RESEARCH   **Open Access**

A new method for traffic forecasting in urban wireless communication network

Kaisa Zhang[*], Gang Chuai, Weidong Gao, Xuewen Liu, Saidiwaerdi Maimaiti and Zhiwei Si

Abstract

With the development of wireless devices and the increase of mobile users, the operator's focus has shifted from the construction of the communication network to the operation and maintenance of the network. Operators are eager to know the behavior of mobile networks and the real-time experience of users, which requires the using of historical data to accurately predict future network conditions. Big data analysis and computing which is widely adopted can be used as a solution. However, there are still some challenges in data analysis and prediction for mobile network optimization, such as the timeliness and accuracy of the prediction. This paper proposes a traffic analysis and prediction system which is suitable for urban wireless communication networks by combining actual call detail record (CDR) data analysis and multivariate prediction algorithms. Firstly, a spatial-temporal modeling is used for historical traffic data extracting. After that, causality analysis is applied to communication data analysis for the first time. Based on causal analysis, multivariate long short-term memory models are used to predict future data for CDR data. Finally, the prediction algorithm is used to process real data of different scenes in the city to verify the performance of the entire system.

Keywords: CDR, Spatial-temporal modeling, Causality analysis, LSTM multivariate time series forecasting

1 Introduction

Communication technologies and devices coincide with big data analysis and computing technologies are developing rapidly. Lots of researchers from various domains begin to use the data from the wireless communication network, among which, the call detail records (CDR) data is one of the most measured data sources [1]. The current application of CDR data analysis is quite extensive, including calling pattern, criminal investigation, and urban computing. A CDR is an information structure storing applicable information about a given telephonic activity including a customer of a network framework. A CDR, as a rule, contains spatial and temporal data [2]. Therefore, it is also very important for network analysis and optimization. Through the analysis and modeling of CDR historical data, we can predict the trend of CDR data in the future, so as it plays a decisive role in the technical adjustment of resource allocation and load balancing. This is also the starting point of this

paper. Meanwhile, there are two major challenges in network optimization:

1. Avoiding the dissatisfaction of user experience in advance, which is also the main concern of operators.
2. Because the traffic levels of different scenarios in the city are not same, for example, the traffic volume of a station is higher than that of a residential area, which causes the prediction algorithm to be not very adaptable to the whole urban communication network.

In this paper, in order to solve these problems, we extract CDR data through spatial-temporal modeling based on grids of urban communication network. The extracted data would be used for further causality analysis. Causality is mainly reflected in space, which means the causality of traffic variation trend between predicted area and surrounding areas. From the perspective of data analysis, according to the Granger causality test, the data of surrounding areas leading to traffic changes in the prediction area are obtained. Then, the extracted traffic

* Correspondence: kaisa@bupt.edu.cn
Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, People's Republic of China

data of adjacent grids which have strong causality, and the greatest impact on the prediction area are used for the final multivariate time series prediction. LSTM is used for multivariate time series prediction.

For urban wireless communication networks, grid partitioning is a common method of spatial-temporal modeling. On this basis, the traffic of each grid is changing in real time, and the trend of variation is different. Therefore, the urban network can be regarded as a dynamic spatial-temporal system. It is the dynamic changes of traffic that the network optimization should be timely adjusted to ensure the user's experience. The purpose of this method is to forecast the traffic changes in each region and to provide reference for resource allocation and load balancing. Accurate and timely adjustment of a resource allocation strategy and network optimization requires high-accuracy prediction. The purpose of this paper is to improve the accuracy of traffic forecasting. A traffic forecasting system suitable for urban communication network is proposed. Moreover, the effectiveness (short-term prediction with low complexity) and the universality (multi-scenario of cities) of the algorithm are guaranteed.

To the authors' knowledge, it is the first time that causality analysis has been introduced into the analysis of communication data. In fact, causality is crucial for time series analysis. Causal analysis has always been an important part of economics, because economists are not concerned about how things happen together, but about how the relationship between things is transmitted. In the prediction of multivariate time series, how to select multiple time series is the most noticeable problem in data extraction. Correlation analysis is the most common method for selecting sequences. From the perspective of forecasting, correlation analysis and causality analysis are all possible, but from the perspective of data analysis, the results of causality analysis are more convincing and the results of multivariate time series forecasting will be more accurate through the combination with causality analysis. For example, there are two events, A and B. The results of correlation analysis of the two events are highly correlated. When A leads to B, we use B to fit the model based on previous correlation analysis to predict A. But when A has not happened, using B to predict A is not actually a prediction of the future, but an inference of known events. Causality analysis can make us know which is the cause and which is the result. We find that B cannot predict A well through causality analysis, but B can be predicted more accurately through A. Back to our previous question, for grid-based urban communication networks, changes in each region of the city are related to a certain extent [3]. So, the causality between grids in this system is very strong, which is also the reason why the author introduced causality analysis.

The Granger causality test is the most widely used method in causality analysis, and it has been applied to reveal interdependence structure in multi-variate time series [4]. The Granger causality test is also derived from economics, but it has been gradually used in other fields and has been popular in recent decades. Whether it is economics [5], meteorological science [6], or neuroscience [7], it has a wide range of applications. Moreover, the algorithm has also made some progress on the basis of the original [8].

According to the results of causality analysis, LSTM algorithm is used to predict multiple time series. LSTM algorithm is a special recurrent neural network (RNN) structure. It has long short-term memory (LSTM) units. RNN composed of LSTM units is usually called LSTM network [9]. The difference between LSTM and traditional RNN neural network is that each neuron in LSTM is a memory cell. LSTM links previous data information to the current neuron. Each neuron contains three gates: input gate, forgetting gate, and output gate, which are different from RNN. Using these internal gates, LSTM can solve the problem of long-term data dependence, especially for the prediction of data in a communication network, because the prediction of communication network requires high timeliness and low complexity. Timely fault handling is necessary for network optimization.

According to the characteristics of urban wireless communication network, the authors adopt multivariable prediction based on spatial-temporal model. Multivariable comes from the causality analysis of different grids in the spatial-temporal model. Through the causality test of the previous step, the time series with strong causality is extracted and used for multivariable LSTM prediction. Compared to the LSTM algorithm, multivariable LSTM is embodied in multiple input data, and the output data is the future value of the main prediction time series. Specific introduction of LSTM and its application in this paper will be given in the following chapters. Finally, by comparing the performance with other algorithms and non-causal analysis algorithms, the performance of the prediction system proposed by the author is verified. The algorithm is simulated in three scenarios in the city, which verifies the universal applicability of the algorithm in the urban communication network.

Through the combination of data processing of spatial-temporal modeling, causal analysis, and multivariate time series prediction algorithms, a prediction system for urban communication networks can be obtained. The main contributions of this paper are as follows:

1. A traffic prediction system suitable for urban communication networks is proposed, and it is universal for all urban scenes.

2. Causal analysis for CDR data is proposed for communication network analysis and to enhance the accuracy of multivariate time series prediction.
3. Multivariate time series LSTM prediction method based on spatial-temporal model of communication networks is proposed.

