LSTM Deep Learning Model for Workload Prediction in Massive Machine Type Communication Networks

Sarder Fakhrul Abedin, Md. Shirajum Munir, Choong Seon Hong
Department of Computer Science and Engineering, Kyung Hee University, Korea
Email: {saab0015, munir, cshong}@khu.ac.kr

Abstract

In this paper, the problem of workload prediction for the massive machine type communication (mMTC) networks in a remote monitoring system has been studied for enabling proactive resource scheduling. Therefore, first, we identify the key wireless network environment features and preprocess the features using the Density-based spatial clustering of applications with noise (DBSCAN) algorithm. As a result, the mMTC sensors and devices are clustered in a group and labeled as monitoring zones which reduces the sparsity of the observation space. Second, we apply a deep learning model, named Long short-time memory network (LSTM) to predict the workload of different monitoring zones in the mMTC network over the time horizon. Finally, we perform the simulation analysis on a real wireless environment dataset and show the effectiveness of the proposed approach.

1. Introduction

Over the years, the Internet of Things (IoT) based data services have become prominent and the number of IoT devices in the network continues to grow exponentially [1]. Therefore, the IoT data services and applications has been broadly categorized into three main types such as, enhanced mobile broadband (eMBB), ultra-reliable and low latency communication (URLLC), and massive machine type communication (mMTC) considering the dynamics of the traffic and stringent quality of service (QoS) requirements [2][3]. Moreover, in the era of big data, careful planning is necessary for the mMTC data services that generate a massive amount of environmental data for different IoT applications such as, remote monitoring system [4], ambient assisted living system [5], medical cyber physical system [11], etc. Furthermore, the mMTC data services are necessary to provide up-to-date information at any remote monitoring sites for rigorous data analysis.

In [6], the authors have proposed LSTM for finding the temporal energy demand and reduced the root mean squared error (RMSE). In [7], the authors proposed a recurrent neural network based energy demand prediction in the smart-grid framework. In [8], the authors proposed a short-term traffic forecast in the intelligent transportation system.

The main contribution of the paper is to propose, first, a Density-based spatial clustering of applications with noise (DBSCAN) algorithm to discretize the monitoring zones in the mMTC network. Second, using the preprocessed data, we apply the long short-term memory (LSTM) network based workload prediction for the mMTC network considering a remote monitoring environment. Finally, we perform experiments on real workload dataset for evaluating the efficacy of the proposed approaches.

2. LSTM Model for Workload Prediction

In Fig. 1, we consider massive machine type communication network m number of mMTC wireless gateways $M = \{1, \dots, M\}$ are deployed for remote environment monitoring. Each $m \in M$ aggregates environmental sensing data from as set of mMTC sensors and devices K which are resource constrained devices use the short-range communication medium for data transmission. Therefore, we consider each mMTC gateway works as aggregator for different monitoring zones Z_K. The on-site controller O designates the monitoring zone using the clustering method considering different deployment features. Since the buffer capacity of the mMTC devices are limited, therefore, the mMTC devices transmits the aggregated data from different data sources to the nearest SBSs $n \in N$ for enhanced application specific data services.

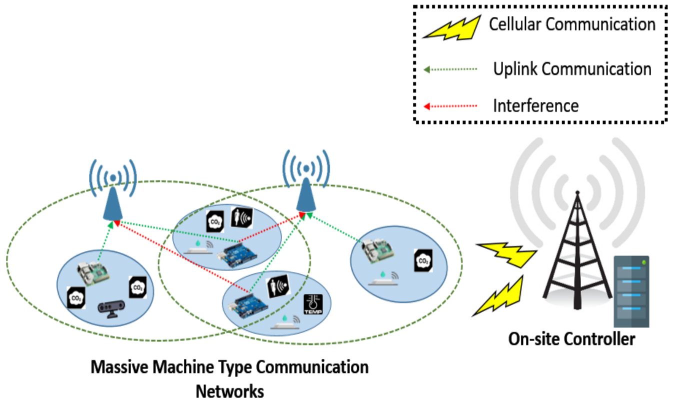

Figure 1 System model for workload prediction in mMTC

Since the data from different observation zones Z_K at each $m \in M$ arrives in a sequential pattern, the data offloading requests from $m \in M$ to $n \in N$ also arrives in a sequential pattern. Therefore, we model the LSTM model for real-time prediction of request workload of Z_K to $n \in N$. We also consider the uplink co-channel interference between the mMTC gateways and the SBSs. This feature is covered in the dataset that we have utilized for the workload prediction. In the experiment results we discuss about different aspects of the dataset.

