Turkish Journal of Computer and Mathematics Education Vol.12 No.12 (2021), 1257-1264

Research Article

Development of Hybrid model for data prediction and improving network performance in wireless sensor networks

[1]Anand Dohare, [2]Tulika

[1] Research Scholor, SHAUTS University, Allahabad, India.
[2]Assistant Professor, SHAUTS University, Allahabad, India.
[3]Assistant Professor, School of Computer Science & Engineering, Galgotias University, Greater Noida, Inida
[1]dohare.anand@gmail.com, [2] tulika.tulika@shiats.edu.in, [3]basetty.mallikarjuna@galgotiasuniversity.edu.in

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 23 May 2021

Abstract
In last decade lots of work done in the field of data prediction, various approaches used for data prediction like ARMA, Kalman Filter, LMS and recently work on Deep Learning based LSTM approach. In current scenario sensor is best resource for data collecting and used in various place like border surveillance, security purpose, monitoring and so on. The limitation of sensor nodes are low battery power, low memory and low computation power.
In this paper we proposed a hybrid model which is work in two steps in first step Deep learning-based LSTM approach for data prediction in wireless sensor network and in second step apply feed forward filter with gateway to improve network performance, simulation process is done by Spyder 3.8 with intel dataset it gives better result.

Key word: - ARMA, LSTM, Deep Learning, Data Prediction, LMS, Gateway.

Introduction:

Wireless sensor network is only single choice to collecting information, data gathering and monitoring environmental and physical condition since 1940, it is frequently sense temperature, humidity and vibration So widely used in all sensitive places including military for border surveillance, in agriculture for Crop protection, in house security used as burglar alarm as well as used in most of daily uses appliances and traffic monitoring[1]. Challenges of sensor nodes are low battery power, low memory space and low computational speed. Lots of work has been done to overcome such challenges but day by day increasing of data transmission speed and huge number IoT based appliances, challenges are not overcome. Data prediction method is best solution for decreasing such type of problem. Data prediction is done by three approaches time series forecasting, stochastics and algorithms approach. Stochastics approach mostly used in probability density function while algorithm used in heuristic function, data prediction in wireless sensor networks widely used time series forecasting in this approaches LMS, Gray model and Kalman filter widely used[5].

Since last few year wireless sensor network widely used in domestic appliances and other sensitive areas like military and also for space research with high speed data transmission, so it is necessary to modify existing prediction method, however researchers continuously working for implementing of machine learning concept in data prediction. Machine learning is very strong strategy for providing accurate result with large scale of data as well as improving network performance. Nowadays, most of researchers and scientist work to make machine as intelligent by using Artificial intelligence. AI is very broad area most of software as well as other company divert to work with this like Google Alexa and other. Machine learning is a part of AI and further divided in supervised and unsupervised learning and also subdivided in Deep learning, reinforcement learning and CNN which provide accurate result at acceptance level with large scale of data without need of prior information.

The rest of the paper is organized as follows: the next section introduced related work and third section include brief description of proposed model of data prediction using deep learning and forth section compiled simulation result and conclusion of proposed approach.

Related work

Data prediction is widely used in wireless sensor network for traffic monitoring, weather forecasting, financial prediction and space research. Most popular and oldest approach for data prediction is Kalman filter based dual prediction which is also used by NASA in Apollo mission for successful returning of Neel Arm strong from moon surface in 1969, in this approach each node work as a filter and estimate next measurement at same time the sink node or base station also work with measured value that is threshold value measured by sink node and compared with predicted value if both value are same the data will stored otherwise discard it.

Olston [8] proposed TRAPP based data reduction used query-based optimization and deliberated different models of the sensor node, it performs data transmission between synchronous intervals, it is mainly focused on energy savings but not solve the major challenging issues. Rajagopalan [9] survey on various data normal operational techniques and stochastic models to reduce the communication overhead, remove data redundancy and increase the battery lifetime the data need not be reported on every sink node where the sink node does not give the information to the root node, this model fails on aggregated data in this model lack communication and also cannot improve memory space. Biljana Risteska Stojkoska [13] explain the LMS based data prediction method, in this process used previous knowledge of data, it implemented on variable step-size method and proved 95% aggregated data has improved on the cluster network. El-Telbany [16] suggest to increase the battery life and improve energy consumption, while transmission if any node failure source and destination provides the same prediction values, the proposed methodology also used the same prediction approach. Seyed Ahmed [9] proposed a A Hybrid Prediction Model for Energy-Efficient Data Collection in Wireless Sensor Networks based on decision tree (DT), autoregressive integrated moving average (ARIMA), and Kalman filtering (KF) methods which is mainly used for decrease unnecessary data transmission by predict the data sampling requirement of sensor nodes. To perform data sampling predictions in the WSNs efficiently, clustering and data aggregation to each cluster head are utilized, mainly to reduce the processing overheads generating the prediction model.