2 Urban communication network prediction system

The urban communication network prediction system proposed by the author is mainly divided into three modules: data processing module, causality analysis module, and LSTM prediction module. The overall framework of the system is shown in Fig. 1. In the data processing module, the spatial-temporal modeling is completed, and the corresponding time series are extracted according to the spatial information. In the causality analysis module and prediction module, these data are further analyzed and processed. The flow chart of the system is shown in Fig. 2. The specific content of each module will be introduced in the simulation section.

3 Algorithm simulation

3.1 Data processing

Nowadays, cellular network technology is the most widely used technology in the world. A communication system is composed of several base stations. Mobile devices that receive strong enough signals in the base station area are connected to the network and thus can be used for communication [10]. The design of the network determines the size of each community. The size of micro-cellular community in urban environment is
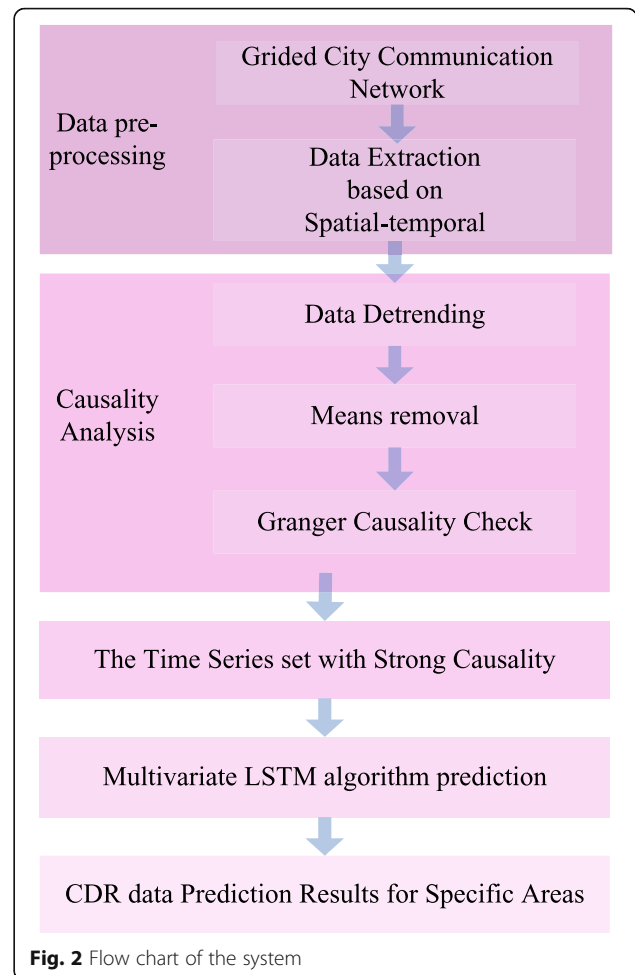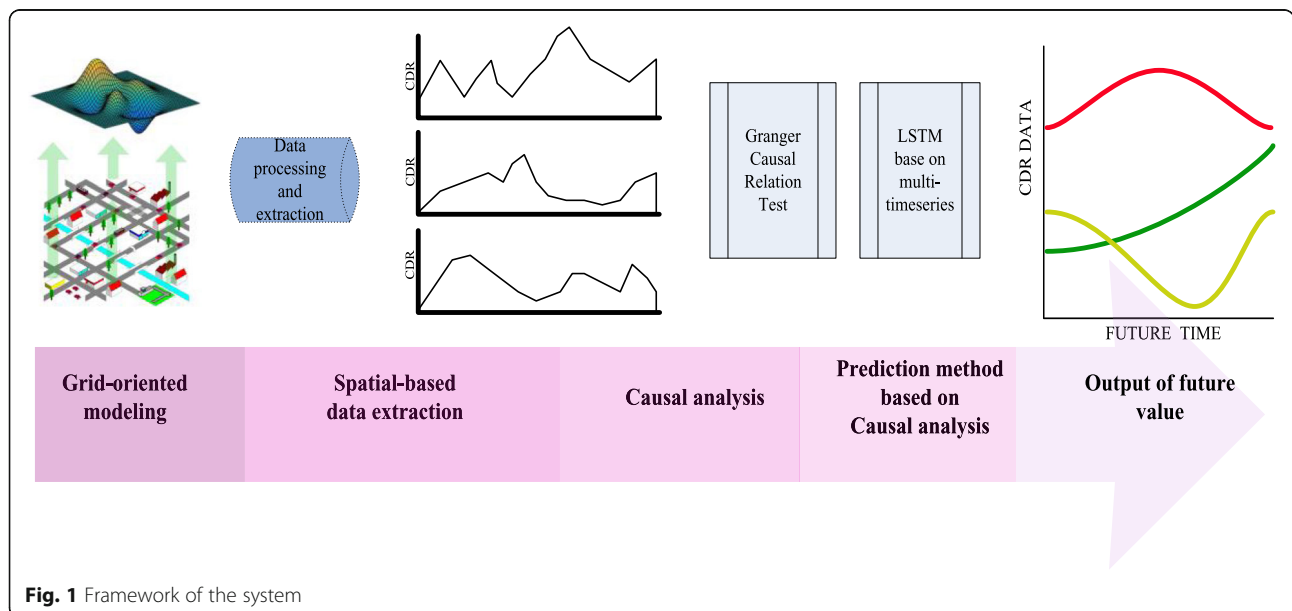

Fig. 2 Flow chart of the system

Fig. 1 Framework of the system

generally 300 m, and some macro-community in rural environment can reach 30 km [11]. All adjacent cells are overlapping, allowing a continuous connection to the network when the mobile equipment is moving. Many adjacent cells are grouped in zones identified by a local area code (LAC) [11]. Operators will keep detailed records of mobile devices in use. These records are called call detail records. CDR data generally includes time stamp, cell number, IMEI (International Mobile Equipment Identity), and time type. This information is highly correlated in spatial-temporal.

The data used in this paper is the CDR data provided by Telecom Italia. Traffic data in Milan on November 2, 2013, was adopted. The statistical time granularity is 10 min [12]. Geographical grids for data records are defined. The map of the city is divided into 100×100 grids. As shown in Fig. 3, each grid has a unique square ID covering an area of 235×235 square meters. The CDR data format is shown in Fig. 4 [12]. In this paper, we mainly analyze voice traffic data, so for each sampling point, we add the call-in and call-out data to constitute the current traffic.

In the study of communication networks, urban scenarios are often divided into four categories: office, station, entertainment, and residential areas [13–15]. In order to improve the universality of the algorithm, we analyze the entertainment area (Quadrilatro della moda) and the office area (Politecnico di Milano). In addition, traffic often increases dramatically in the scenario where major events occur.

Because our causality analysis is based on the causality between the main prediction area and the adjacent grid, our data extraction is centered on the prediction area, and traffic data is also extracted for the four adjacent regions. As shown in Fig. 5, the red area is the selected three scenes, and the surrounding blue grid are the areas for joint analysis.

