a single hidden layer (shallow LSTM). The input layer that is composed of 50 cells. The obtained results are compared with those of the original version of ADAPCAST. Table I summaries the simulation parameters.

Parameter	Value
Simulation duration(frames)	[50,250]
Unit of incoming energy E_r(mj)	100
θ	0.1,0.5,0.9
Threshold of effective energy E_{thr}(mj)	250
Time_Slot length (ms)	5
Frame length (Time_Slot)	20
environment	Python , Theano

TABLE I

Simulation parameters.

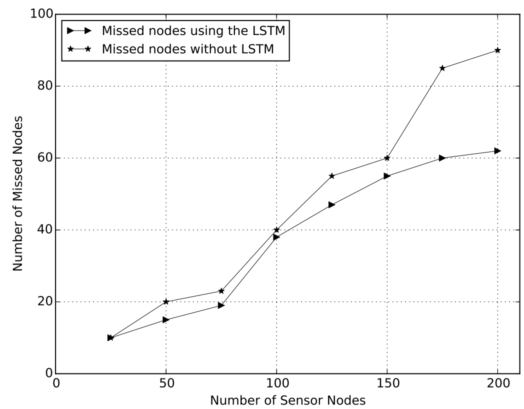

Fig. 4. Average number of missed nodes vs. number of nodes

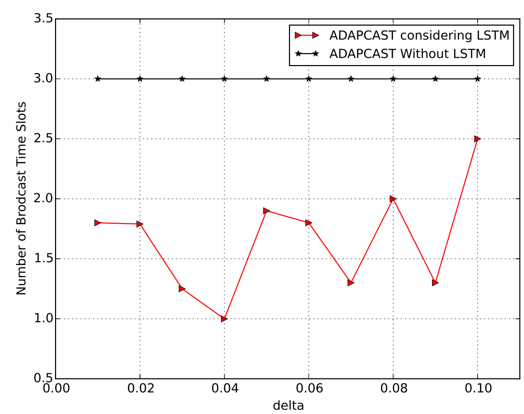

Fig. 5. Average Number of Broadcast Time-slots Vs. δ values

The purpose of the simulation tests is to investigates the number of missed nodes and the number of broadcasting time-slots required for maximizing the nodes coverage, as well as the impact of the model sensitivity parameter (δ). In ADAPCAST, the geographical locations and the weather condition make the gain in terms of incoming energy slightly different from a node to another. This is modled by the parameter θ. the latter is used with random values $\in$ [0.1, 0.9] in each scenario, and average values of the performance parameters are plotted. In a further step of the study, the performance of the proposed LSTM model under harsh weather conditions