All of above work for data prediction are properly working for improving battery power of sensor node as well as data transmission and fault detection in wireless sensor network without use of any machine learning or deep learning algorithm but in case of environment monitoring, weather forecasting and border surveillance we can apply deep learning algorithm for data features for prediction. For such type of prediction CNN and LSTM widely used however RNN also used in some places but it is mostly used in time series analysis. LSTM is special type of recurrent neural network able to learning and memorizing for long term. For reducing the energy consumption in wireless sensor network, Hamed Nazaktbar [10] used Reinforcement learning technique to improve prediction rate by the help of dual prediction such type of model leaned environmental signals collection. A time series predictive model proposed by Adrien Russo[11] this model applicable wireless sensor networks based on self- organizing algorithm which is used for managing system integrity and anomaly monitoring, in this system when any node will not properly working or stop data collecting, the value simultaneously change according to predictive model. Ki-Seong Lee [12] proposed Deep learning–based real-time query processing for wireless sensor networks in this work a monitoring framework used for wireless sensor network to streaming data analysis based on deep learning. This model is applicable where time requirement is strictly followed and result of data analysis is required in time bond while complete data are not received. Deep learning query generator, Query processor and deep learning predictor are used for correcting data. Cheng et al [13] also suggest a data prediction model for wireless sensor network using multi-node and multi-feature where multi nodes refer to sensor node, based on bidirectional long short-term memory (LSTM) network in this work he quantify the similar features between various sensory node and construct a bidirectional LSTM based model which is used temporal correlation between the sensory data and its historical data, the spatial correlation of the sensory data between different nodes, and the low data quality caused by the transmission error of the sensor network. For finding the damage of wind turbine, Alves Maicon Melo [14] proposed a method using time series forecasting with ARIMA and Fuzzy logic to finding effect of temperature on damage of wind turbine, in this model wireless sensor is collect the influence of ambient

temperature on turbine blade and predict the value if temperature is higher than saturation point then it may be possible to damage turbine blade. Q. Liu [15] proposed a method to predict traffic speed farecasting in this method CNN is used to improve LSTM prediction method used three-dimensional data matrices constructed by traffic flow, speed, and occupancy in 15 minutes interval data collected compared with other prediction mechanism. Sinha et al. [16] proposed a data aggregation and prediction model TDPA based on time data prediction. The model generates an estimate of future data to analyze the prediction error and this error is uses the predicted value to save transmission energy consumption when the prediction meets a predefined threshold. Fu et al. proposed a model for predicting traffic flow using long short-term memory networks (LSTM). They compare the performance of ARIMA and LSTM in predicting traffic flow problems and prove that LSTM has certain advantages in traffic flow prediction. Fu et al. [17] proposed a model for predicting traffic flow using long short-term memory networks (LSTM). They compare the performance of ARIMA and LSTM in predicting traffic flow problems and prove that LSTM has certain advantages in traffic flow prediction

Tomoki Kawamura [18] proposed a method for monitoring train condition and predicting the power consumption of sensor node in a wireless sensor network. In this paper the frequent changes in train compartment configuration and environmental changes problem are include for These problems often lead to changes in the network configuration and communication environment between wireless sensor, making it difficult to predict their power consumption. Monte Carlo method is proposed to predict power consumption in WSN. F. U. M. Ullah and Seyed Ahmad Soleymani [3,19] proposed prediction mechanism for residential energy consumption that also used long short term memory in bi directional including multilayer with CNN, this method working in three steps in first step data is collected and preprocessing it then in second step a CNN and M-BDLSTM used for input data and find predicted value which is tested on house hold electricity power consumption.

The Proposed Model

Various model were proposed for data prediction in wireless sensor network, some popular model are Kalman Filter, Gray model, LMS based model these model are used in past but recently researchers work in the field of data prediction and improving network performance using machine learning so more accuracy is found, in continuation of this work we proposed a data prediction model for wireless sensor network using Deep learning. Deep learning focused on creating algorithm that can explain and learn data abstraction at both high and low level, some important applications are game theory and neuroscience etc. Deep learning also support 5G technology with large amount of data and give high accuracy.. Deep learning is also subdivided is various types like CNN, RNN, LSTM, Deep Belief network, Auto encoder and Deep Boltazman machine etc. each algorithm is used for some specific purpose, for data prediction in wireless sensor networks LSTM algorithm is providing better result compare to other mechanism [2].