According to the grid information, we get the corresponding traffic data. Figures 6 and 7 show the traffic changes at the Milan Polytechnic University, San Siro Stadium, and nearby grids on that day. The abscissa in the figure is the sampling time point, and the time granularity is 10 min. Different colors represent different regions. Among them, the blue-centered area is also the traffic change trend of the analyzed area.
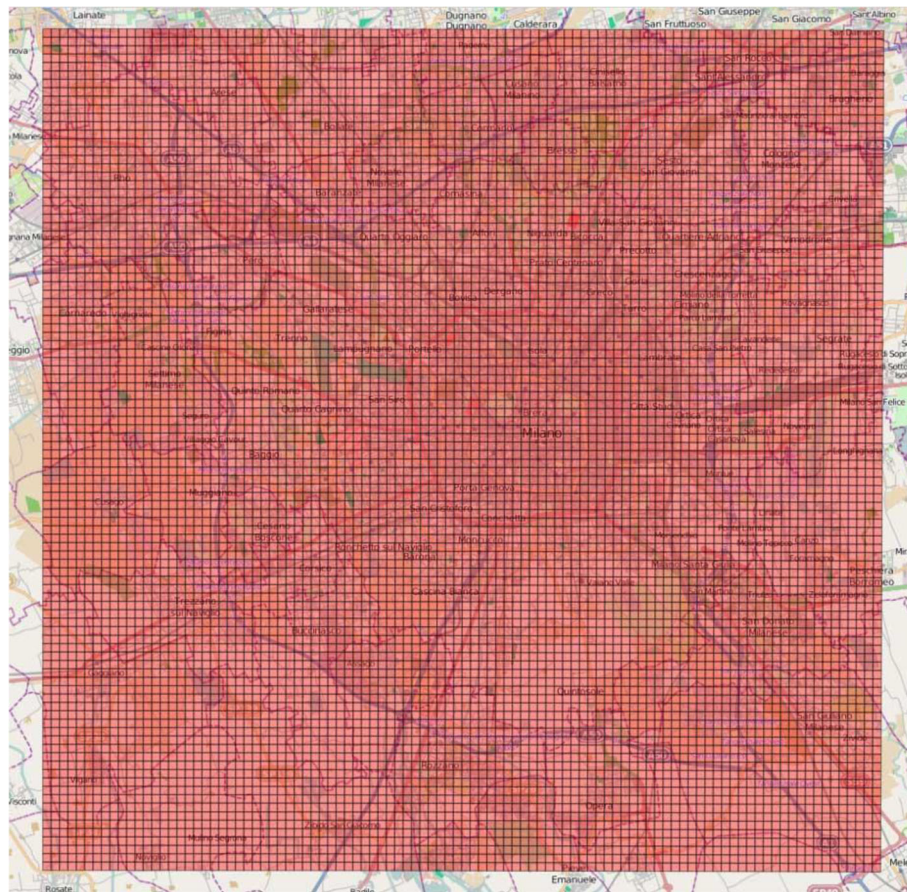

Fig. 3 Geographical grids for Milan

Time interval	Square id	SMS-in activity	SMS-out activity	Call-in activity	Call-out activity	Internet traffic activity
4/11/2013-0	6164	0.142298	0.181244	0.063178	0.041632	9.462406
4/11/2013-10	6164	0.112873	0.258088	0.091327	0.007971	8.652675

Fig. 4 CDR data format

3.2 Granger causality test

Causality can be defined by the dependence between variables, that is, the variable as the result is determined by the variable as the cause, and the change of the cause variable causes the change of the result variable. Granger points out that if one variable X is not helpful in predicting another variable Y, then X is not the cause of Y; on the contrary, if X is the cause of Y, two conditions must be satisfied [16]:

Firstly, X should be helpful in predicting Y, that is, in the regression of Y's past value, adding X's past value as an independent variable should significantly increase the explanatory power of regression.

Secondly, Y should not be helpful in predicting X. The reason is that if X is helpful in predicting Y and Y is helpful in predicting X, there may be one or more other variables, which are both the cause of X change and the cause of Y change.