A memory cell is the core unit of a LSTM network[6] where the input of the historical data at a time t is fed in the cell of the LSTM. The components of the LSTM cell is defined as forget gate f_t, candidate layer $\hat{C}_t$, input gate I_t, hidden state H_t and memory state C_t.In this model we apply sigmoid function for input gate, forget gate, output gate and tanh function for candidate layer. Moreover, the hidden state and memory state are represented as vectors. The input of the LSTM cell $X_t = \{x_1, \ldots, x_j\}$ is based on the preprocessed workload data features x_j from the observation zones. The forget gate at t is defined as,

$$f_t = \sigma(X_t * U_f + H_{t-1} * W_f) \quad (1)$$

In (1), U_f and W_f are the weights of the forget gates. In the candidate layer we apply the tanh function as the squash function which is represented as,

$$\hat{C}_t = \tanh(X_t * U_c + H_{t-1} * W_c) \quad (2)$$

In (2), U_c and W_c are the weights of the forget gates in the candidate layer. The input gate I_t is defined as,

$$I_t = \text{sigmoid}(X_t * U_i + H_{t-1} * W_i) \quad (3)$$

In (3), sigmoid function squashes the input vector where U_i and W_i are the weights of the input gates. Subsequently, the current state cell memory is defined in (4) as,

$$C_t = f_t * C_{t-1} + I_t * \hat{C}_t \quad (4)$$

The output of the LSTM cell is squashed by the sigmoid function and represented as,

$$O_t = \text{sigmoid}(X_t * U_o + H_{t-1} * W_o) \quad (5)$$

In (5), U_o and W_o are the weights for the output gate. Finally, the hidden state is defined with the tanh function in (6) as,

$$H_t = O_t * \tanh(C_t) \quad (6)$$

At each time t, C_t and H_t are transmitted to the next time $t+1$ and the whole process is repeated. The LSTM model is executed at the on-site controller and the trained model is dispersed to the mMTC gateways. The mMTC gateways then can utilize the prediction model to proactively allocate the network resource to the different observation zones.

Table 1: Experiment Settings

Number of features	19
Test dataset size	0.33
Epochs	30
Batch size	512
Learning rate	0.001
Decay	1e-6
Optimizer	Adam optimizer [9]

3. Experiment Results

In the experiment, we use the dataset [10] for predicting the workload of the mMTC network. The dataset contains various KPI performance measurements which are discussed in detailed in [10]. The preprocess the data using the DBSCAN where different mMTC sensors and devices are clustered based their location, SNR, and speed. Then we frame the dataset as a supervised learning problem and the input variables are normalized.

The we perform integer encoding on the zone label feature obtained from the DBSCAN. Subsequently, we define and fit the model where we split the dataset into training and testing dataset. The experimental settings are discussed in table 1. The loss of the forecast results is measured using the mean-squared-error (MSE).

In Fig. (2), we evaluate the workload prediction accuracy using the mean-squared-error where the training accuracy is 76.7% whereas the testing accuracy is 76.4%.

In Fig (3), we show the workload prediction for a single zone over different 100 timeslots. The arrival of the workloads is designed as binary variable where 1 means