Our proposed model work in two steps

1. Data Prediction
2. Improving Network Performance

Data Prediction: Sensor node is small low battery powered device, it collects data from outside the world or surrounding area in the form of temperature, humidity, light or other way. Maximum energy of sensor node is consumed in transferring the captured data to base station, it is directly affected to network performance as well as consuming extra power, the challenges is to reduce transmission. Our proposed model is helpful to reduce data transmission by using LSTM algorithm which is special kind of Recurrent neural network (RNN). Rather than other deep learning algorithm RNN, CNN, LSTM have good ability to preserve long term data and avoiding vanishing gradient problem. It works like other neural network having mostly three layers input layer, output layer and hidden layer but LSTM have also forgot Gate and Cell state.

the chain structure of LSTM including three gates. The cell of LSTM is divided in five stage.

A. Cell state (C_t)
B. Hidden state (H_t)
C. Input gate (I_t)
D. Forget Gate (f_t)
E. Output Gate (F_t)

The Cell state is represented by the C_t is an internal memory of cell used for storing long term as well as short term memory while Hidden state H_t is output state which is used for finding current state, previous hidden state and current cell information which is compared with desired output or future coefficient, mainly used for predicting the value, next input gate I_t decide that which information passes to cell state, forget gate F_t decide how much information is passes to current state from previous state to current state. The most important output gate F_t decide how much data is send to previous hidden cell to decide which memory is selected short term or long term. We explain in details how data will be collected from each cell and predict accordingly. Let $X_{t-1,}$ $X_{t,}$ X_{t+1} are the input stat, current state and next state accordingly but in this case input state is concinnated with current state again $Y_{t-1,}$ $Y_{t,}$ Y_{t+1} are the various output respectively. U_t will be the weight vector for hidden layer. In the first step current input is concinnated with previous hidden state and fed into the forget layer with removing non relevant information and combining with input layer for deciding the data for new cell state. By using the equation, we can calculate $C_{t\,,}$ $H_{t,}$ $I_{t,}$ $f_{t,}$ F_t respectively

$$C_t = f_t*C_{t-1}+i_t*C'_t \qquad 1$$

In above equation we calculated C_t which is input cell and combination of f_t and i_t received

$$f_t = \boldsymbol{\sigma\,(W_f.[C_{t-1,}\,H_{t-1,}\,.X_t] + b_f)} \qquad 2$$

$$\acute{\imath}_t = \boldsymbol{\sigma\,(W_i.[C_{t-1,}\,H_{t-1,}\,.X_t] + b_i)} \qquad 3$$

$$o_t = \boldsymbol{\sigma\,(Wo.[C_{t-1,}\,H_{t-1,}\,.X_o] + b_o)} \qquad 4$$

in equation 2, f_t is calculated by using of sigmoidal function with dot multiplication of wait matrix $\mathbf{W_f}$ and input matrix of $\mathbf{C_{t-1,}\,H_{t-1,}\,.X_t}$ also adding bias of $b_{t.}$ on same way $\acute{\imath}_t$ and o_t will calculated, but after calculating f_t, we apply tanh function to calculating h_t and this function will be found by multiplication of

o_t with C_t which will already calculated previously. So h_t will given by the equation 5.

$$h_t = o_t*\tanh(C_t) \qquad 5$$

on computing foregate layer candidate layer and input layer calculate cell state with the help of previous vector after that pointwise multiplication of output and new cell state will gives the new hidden state.

In simple way we explain whole process step by step

1. In first step it is decided by forget gate using sigmoidal function which information is flow in cell state $C_{t.}$
2. In second step which information is stored in cell state, decided in two step first using sigmoidal of input layer for updating value and in second step tanh layer create new vector for added in previous state and provide current state its C_{t-1} to C_t .
3. In third step multiplying old state with f_t for deciding earlier state.
4. In last step we decide actual output value or future state. This output is based on our cell state.