Now, people generally call this causality defined from the perspective of prediction Granger causality. Specifically speaking, for the data we extracted, let us assume that X is traffic data in the central area, which is sampled at different time points $\{X_1, X_2, X_3, ...X_n\}$. Where n is the total number of training set samples. Time series Y is traffic data in one of the adjacent regions, from $\{Y_1, Y_2, Y_3, ...Y_n\}$. Now, we use the past of X to predict the future
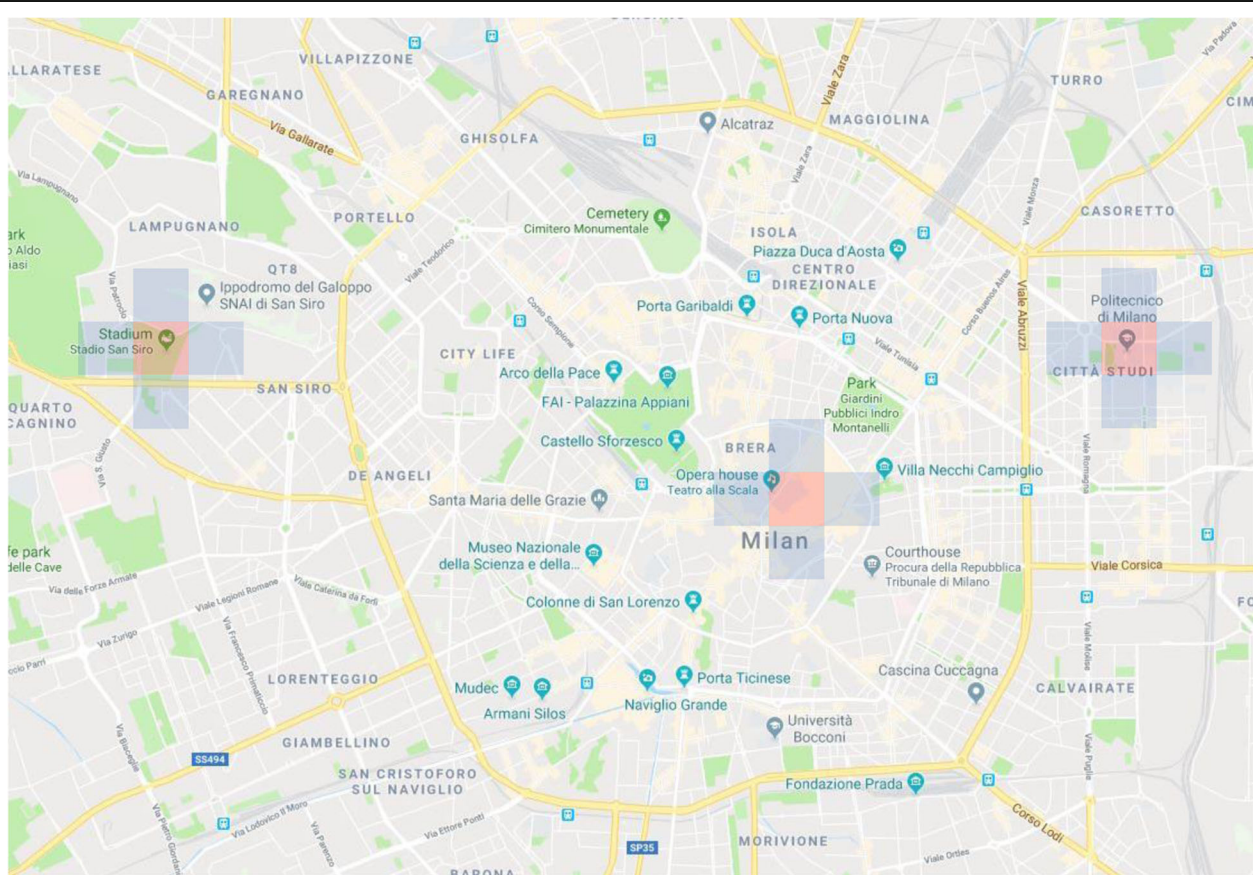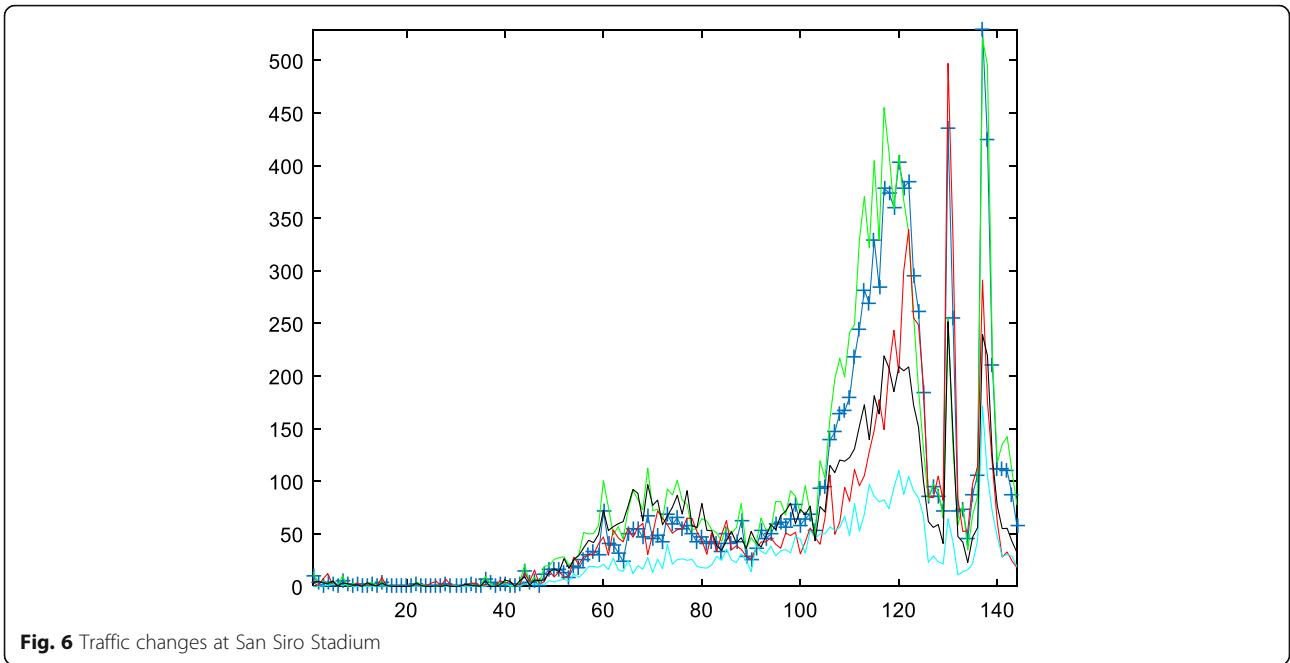

Fig. 5 Simulation areas

Fig. 6 Traffic changes at San Siro Stadium

of X. For example, we use $X_1\sim X_{n-j}$ (which is the past value of Y) to predict $X_{n-j+1}\sim X_n$ (which is the past value of X). In the process of prediction, we produce an error of δ_1 and then regard this error as the first result we get.

Then, we use the past of X and Y to predict the future of X, such as $\{X_1\sim X_{n-j}|\ Y_1\sim Y_{n-j}\}$ to predict $X_{n-j+1}\sim X_n$, and an error of δ_2 is generated in the process of prediction. If δ_1 is less than δ_2, that is to say, the combined prediction error of X and Y is less than the prediction

error of X itself, then it must be because Y is helpful to the prediction of X, so the prediction error is reduced. In this case, we call Y Granger cause to X [17].

So, whether variable X is the Granger cause of variable Y is testable. But some processing of data is needed before testing.

One of the basic conditions of Granger's test is the stability of the series, so before the causality test, we first ensure that these time series are stable, and the time series in the communication network are often complex

Fig. 7 Traffic changes at Milan Polytechnic University
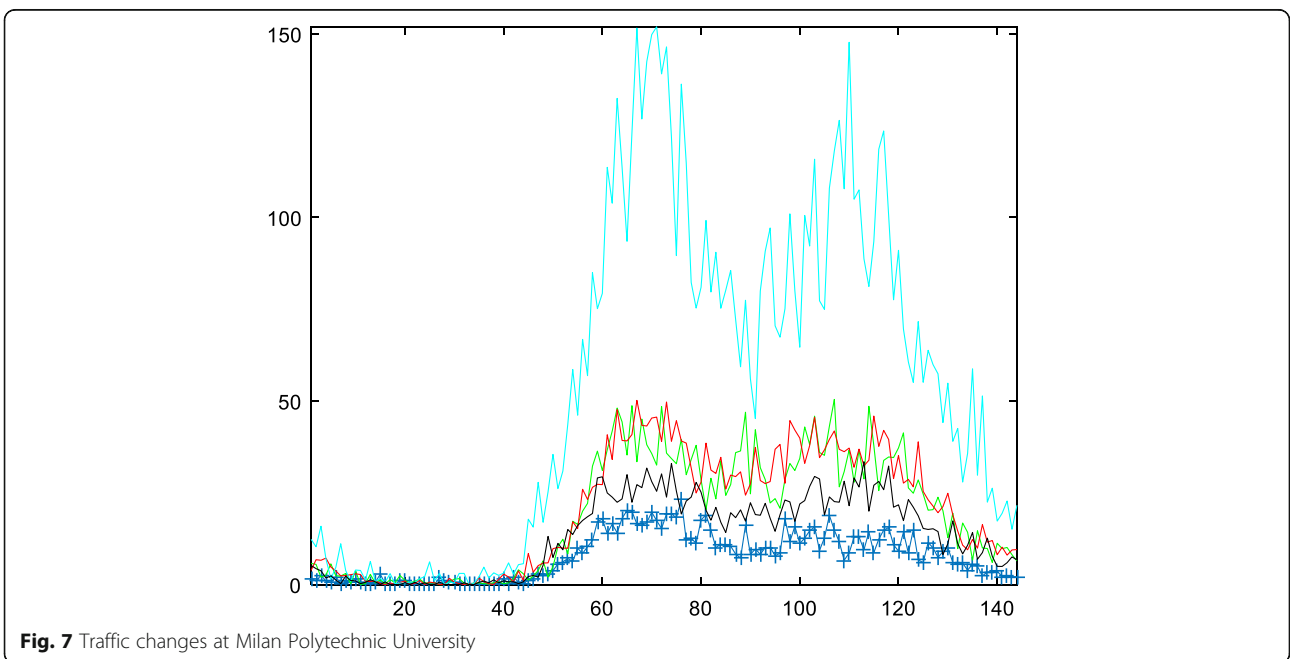
and non-stationary. This requires some pre-processing before verification. The arithmetic flow chart of the causal analysis module in this paper is shown in the Fig. 8 [18].

