A Prediction Model of IoT Data Using Long Short-Term Memory Neural Network

Meiyu Wen[1,2], Dandan Che[1,3], Jean-Pierre Niyigena[1,3], Ruohan Li[1], and Qingshan Jiang[1]

[1]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

[2]Software Institute of Nanjing University, Nanijing 210093, China

[3]Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Beijing 100049, China

my.wen@siat.ac.cn, dd.che@siat.ac.cn

Keywords: component; commercial hotel kitchens; data analysis; prediction models; anomaly detection

Abstract: In order to ensure the freshness stability of different ingredients in the cold storage room of commercial hotels' kitchen, the real-time monitoring of temperature and humidity is required. Thus, the establishment of a model to predict the temperature and humidity for future and for conducting early warning analysis on the temperature that may exceed the threshold is needed, so that relevant personnel can take defense measures before the temperature changes drastically. This paper detects the processing of abnormal value and the missing value of temperature and humidity according to the sensors’ receiving time. Long Short-Term Memory (LSTM) model is used for temperature and humidity time series prediction. Then, the result is compared with the prediction result using traditional statistical model of Autoregressive Integrated Moving Average (ARIMA). The final findings show that the predictive accuracy of the LSTM model is significantly better than the traditional model of ARIMA and the final temperature prediction result error is quite small.

1. Introduction

The structure of the cold storage of commercial hotels is complex. As a result, it usually affects the signal penetration ability of wireless equipment and the data collector devices may receive unstable signal [1]. Considering that a cold storage of a five-star hotel requires high-level performance, it is necessary to detect abnormal values generated during data transmission of sensors, so that the integrity of data is ensured.

The factors that lead to structure complexity of cold storage are the requirement of a temperature much lower than the temperature of normal cold storage and the double cold temperature storage features such as fresh-keeping and freezing. In order to ensure that there is adequate food, hotels will often replenish fresh food and they need to consume a large amount of food every day [2]. Therefore, the cold storage is frequently accessed and disturbed by human factors. Different kinds of food are stored in cold storage, and in order to keep the freshness of food, strict temperature and humidity control standards must be adopted. In addition, temperature and humidity required by different kinds of food is also very different. In that case, it is necessary to establish a prediction model based on real-time temperature and humidity data and it should have ability to predict the value of future data [3]. According to the predicted values, potential abnormal data can be detected in advance so that relevant personnel can take defense measures to prevent the acute temperature change from happening to ensure stability of the cold storage environment and freshness of food.

This paper proposes a prediction model of abnormal value and missing value based on LSTM and ARIMA model using the temperature and humidity data from commercial hotels’ kitchen cold storage. Moreover, this article compares multiple evaluation indexes of the two models.

  DOI: 10.25236/isaicn.2019.076

2. Related Work

In commercial hotel, the kitchen cold storage is vital to maintain the freshness of food by strictly keeping temperature and humidity stable [5]. Therefore, it is necessary to establish a reliable prediction method so that the responsible personnel can take relevant defense measures before the environment changes abnormally. It is necessary to monitor the abnormal values of the data transmission process of temperature and humidity sensors before establishing the prediction model to ensure the accuracy and effectiveness of data transmission. Currently, wireless sensors can record real-time change of temperature and humidity. Besides, there are abundant matured applications of wireless sensors in China and aboard in the field of temperature and humidity detection of agricultural environment. However, the outliers (*abnormal value* and *missing value*) caused by the loss of those wireless sensors data transmission of devices are not detected in these applications [4].

There are many existing methods of analyzing and forecasting Internet of Things temperature and humidity data. Most of them are based on traditional machine learning methods. However, there are also traditional methods based on the self-regressive moving average model for predicting time series data such as ARIMA [6]. Following by the development of deep learning, many new models that are capable to process time series data are introduced. As an example, Recurrent Neural Networks (RNN) introduces the concept of timing into its network structure, which allows it to effectively process time series data. Long Short-Term Memory (LSTM) neural network, which is developed on this basis, not only takes the temporal correlation between time series into account, but also effectively avoids Gradient Explosion and Gradient Disappearance problem in the training process [7][8][9]. LSTM can make full use of time series data which Internet of Things temperature and humidity real time data can also be regarded as.

3. Prediction Method for IoT Data

As shown in Figure 1, our prediction model of the data is implemented in four steps: data acquisition, pre-processing the collected data, which includes the detection of outliers and the filling of missing values. The pre-processed data is used for model prediction. In this paper, two models of LSTM and ARIMA are used. Finally, the results are analyzed and the evaluation indexes of the two models are compared.