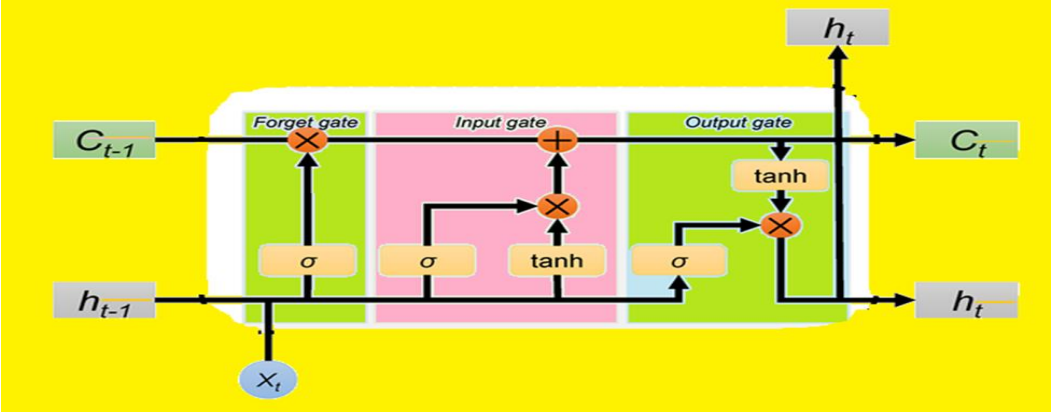

Figure 1: LSTM Network Architecture

We got predicted value from equation 5.

Improving Network Performance

Next important things are to improving network performance which is handled by applying feed forward approach for data processing. Feed forward approach is lie to improving QoS, security as well as network performance [7]. A Gateway have significant importance, it has decision forward approach and basic function of gateway is to filtering the data which decide the received packed is to be forward or not if need to forward then send to base station otherwise reject it. Routing and traffic aggregation are major function of gateway and it also responsible for storing, computing as well as deciding that where data is sending or receiving from concern path [6]. These features of gateway will minimize energy consumption of sensor node and improve network performance. In figure 1 feed forward approach is showing, data is collected is various sensor nodes and send to data center but gateway have important role between data center and sensor node. Only specify data is to be send to data center after confirm from feedforward approach, the request is sent to datacenter and wait for reply message, if reply message is not receiving in given time interval then packet is not sent to data center. Hence security will increase and unnecessary bourdon of data transmission also decrees it will improve network performance.

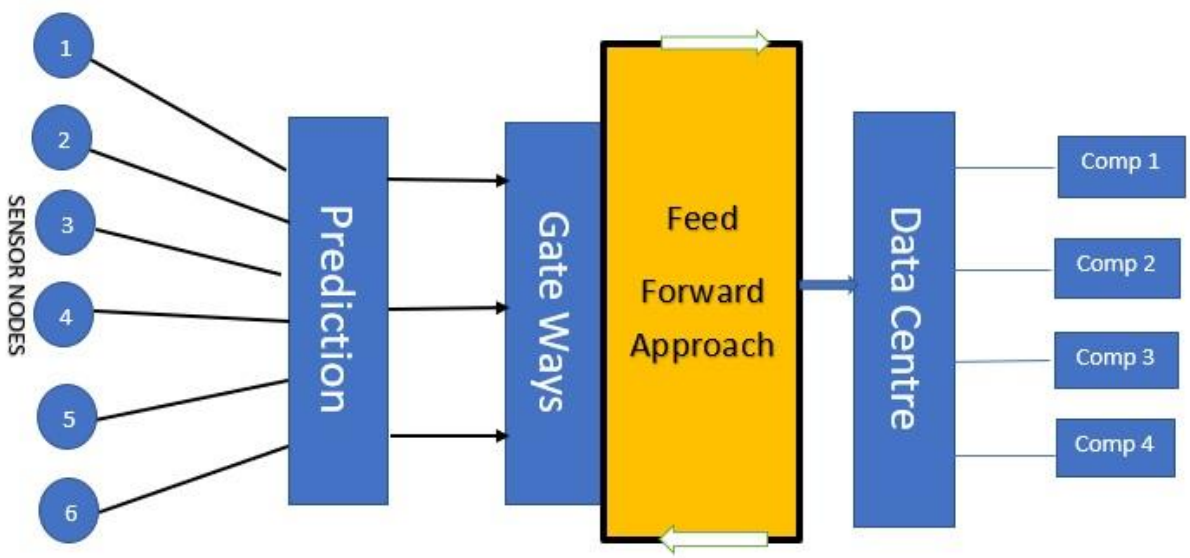

Figure 2. Feed forward approach with gateway

In figure 2 all sensors node collect data from outside world and transferring to gateway through LSTM based prediction strategy and applying feed forward approach to confirm data requirement and authentication after confirmation received from data center and got acknowledgement, the data will send to data center otherwise it will rejected at gateway level. It is useful to reducing unnecessary data transmission and increasing network performance.