The first part is de-trending. The sequence with obvious trending is non-stationary, so we remove the trend from the sequence. Granger test requires data to fluctuate around the horizontal axis, so demeaned processing is needed. After processing in this part, take the data in stadium area as an example, the processed data is shown in the Fig. 9.

Subsequently, unit root test is done for time series data. If the conclusion is that the time series has unit root, then we can be sure that the sequence must not be stationary, and then carry out the subsequent differential processing. On the contrary, subsequent analysis is carried out through AIC criteria.

Our goal is to search lag in the range of $1 \sim n$ to minimize the value of AIC. And the lag that makes the AIC minimum is the order lag we want. The last step is the normal distribution test and consistency test [18]. Since error obeying normal distribution is a prerequisite for solving regression problems by least square method, the purpose of normal distribution test is to detect whether the residual after regression is obeying normal distribution. If it is not obeyed, the data does not satisfy

Fig. 8 Flowchart of the causal analysis

the precondition of using least squares method and the basis of solving Granger causality.

For consistency test, when the data points of time series are regressed. It is not possible to determine whether the theoretical and actual values obtained by regression come from the same distribution. At this time, consistency test should be adopted. If the conclusion of the consistency test shows that the gap between theoretical value and actual value is small, the regression results are good.

So far, we have completed all the processing steps before causality test. In fact, after such processing and analysis steps, we can do a Granger causality test which is somewhat demanding for data for a complex time series in urban communication network. In causality checking, we proceed step by step [19].

Step 1: Test the original hypothesis: X is not the Granger cause of Y. First, we estimate the following two regression models:

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{i=1}^{q} \beta_i X_{t-i} + \varepsilon_t \qquad (1)$$

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \varepsilon_t \qquad (2)$$

Among them, α_0 denotes the constant term, P and Q are the maximum lags of Y and X, respectively, ε_t is the white noise. Then F-statistics are constructed by sum of residual squares of two regression models which are RSS_u and RSS_r.

$$F = \frac{(RSS_r - RSS_u)\big/ q}{RSS_u \big/ (n-p-q-1)} \sim F(q, n-p-q-1) \qquad (3)$$

Among them, n is the number of samples. The original hypothesis can be tested by function 3. If $F \geq F_\alpha\ (q, n - p - q - 1)$, then $\beta_1, \beta_2, \dots \beta_q$ is significantly not 0. We should reject the hypothesis that X is not the Granger cause of Y; on the contrary, we cannot reject this hypothesis.

Step 2: Exchange the positions of Y and X, and test the original hypothesis in the same way: "Y is not the Granger cause of X change".

Step 3: To reach the conclusion that "X is the Granger cause of Y," we must reject the original hypothesis that "X is not the Granger cause of Y" and accept the original hypothesis that "Y is not the Granger cause of X."

We use this method to analyze the traffic sequence of each region. In this paper, the data are divided into test set and training set. The training data are traffic data from 0:00 to 18:00, and the data from 18:00 to 24:00 are
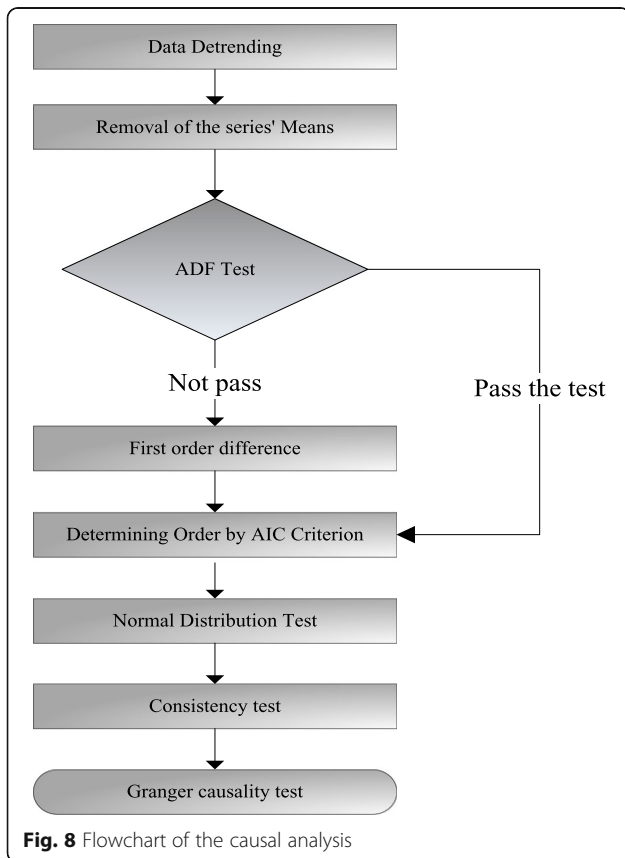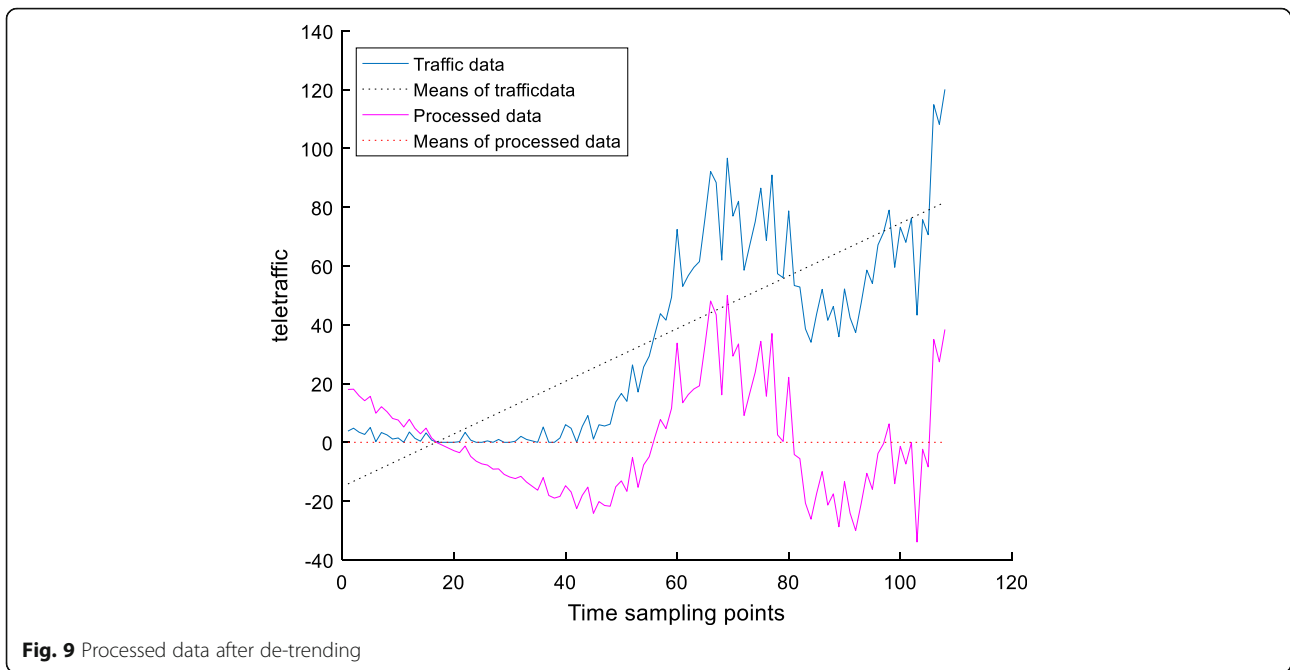
Fig. 9 Processed data after de-trending