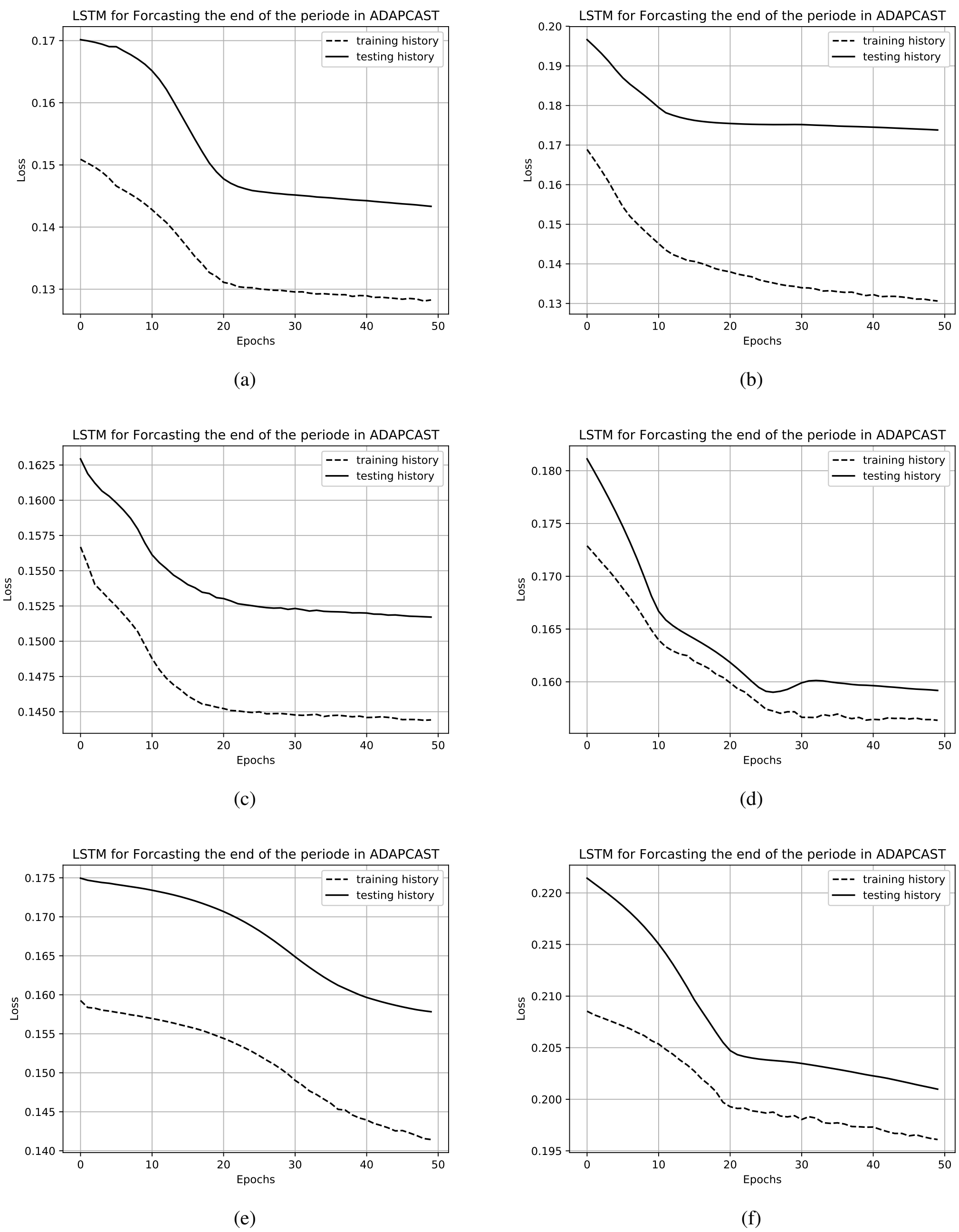

Fig. 6. Loss : (a): $\delta = 0.01$, (b) :$\delta = 0.02$, (c): $\delta = 0.05$, (d): $\delta = 0.06$, (e) $\delta = 0.09$, (f): $\delta = 0.1$

(i.e, rainy weather) is investigated, where θ is set to 0.1. Fig 4 shows how the proposed LSTM helps reducing the number of missed nodes (averaged per time frame). Notice the increasing gap between the two plots with the rise in the number of nodes. Ultimately, the LSTM allows to reduce this number from 90 to 60 for 200 nodes.

Fig. 5 plots the average number of the selected broadcast time-slots per time frame vs. δ. Here θ has been set to 0.1 (simulate raining condition) while the number of nodes has been set to 25 nodes. The figure shows clear improvement by the LSTM, which allows to define schedules that use low number of time-slots according to the sensitivity of the model expressed by δ. The enhanced ADAPCAST takes between 1 to 2 time slots on average, while the standard ADAPCAST uses 3 on average.

Fig.6 evaluates the performance of the proposed LSTM in terms of the loss function between the test set and the validation set of the entire historical data-set. This is composed of the results related the number of missed nodes per frame, the incoming energy, the effective energy, the number of broadcasting time-slots, and the gap between the optimal (desired) observations and the generated ones expressed in terms of Hamming distance [21]. The results reported in Fig.6 confirm that the training of the proposed LSTM model for different values of δ improves the efficiency of the learning model in terms of the accuracy in prediction. The results also confirm the predictive (testing) values are in line with those of the training set in all scenarios. However, note that overfitting is still possible in case of severe changes in the weather conditions causing random picks to Θ which was not investigated in this study.

V. Conclusion & Perspectives

This paper we have dealt with enhancing periodic slotted MAC (medium access control) based broadcast policies in EHWN (energy harvesting wireless networkss) and have explored the use of forecasting methods for optimizing the broadcast scheduling. An LSTM (long short term memory) model was been proposed for this purpose. The model allows to dynamically decide about the end of broadcasting period and the need of rescheduling in response to the variable changes in environmental conditions (e.g., weather). The model was been described when applied to our previous broadcast protocol (ADAPCAST) as a case study. However, it applies to any time-sloted broadcast policy that can take advantage of forecasting to decide about scheduling updates. The proposed model has been evaluated by simulation and its impact on enhancing ADAPCAST has been investigated. The obtained results prove the efficiency in terms of decreasing number of missed nodes and the number of time-slots that are required for broadcasting. This confirms that the efficient prediction of the optimal broadcast cycle period (time frame) reduces the impact of the divergence of the energy harvesting parameters (derived with the HMMs of the WDs in case of ADAPCAST). This ultimately enables reaching a maximum number of nodes with minimum retransmissions of the broadcast messages (broadcast count). As a perspective, we plan to design a multi-layer LSTM for more accuracy, and examine cases of overfitting in case of severe changes in the weather conditions causing random picks to Θ. These scenarios are not likely to be realistic, and if so extremely rarely, but evidence based analysis is needed. We are also exploring the use of realistic datasets to evaluate the proposed LSTM for this purpose.