Algorithm: Feed forward Approach for Data Processing.

Step 1: Initialize the sensor nodes at network.

Step2: Sensor nodes collect the data and forwards to the Prediction filter then gateway.

Step 3: Gateway forwards the request message to the datacenter.

Step 4: Datacenter responds with a feedback about the timestamp of the data required

Step5: Gateway collects the data at particular time requested by the server and forwards to datacenter.

Step6: User can access the data any time that are stored in the datacenter.

Simulation and Result

The proposed model is tested with python (Spyder 3.8) for data prediction with intel data set on 50000 data with temperature and epoch graph, it gives 18.36 ERMS value on applying LSTM algorithm with training data and 23.24 ERMS value on testing data. Testing and training data are separated with 25:75 ratio of whole data while simulate same data with another prediction approach ARMA it gives 28.72 ERMS value with same factor temperature and epoch. On other hand network performance is also improved and gives batter result with using feedforward approach with gateway as compared to without applying feedforward approach and gateway also monitoring service regularly.

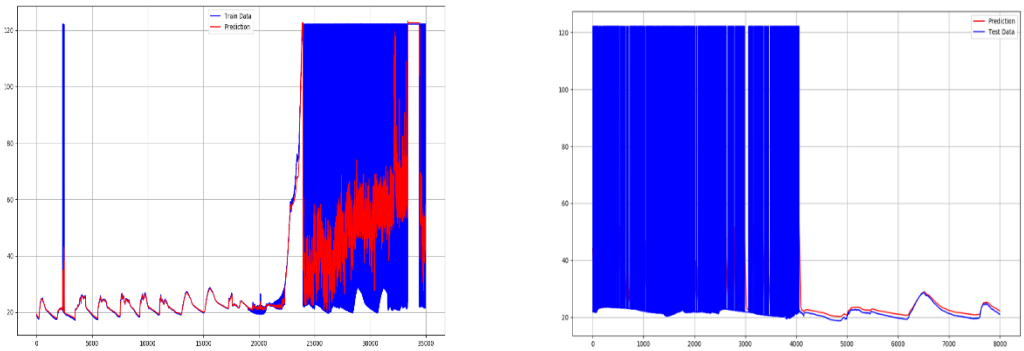

Figure (a) ERMS value with Training data set (b) ERMS value with testing data set.

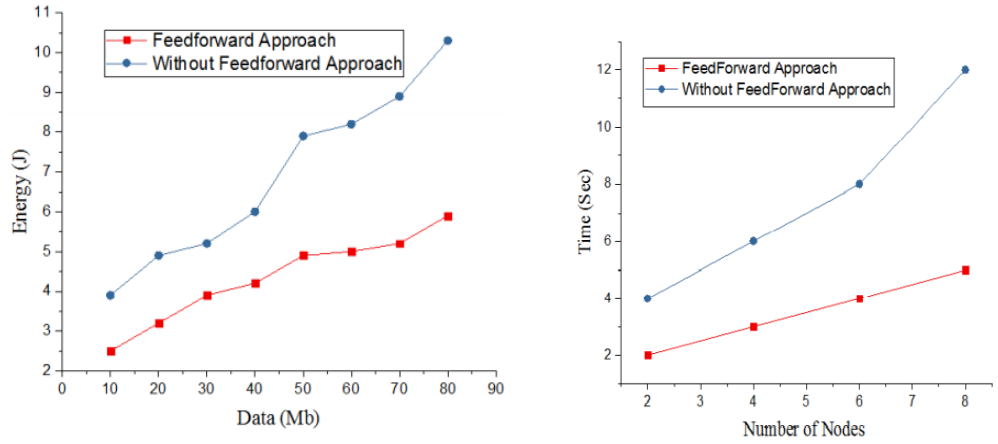

Figure (a) Energy consumption of the sensor nodes (b) Service discovery of sensor nodes

Conclusion:

Data prediction using LSTM provide better performance rather than other method and feedforward approach with gateway is help to improve network performance. Result is proven with Spyder 3.8 with intel data set on calculating Root mean square error with temperature and epoch.