test data. According to the spatial distribution, multiple time series can be obtained. Causality checking is carried out among multiple time series to obtain the causality between traffic in adjacent areas and traffic changes in central areas. It is equivalent to finding out the "cause" of traffic change in the main forecast area and combining the traffic data of the main forecast trend with these causal data to analyze and forecast the future value of the main area.

Fig. 10 Causality topological diagram of San Siro Stadium
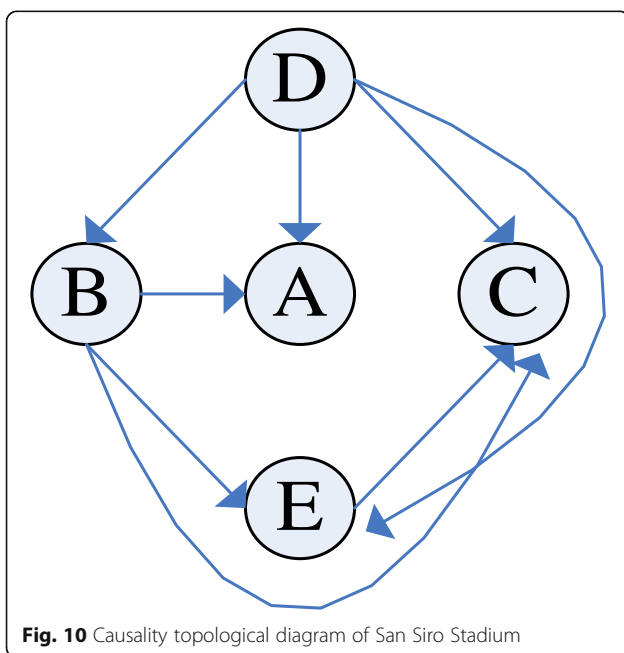
After causality checking, causality diagrams showed in Fig. 10 can be obtained. Causality diagram is a topological diagram of causality [20]. Each vertex represents a different area. In the diagram, the causal diagram is based on the San Siro Stadium (a major event scenario). A is the central area, that is, the stadium. It is also the area where we will ultimately predict the traffic. The arrow represents the impact on the data of an area. As can be seen from the Fig. 10, it is the data of B and D regions that affect the data of central area A. Therefore, the data from these three regions are used for subsequent multivariate time series prediction which will be described in detail in the next chapter.

3.3 LSTM prediction algorithms for multivariate time series

According to the multivariate data obtained from causal analysis, the traditional linear model is difficult to solve the multivariate or multi-input problem, while the neural network such as LSTM is good at dealing with the problem of multiple variables, which makes it helpful to solve the problem of time series prediction. So multivariate LSTM algorithm is used for prediction.

LSTM is a special form of RNN, which is widely used in time series analysis, especially for multivariate time series analysis. The difference between LSTM and RNN is that every neuron in LSTM is a memory cell. LSTM stores previous information in current neurons. As
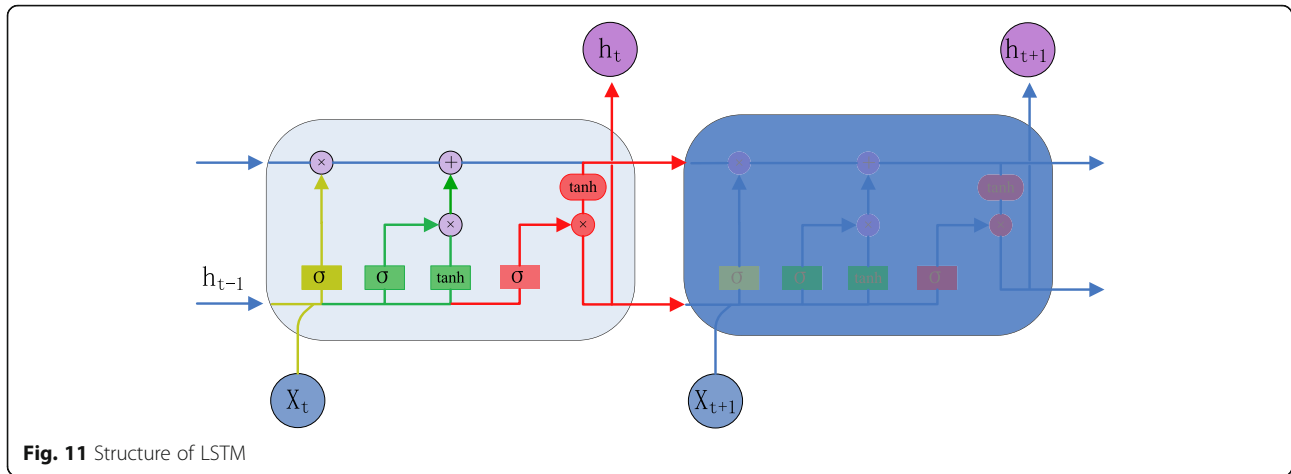
Fig. 11 Structure of LSTM

shown in Fig. 11, each neuron contains three gates: input gate, output gate, and forgetting gate. Through these internal doors, the long-term dependency problem can be solved [9, 21].