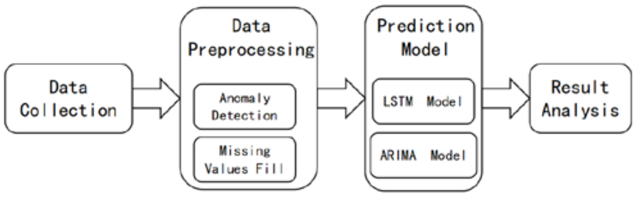


Figure 1. Model prediction flow chart

3.1 Data acquisition process

Firstly, real-time detection data needs to be acquired through data acquisition. The acquisition and transmission process is divided into six steps as Figure 2 shows. Real-time detection data is acquired from commercial hotels' kitchen cold storage sensor and is then transferred to gateway, codec server, data center and finally independent database for query and analysis.

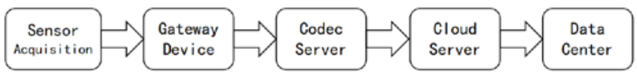


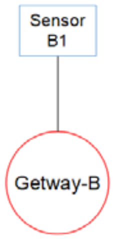
Figure 2. Data collection flow chart

According to the data transmission process, the temperature and humidity data of 24 sensors from January 2018 to May were collected. The temperature and humidity sensor A placed in the hotel cold storage is selected as the research object, and the temperature variable is taken as the prediction object. The data transmission interval of the sensor is 3 minutes. The temperature data from May 15 to 18 is selected as the training set, and a model is created to predict the temperature data of the first half of May 19.

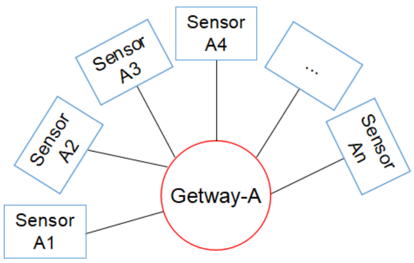
3.2 Anomaly Detection

The network topology of the gateway with the temperature and humidity sensors of the cold storage of commercial hotels kitchen can be divided into the following two types:

- One-to-one structure: A gateway only transmits data from one sensor, as shown in Figure 3(a).
- One-to-many structure: A gateway needs to transmit data from multiple sensors, as shown in Figure 3(b).



(a)



(b)

Figure 3. Gateway sensor topology

The transmission of a gateway's data can be judged by the following expressions:

$$F_A(x)=\begin{cases}0, & x\notin T\\ 1, & x\in T\end{cases} \tag{1}$$

Among them, T represents the time period to be detected, x represents the actual receiving time of gateway. When $x \notin T$ and gateway A does not transmit data to the corresponding sensor in a time period, $F_A(x)$ is recorded as 0; when gateway A has transmitted data to the corresponding sensor in a time period, $F_A(x)$ is recorded as 1.

Therefore, for a gateway with only one sensor, it is only needed to detect the loss of packets transmitted by the gateway. For a gateway with many sensors, it is needed to detect the loss of packets transmitted by the gateway first. If there is no loss of packets, the corresponding sensors of the gateway do not lose any packets, and if there is loss of packets, it is necessary to continue pinning. The packet loss is detected by sensors. The data selected in this paper are all missing data.

3.3 Missing values filling

The commonly used methods of filling missing values include mean, median, mode, random number filling, etc. Because the cold storage environment is relatively stable, the temperature value is not significantly determined by short-term factors. In view of the small amount of missing data in the data transmission process selected in this paper, the average filling method is adopted.

Assuming that the cold storage temperature sensor transmits the data once every T minutes and the missing value X_T is expressed, the average value of the last day's data that needs to be filled in time T is filled in X_{ave},with $X_T = X_{ave}$.

3.4 Long Short-Term Memory Neural Network

LSTM is a special kind of RNN. The core idea of LSTM is to select and remove or retain information to the memory unit through a structure called "gate", which avoids the problem of gradient dispersion and gradient explosion in the time dimension. It can flexibly apply the timing characteristics of related learning tasks [10].