References

1. Vandana Jindal," History and Architecture of Wireless Sensor Networks for Ubiquitous Computing", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 7, Issue 2, February 2018, ISSN: 2278 – 1323.
2. S.N. Mohanty, E.L. Lydia, M. Elhoseny et al "Deep learning with LSTM based distributed data mining model for energy efficient wireless sensor networks" https://doi.org/10.1016/j.phycom.2020.101097 1874-4907/© 2020 Elsevier B.V. All rights reserved.
3. Seyed Ahmad Soleymani" A Hybrid Prediction Model for Energy-Efficient Data Collection in Wireless Sensor Networks" Symmetry 2020, 12, 2024; doi:10.3390/sym12122024, www.mdpi.com/journal/symmetry.
4. Xiaofei Yan" Real-Time Identification of Smoldering and Flaming Combustion Phases in Forest Using a Wireless Sensor Network-Based Multi-Sensor System and Artificial Neural Network" Sensors 2016, 16, 1228; doi:10.3390/s16081228.
5. Mohammed E. El-Telbany" An Energy-Efficient Wireless Sensor Networks Utilizing LMS Filter and Matrix Completion" International Journal of Applied Engineering Research, ISSN 0973-4562 Volume 12, Number 5 (2017) pp. 591–597.
6. Nef, M.A., Perlepes, L., Karagiorgou, S., Stamoulis, G.I. and Kikiras, P.K., 2012. Enabling qos in the internet of things. In Proc. of the 5th Int. Conf. on Commun., Theory, Reliability, and Quality of Service (CTRQ 2012) (pp. 33-38).
7. Krishna, P. Venkata, et al. "Learning automata based decision making algorithm for task offloading in mobile cloud." Computer, Information and Telecommunication Systems (CITS), 2016 International Conference on. IEEE, 2016.
8. Olston, C., & Widom, J. (2005). Efficient monitoring and querying of distributed, dynamic data via approximate replication. IEEE Data Engineering Bulletin.

9. Seyed Ahmad Soleymani and Shidrokh Goudarzi " A Hybrid Prediction Model for Energy-Efficient Data Collection in Wireless Sensor Networks" Symmetry 2020, 12, 2024; doi:10.3390/sym12122024.
10. H. Nazaktabar, K. Badie, and M. N. Ahmadabadi, ''RLSP: A signal prediction algorithm for energy conservation in wireless sensor networks,'' Wireless Netw., vol. 23, no. 3, pp. 919–933, 2017.
11. A. Russo, F. Verdier, and B. Miramond, ''Energy saving in a wireless sensor network by data prediction by using self-organized maps,'' Proc. Comput. Sci., vol. 130, pp. 1090–1095, 2018.
12. Ki-Seong Lee "Deep learning–based real-time query processing for wireless sensor network" International Journal of Distributed Sensor Networks 2017, Vol. 13(5)
13. Cheng et al "Data prediction model in wireless sensor networks based on bidirectional LSTM" EURASIP Journal on Wireless Communications and Networking (2019) 2019:203
14. M. M. Alves, L. Pirmez, S. Rossetto, F. C. Delicato, C. M. de Farias, P. F. Pires, I. L. dos Santos, and A. Y. Zomaya, ''Damage prediction for wind turbines using wireless sensor and actuator networks,'' J. Netw. Comput. Appl., vol. 80, pp. 123–140, Mar. 2017
15. Q. Liu, B. Wang, and Y. Zhu, ''Short-term traffic speed forecasting based on attention convolutional neural network for arterials,'' Comput.-Aided Civil Infrastruct. Eng., vol. 33, no. 11, pp. 999–1016, 2018
16. A. Sinha, D.K. Lobiyal, Prediction models for energy efficient data aggregation in wireless sensor network. Wireless Personal Communications 84(2), 1325–1343 (2015)
17. R. Fu, Z. Zhang, L. Li, Using LSTM and GRU neural network methods for traffic flow prediction (2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, 2016), pp. 324–328 29.
18. T. Kawamura, S. Ryuo, and N. Iwasawa, ''Power consumption prediction method for train-health monitoring wireless sensor networks,'' Electron. Commun. Jpn., vol. 101, no. 6, pp. 24–32, 2018
19. F. U. M. Ullah et al.: "Short-Term Prediction of Residential Power Energy Consumption" Digital Object Identifier 10.1109/ACCESS.2019.2963045
20. H. Cheng et al.:" Multi-Step Data Prediction in Wireless Sensor Networks" Digital Object Identifier 10.1109/ACCESS.2019.2937098