The yellow part of the picture is the forgetting gate (Eq. 4). The first step in LSTM is to decide what information we will discard from the cell state. This decision is made through a layer called the forgetting gate. The gate reads h_{t-1} and X_t and outputs a value between 0 and 1 to each number in the cell state (Eq. 5). 1 means "complete reservation" and 0 means "complete abandonment". σ denotes sigmoid function.

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, X_t] + b_f\big) \qquad (4)$$

$$S(t) = \frac{1}{1 + e^{-t}} \qquad (5)$$

The green part is the input gate, which determines how much new information is added to the cell state. There are two steps to achieve this: first, a sigmoid layer called "input gate layer" determines which information needs to be updated (Eq. 6) and second, a tanh layer generates a vector, which is the alternative content for updating. Cell status is updated by combining these two parts (Eq. 7).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \qquad (6)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \qquad (7)$$

The red part is the output gate. First, we run a sigmoid layer to determine which part of the cell's state will be exported (Eq. 8). Next, we process the cell state through tanh (to get a value between − 1 and 1) and multiply it with the output of the sigmoid gate (Eq. 9). Eventually, we only output the part of the output we determined [22].

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \qquad (8)$$

$$h_t = o_t * \tanh(C_t) \qquad (9)$$

The author fits our LSTM model with keras, a deep learning library [23]. Take the stadium scene data as an example. The first hidden layer defines a LSTM with 50 neurons and a neuron in the output layer used to predict contamination. The time step of input data is 1, which has three characteristics, namely, three input variables, the data of the central region and the data of the two regions which cause the change of the data of the central region after causal analysis. The output variable is the prediction result of traffic volume in the central area. The data are divided into test set and training set as mentioned above.

4 Results and discussion

In order to fully verify the performance of the algorithm, the author compares two other prediction methods, one is correlation-based multivariate LSTM. In this model, the data of different grids in the same scene are correlated with the data of the central area through Pearson correlation coefficient calculation [24]. After that, the highly correlated data are used for subsequent multivariate time series prediction. Another method is the VARIMA (vector autoregressive integrated moving average) model [25], which is often used for prediction. When the VARIMA algorithm is applied, all grid data in the same region are input data, and the output is the prediction of the future value of the central grid data of the region. Both the two multivariate time series prediction algorithms are considered as good choices to improve prediction accuracy.

As can be seen from Figs. 12, 13, and 14, the LSTM prediction algorithm based on causal analysis proposed by the author has higher accuracy than other algorithms.
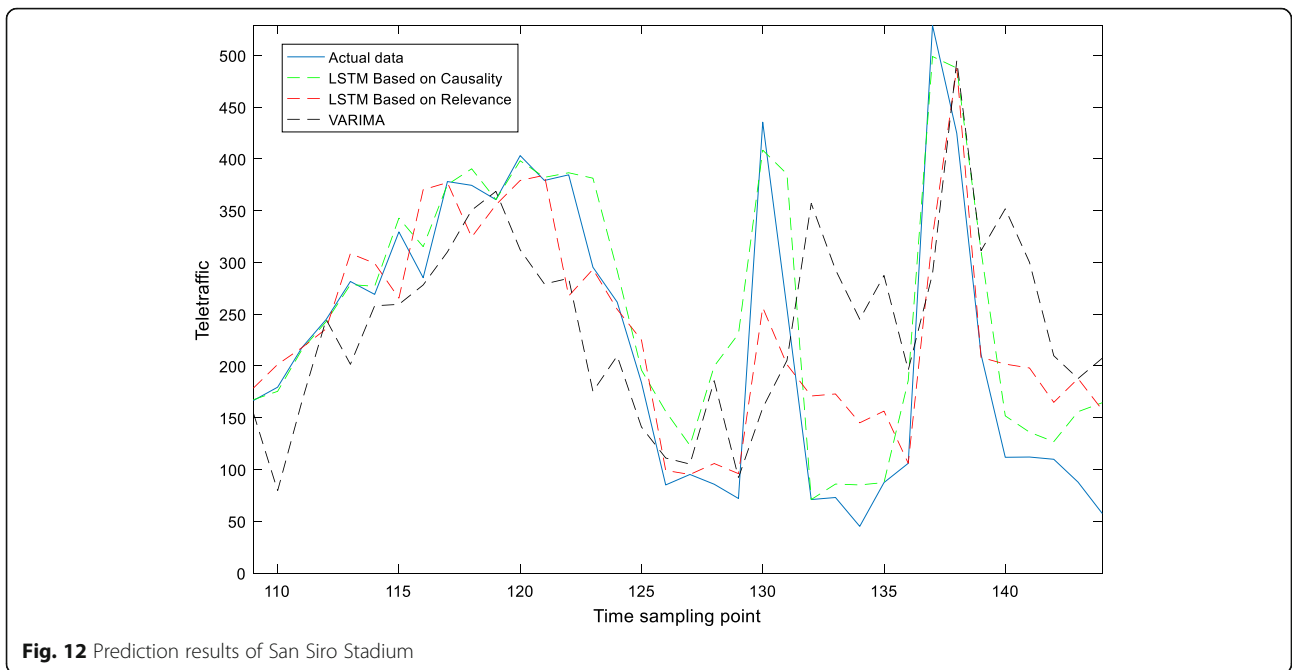
Fig. 12 Prediction results of San Siro Stadium

In order to make the perception performance more intuitive, the author calculated RMSE and MAE in three different scenarios using three prediction methods.

Since the values of RMSE and MAE depend on the magnitude of the test data and the traffic gap between different areas in urban communication network is very large, the values of RMSE and MAE in different areas of the table vary greatly, but from Table 1, the accuracy of the proposed algorithm is higher in three areas. In addition, it can be seen from the table that the use of causality method is obviously more accurate for stadium scenes, that is, the condition of the sudden increase in traffic. The VARIMA algorithm seems not suitable for forecasting, because this simulation belongs to short-term forecasting, which is not the advantage of VARIMA. As mentioned above, the prediction algorithm with high accuracy and universality is very important for urban
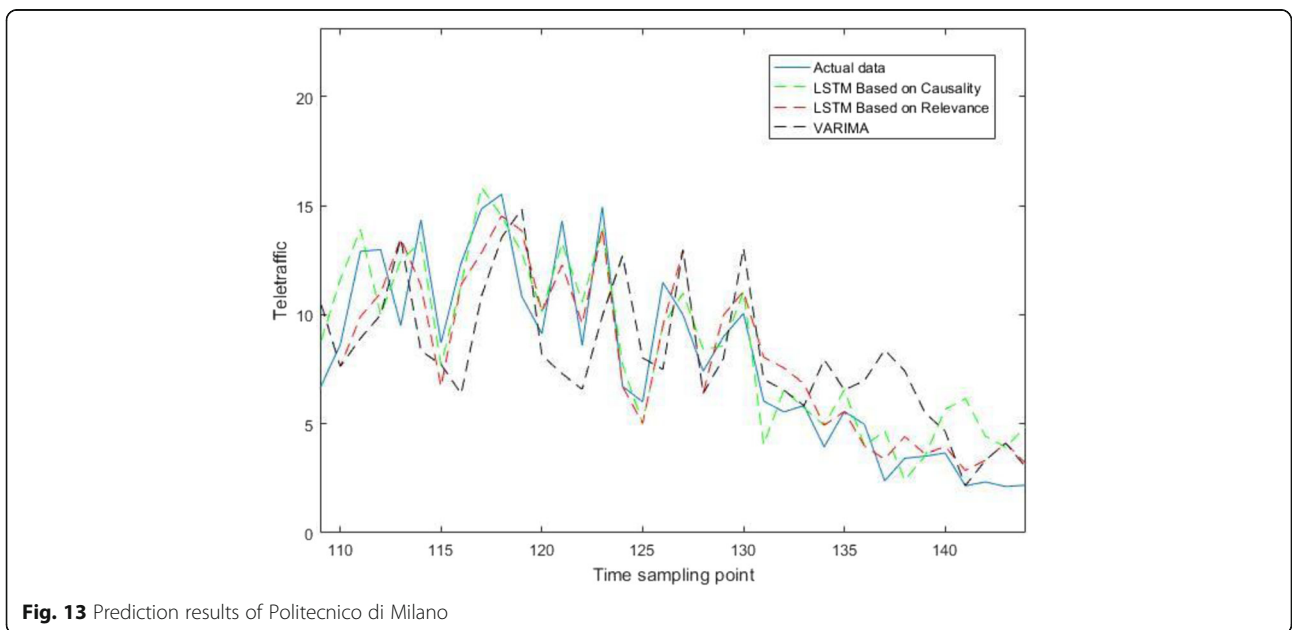