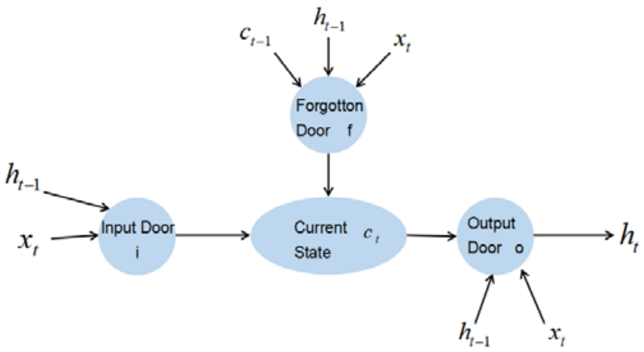


Figure 4. Schematic diagram of a LSTM cell structure

Figure 4 illustrates the schematic diagram of a LSTM unit structure. Each memory unit consists of one tuple and three gates: input gate, forgetting gate and output gate. The forgetting gate and the input gate are the core of LSTM network structure. The function of forgetting gate is the useless information, forgetting gate and input before "forgetting" of circular neural network. Doors can more effectively decide which information should be forgotten and what information should be retained [11].

The formula for each "door" is defined as follows [12]:

Input door:

$$i = sigmoid(W_i[h_{t-1}, x_t]) \tag{2}$$

Forget gate:

$$f = sigmoid(W_f[h_{t-1}, x_t]) \tag{3}$$

New state:

$$c_t = f \bullet c_{t-1} + i \bullet z \tag{4}$$

Output door:

$$o = sigmoid(W_o[h_{t-1}, x_t]) \tag{5}$$

Among them, x_t represents the current input, h_{t-1} for the output of the previous moment, c_t for the current state, W_i, W_o, W_f is a parameter matrix with three dimensions (2n, n).

The input gate decides which information is added to the state c_{t-1} and generates the state c_t according to the current input x_t and the output of the previous moment h_{t-1}; the forget gate decides which part of the information needs to be forgotten according to x_t and h_{t-1}; the output gate decides the output of the current moment according to the latest state c_t, the output h_{t-1} of the previous moment and the current input x_t.

The network structure of this paper is shown in Figure 5. The temperature data is input from the input layer and enters the single-layer LSTM structure. The number of nodes is 16, and the node is randomly removed using the DRPOPOUT layer to prevent over-fitting. The last layer is the fully-connected layer (DENSE), then output the prediction result. According to the trained model, the multi-step prediction is used to obtain the temperature value of the next t step time point.

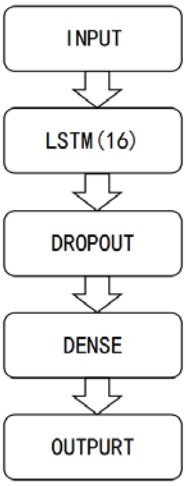


Figure 5. Network structure of prediction model

3.5 Autoregressive Integrated Moving Average Model

ARIMA is short for Autoregressive Integrated Moving Average Model. It was put forward by Box and Jenkins in the early 1970s as a famous time-series approach prediction method, so it is also called Box-Jenkins model and Box-Jenkins method. ARIMA (p, d, q) is called a differential autoregressive moving average model consisting of AR and MA. AR is an autoregressive model with its parameter p, MA is a moving average model with its parameter q and D is the degrees of differencing to make the time series stationary. The so-called ARIMA model refers to the model which transforms the non-stationary time series into stationary time series and then regresses the dependent variable only to its lag value and the present value and lag value of the random error term. ARIMA model includes moving average process (MA), autoregressive process (AR), autoregressive moving average process (ARMA) and ARIMA process according to whether the original sequence is stable or not and the different parts of regression.

In this paper, we judge data as a stationary time series according to the trend of temperature and related analysis diagram. The model identification step selects the ARMA model order as ARMA (2, 2), the AR model order is AR (2), and the MA model order is MA (2).

4. Result analysis

4.1 Long term Memory Neural Network Model Result

The LSTM model is built using Python language. Loss function is set up by: loss = 'mean_square_error', batch size = 1, optimizer = 'adam', period: epochs = 100. The prediction results are as in Figure 6:

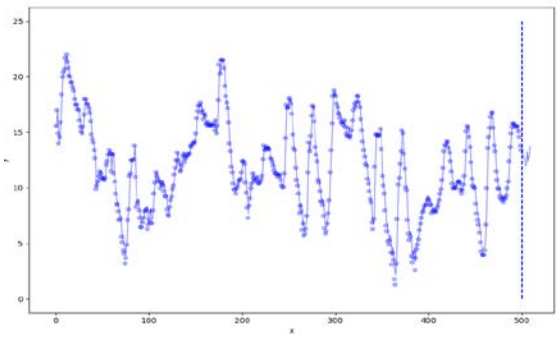


Figure 6. LSTM model fitting

As shown in Figure 6, the circle represents the actual value, the straight line represents the forecasting value and the ten-step forecasting results after the dotted line are less than 0.1.