References

[1] M. Khiati and D. Djenouri, "Adapcast: Adaptive learning-enforced broadcast policy for solar energy harvesting wireless sensor networks," *Computer Networks*, vol. 143, doi:10.1016/j.comnet.2018.07.016, Jul 2018.

[2] T. Guo, T. Lin, and N. Antulov-Fantulin, "Exploring interpretable lstm neural networks over multi-variable data," Tech. Rep., May 2019.

[3] P. Joseph, H. Jason, and C. David, "Versatile low power media access for wireless sensor networks," in *Second ACM Conference on Embedded Networked Sensor Systems (SenSys)*, Baltimore, Maryland, USA, November 2004, pp. 95–107.

[4] B. Michael, gary V. Yee V, E. Anderson, and H. Richard, "X-mac: A short preamble mac protocol for duty-cycled wireless sensor networks." in *ACM SenSys*, Boulder, Colorado, 2006, pp. 307–320.

[5] B.Yildiz, J.I.Bilbao, and A.B.Sproul, "A review and analysis of regression and machine learning models on commercial building electricity load forecasting," *Elsevier Renewable and Sustainable Energy Reviews*, vol. 73, pp. 1104–1122, June 2017.

[6] O. Oyelade, Oladipupo, and Obagbuwa, "Application of k means clustering algorithm for prediction of students academic performance," *(IJCSIS) International Journal of Computer Science and Information Security*, vol. 7, September 2010.

[7] T. L. Griffiths and J. B. Tenenbaum, "Predicting the future as bayesian inference: People combine prior knowledge with observations when estimating duration and extent," Technical Report, September 2011.

[8] Y. Cheng, C. Xu, D. Mashima, V. L.L., and T. Wu, "Powerlstm: Power demand forecasting using long short-term memory neural network," *International Conference on Advanced Data Mining and Applications(ADMA)*, pp. 727–740, October 2017.

[9] S. Selvin, E. A. Gopalakrishnan, V. R, and V. K. Menon, "Stock price prediction using lstm, rnn and cnn-sliding window model," *International Conference on Advances in Computing*, vol. 132, doi:10.1109/ICACCI.2017.8126078, September 2017.

[10] S. Cui, "Solar energy prediction and task scheduling for wireless sensor nodes based on long short term memory," *IOP Conf. Series: Journal of Physics*, 2018.

[11] B. Wang, L. Zhang, H. Ma, H. Wang, and S. Wan, "Parallel lstm-based regional integrated energy system multienergy source-load information interactive energy prediction," *Complexity, WILEY*, vol. 12, pp. 1–13, 2019.

[12] M. Taksir and S. Aktar, "Data-driven time series based prediction in smart home appliance energy consumption," *International Journal of Computer Applications*, vol. 15, pp. 41–46, 2019.

[13] M. Li, H. Zhang, and P. Qiu, "Predicting future locations with deep fuzzy-lstm network," *AGCS*, vol. 47, doi :10.11947/j.AGCS.2018.20170268, 2018.

[14] R. Laidi and D. Djenouri, "UDEPLOY: user-driven learning for occupancy sensors deployment in smart buildings," in *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom) Workshops proceedings*, 2018, pp. 209–214.

[15] D. Djenouri, R. Laidi, Y. Djenouri, and I. Balasingham, "Machine learning for smart building applications: Review and taxonomy," *ACM Computing Surveys*, vol. 52, no. 2, pp. 24:1–24:36, 2019.

[16] Y. Djenouri, D. Djenouri, and J. C. Lin, "Trajectory outlier detection: New problems and solutions for smart cities," *ACM Trans. Knowl. Discov. Data*, vol. 15, no. 2, pp. 20:1–20:28, 2021. [Online]. Available: https://doi.org/10.1145/3425867

[17] Y. Djenouri, A. Belhadi, J. C. Lin, D. Djenouri, and A. Cano, "A survey on urban traffic anomalies detection algorithms," *IEEE Access*, vol. 7, pp. 12 192–12 205, 2019.

[18] A. Belhadi, Y. Djenouri, G. Srivastava, D. Djenouri, J. C. Lin, and G. Fortino, "Deep learning for pedestrian collective behavior analysis in smart cities: A model of group trajectory outlier detection," *Inf. Fusion*, vol. 65, pp. 13–20, 2021. [Online]. Available: https://doi.org/10.1016/j.inffus.2020.08.003

[19] C. Couvreur, "Hidden markov models and their mixtures," catholic university of Louvain, Faculty of Sciences, TechnicalL Report, 1996.

[20] Z. Hu, J. Zehang, and Y. GE, "Handling vanishing gradient problem using artificial derivative," *IEEE Access*, 2021. [Online]. Available: https://creativecommons.org/licenses/by-nc-nd/4.0/

[21] A. Bookstein, Vladimir, and T. Raita, "Generalized hamming distance," Computer Science Department, University of Turku, Tech. Rep., Octobre 2002.