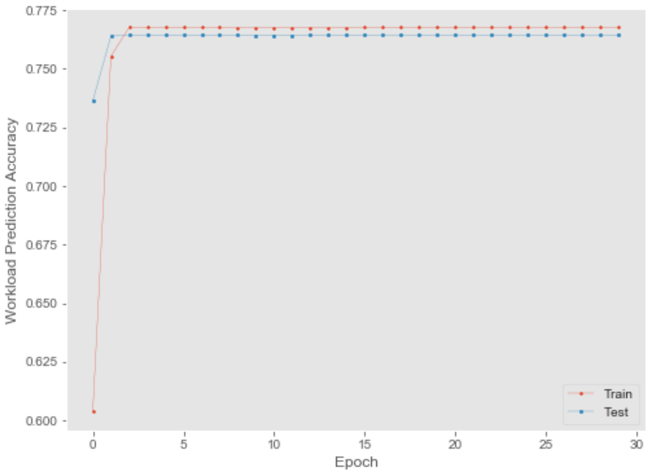

Figure 2 Training and testing accuracy over epochs

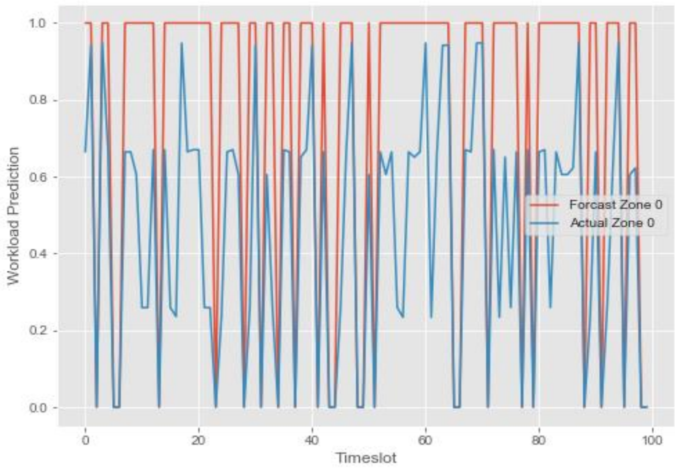

Figure 3 Workload prediction at different timeslots for single observation zone

request and 0 mean no request. There are some discrepancies between the actual arrival of the particular zone and the forecast result which can be reduced by training with large dataset and fine-tuning.

4. Conclusion

In this paper, we have proposed an LSTM model for workload prediction in Massive Machine Type Communication (mMTC) Networks. The design of the LSTM model is suitable for the time-series prediction with heterogeneous workload arrival. The experiment results illustrate the effectiveness of the proposed approach on real workload dataset. In future, we will extend the paper considering the proactive resource allocation using the current model.

Acknowledgement

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the Grand Information Technology Research Center support program (IITP-2018-2015-0-00742) supervised by the IITP(Institute for Information & communications Technology Promotion)" *Dr. CS Hong is the corresponding author.

References

[1] Sarder Fakhrul Abedin, Md Golam Rabiul Alam, Rim Haw, Choong Seon Hong, "A System Model for Energy Efficient Green-IoT Network," The International Conference on Information Networking (ICOIN 2015), Jan 12-14(13), 2015, Siem Reap, Cambodia.

[2] Sarder Fakhrul Abedin, Md. Golam Rabiul Alam, S. M. Ahsan Kazmi, Nguyen H. Tran, Dusit Niyato, and Choong Seon Hong, "Resource Allocation for Ultra-reliable and Enhanced Mobile Broadband IoT Applications in Fog Network," IEEE Transactions on Communications, Vol.67, Issue 1, pp.489-502, January 2019.

[3] Sarder Fakhrul Abedin, Anupam Kumar Bairagi, MD. Shirajum Munir, Nguyen H. Tran, and Choong Seon Hong, "Fog Load Balancing for Massive Machine Type Communications: A Game and Transport Theoretic Approach," IEEE Access, Vol.7, pp.4204-4218, December 2018

[4] Sarder Fakhrul Abedin, Md Golam Rabiul Alam, Anupam Kumar Bairagi, Ashis Talukder, Choong Seon Hong, "UAV-assisted Intelligent Crowdsourcing in Natural Calamity" The International Symposium on Perception, Action, and Cognitive Systems (PACS 2016), Oct. 27-28, 2016, Seoul, Korea

[5] Md Golam Rabiul Alam, Rim Haw, Sung Soo Kim, Abul Kalam Azad, Sarder Fakhrul Abedin and Choong Seon Hong, “EM-psychiatry: An Ambient Intelligent System for Psychiatric Emergency,” IEEE Transactions on Industrial Informatics, pp. 2321-2330, Vol. 12, No. 6, pp. 2321-2330, December 2016

[6] Md. Shirajum Munir, Sarder Fakrul Abedin, Anupam Kumar Bairagi, Sun Moo Kang, Choong Seon Hong, "Temporal Energy Demand Extrapolation for Mobile Edge based on Computational Task in Smart-Grid Framework", KCC 2018, 2017.6.20~6.22.

[7] Md. Shirajum Munir, Sarder Fakhrul Abedin, Md. Golam Rabiul Alam, Do Hyeon Kim, Choong Seon Hong, "RNN based Energy Demand Prediction for Smart-Home in Smart-Grid Framework", KSC 2017.12.20.~12.22.

[8] Zhao, Z., Chen, W., Wu, X., Chen, P.C. and Liu, J., 2017. LSTM network: a deep learning approach for short-term traffic forecast. IET Intelligent Transport Systems, 11(2), pp.68-75.

[9] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).

[10] Raca, Darijo, et al. "Beyond throughput: a 4G LTE dataset with channel and context metrics." Proceedings of ACM Multimedia Systems Conference (MMSys 2018). Association for Computing Machinery, ACM, 2018.

[11] Mowla, Nishat I., Inshil Doh, and Kijoon Chae. "On-Device AI-Based Cognitive Detection of Bio-Modality Spoofing in Medical Cyber Physical System." IEEE Access 7 (2019): 2126-2137.