4.2 Autoregressive Moving Average Model Result

The SAS tool is used to set the q=2 of ARIMA (p, d, q), and the MA (2) model is established. The prediction results are as Figure 7.

As illustrated in Figure 7, the circle represents the actual value, the straight line represents the predicted value, and the ten-step predicted result after the dotted line is half an hour. The fitting effect of the model is poor, and the residual error of the predicted result is greater than 1.

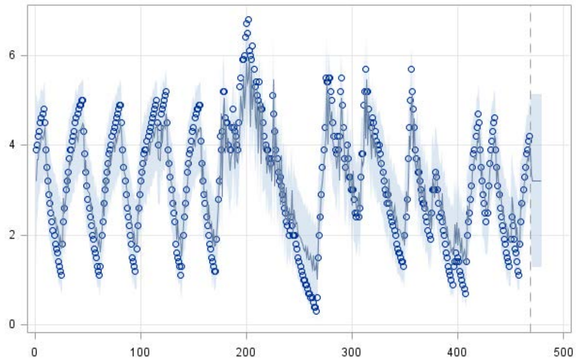


Figure 7. Fitting graph of MA (2) model

The SAS tool is used to set up the p=2 of ARIMA (p, d, q) and establish the AR (2) model. The prediction results are as in Figure 8:

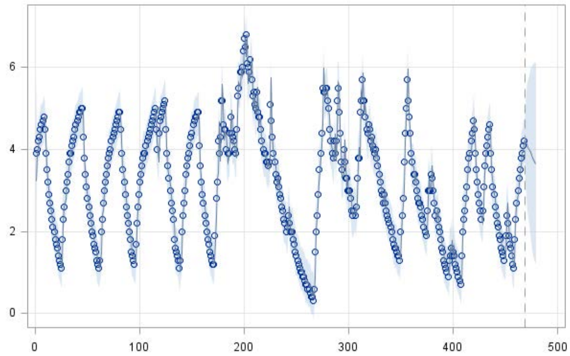


Figure 8. Fitting graph of AR (2) model

As depicted in Figure 8, the circle represents the actual value, the straight line represents the predicted value, and the ten-step predicted result after the dashed line is half an hour. The fitting effect of the model is general. The mean square error of the predicted result is about 0.3.

ARIMA (p, d, q) is set up by SAS tool with P = 2 and q = 2. The ARMA (2,2) model is established. The prediction results are as in Figure 9:

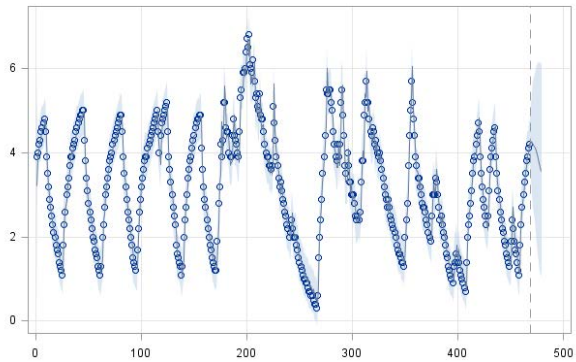


Figure 9. Fitting graph of ARMA (2,2) model

In Figure 9, the circle represents the actual value, the straight line represents the predicted value, and the ten-step predicted result after the dotted line is half an hour. The fitting effect of the model is good. The mean square error of the predicted result is about 0.2.

4.3 Result analysis

As shown in Table I below, the time step t of multi-step prediction temperature is taken as 10. The results of LSTM model are the closest to the actual values, and the results of MA model are not consistent with the actual values.

TABLE I. TABLES OF DIFFERENT MODEL PREDICTIONS

Time	Real value	LSTM predictive value	ARMA predictive value	AR predictive value	MA predictive value
00:03:00	4.4	4.3487	4.2225	4.1943	3.8116
00:06:00	4.5	4.4421	4.1988	4.1318	3.2766
00:09:00	4.7	4.6269	4.1988	4.0419	3.077
00:12:00	4.8	4.7182	4.0482	3.942	3.077
00:15:00	4.8	4.7182	3.9365	3.8421	3.077
00:18:00	4.1	4.0647	3.8103	3.7475	3.077
00:21:00	3.7	3.6781	3.6763	3.6609	3.077
00:24:00	3.3	3.2851	3.5403	3.5832	3.077
00:27:00	3	2.9880	3.4077	3.5145	3.077
00:30:00	2.8	2.7898	3.2827	3.4543	3.077

As shown in Table II below, it is obvious from the table that the residual sum, Mean Square Error (MSE), Root Mean Square Error (RMSE) and Mean Absolute Error rate (MAPE) show a performance of LSTM > ARMA (2,2) > AR (2) > MA (2).

TABLE II. Different models for forecasting and evaluating statistics

Models	Sum of residuals	MSE	RMSE	MAPE
MA	8.3958	1.198155412	1.094602856	0.207889132
AR	1.9875	0.316352179	0.562451935	0.123444593
ARMA	1.7779	0.222571507	0.471774848	0.100860835
LSTM	0.4401	0.004890341	0.069930974	0.010082932

Therefore, according to the temperature and humidity data of kitchen cold storage in commercial hotels, LSTM has the best prediction effect in comparison with ARMA (2, 2), AR (2) and MA (2), which is obviously superior to the other three models.

5. Conclusion

This paper proposes a temperature prediction method approach using deep learning network structure, specifically LSTM. The method performs the detection of *abnormal value* and *missing value* processing among the data received by the sensors. Moreover, it can predict temperature of the next 10 time periods according to the historical data. The experimental results show that the LSTM method has higher accuracy and better prediction effect than the traditional statistical method of ARIMA.

Acknowledgment

This paper is supported by *Shenzhen Discipline Construction Project for Urban Computing and Data Intelligence*; and it is supported by the *Joint Engineering Research Center for Health Big Data Intelligent Analysis Technology*. Also, we would like to thank all the students of the High-Performance Data Mining Key Laboratory, SIAT, Shenzhen, China for their enthusiastic help in the writing of this paper.

References

[1] Oró E, De Gracia A, Castell A, et al. Review on phase change materials (PCMs) for cold thermal energy storage applications[J]. *Applied Energy*, 2012, 99: 513-533.

[2] Changan C . Design of Greenhouse Temperature and Humidity Monitoring System Based on the Internet of Things[J]. *Journal of Changsha University*, 2016.

[3] Yong L , Lu-Lu Y . Development of Temperature and Humidity Monitoring System for Electric Meter Warehouse Based on the Internet of Things Technologies[J]. *Journal of Anhui Vocational College of Electronics & Information Technology*, 2014.

[4] Ruizgarcia L , Lunadei L , Barreiro P , et al. A Review of Wireless Sensor Technologies and Applications in Agriculture and Food Industry: State of the Art and Current Trends[J]. *Sensors*, 2009, 9(6):4728-4750.

[5] Joaquín Gutiérrez, Villa-Medina J F , Nieto-Garibay A , et al. Automated Irrigation System Using a Wireless Sensor Network and GPRS Module[J]. *IEEE Transactions on Instrumentation & Measurement*, 2013, 63(1):166-176.

[6] E Cadenas， W Rivera， R Campos-Amezcua， C Heard. Wind Speed Prediction Using a Univariate ARIMA Model and a Multivariate NARX Model. Energies, 2016, 9(2):109.

[7] Zhang H, Wang Z , Liu D . A Comprehensive Review of Stability Analysis of Continuous-Time Recurrent Neural Networks [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, 25(7):1229-1262.

[8] Haşim Sak, Andrew Senior, Françoise Beaufays. Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition[J]. *Computer Science*, 2014:338-342.

[9] Inc G. Convolutional, long short-term memory, fully connected deep neural networks[C]. *IEEE International Conference on Acoustics. IEEE*, 2015.

[10] Zeyu, Zheng, and Gu Siyu. "TensorFlow: practical Google deep learning framework [M]." (2017): 199-210.

[11] Sak H, Senior A, Beaufays F. Long short-term memory recurrent neural network architectures for large scale acoustic modeling[C]. *Fifteenth annual conference of the international speech communication association*. 2014.

[12] Hochreiter, Sepp, and J. Schmidhuber. "Long short-term memory." *Neural Computation* 9.8(1997):1735-1780.