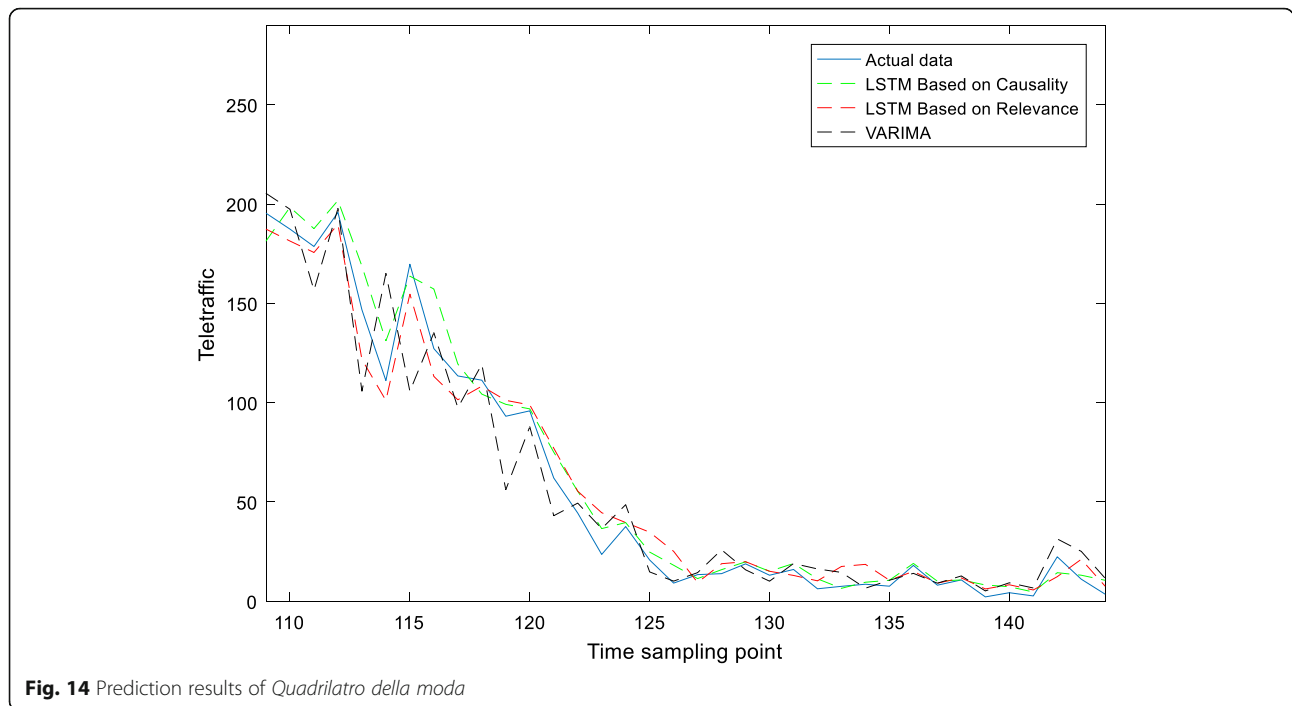**Fig. 13** Prediction results of Politecnico di Milano

Fig. 14 Prediction results of *Quadrilatro della moda*

communication network. The algorithm proposed by the author can satisfy this condition and be used in network maintenance and optimization.

5 Conclusion

Urban wireless communication network is a complex system. To allocate resources reasonably is the key to network optimization. Accurate prediction algorithm is the basis of network optimization. In this paper, the author transformed the problem of network analysis into a time series model through the method of big data analysis and made quantitative analysis of urban communication network while retaining the spatial-temporal characteristics. Firstly, spatial-temporal modeling was introduced through rasterization, and then data extraction was carried out according to this model. In the following data analysis, the Granger causality test was introduced for the first time. Then, multivariate LSTM algorithm was used to predict the traffic. From the simulation

results, the addition of causality can improve the accuracy of multivariate time series prediction, and this prediction system is generally applicable to communication networks. Causality analysis can be used to analyze the dynamic variation of urban wireless communication network more effectively, and more accurate prediction results can be obtained, which will considerably assist to the maintenance and management of the network. As we all know, the most important concern of MNOs (mobile network operators) is the real-time changing of network state. In this paper, the proposed analysis and prediction system for urban communication network can not only make operators better understand the real-time changing of urban network, but also help operators manage the network primely. The author's future work will focus on the specific parameters of the network, which have a large amount of wireless data and complex data types. But only in this case can each user's experience be better concerned in real-time. Hence, it will be the focus in further studies.

Table 1 Error comparison of each algorithm

Error calculation by	RMSE			MAE		
Algorithm in different areas	Quadrilatro Della Moda (office area)	Politecnico di Milano (entertainment area)	Stadio San Siro (stadium with activities)	Quadrilatro Della Moda (office area)	Politecnico di Milano (entertainment area)	Stadio San Siro (stadium with activities)
LSTM based on causality	1.7491	9.3764	56.0866	1.4836	6.6944	36.9817
LSTM based on relevance	1.8073	9.5862	72.1137	1.5028	7.6139	51.7222
VARIMA	3.3652	18.7506	128.3342	2.7729	11.9167	72.1137

Abbreviations

Acknowledgments

Funding

Availability of data and materials

The datasets supporting the conclusions of this article were collected from reference [12].

Authors' contributions

WG conceived and designed the study. KZ and SM performed the simulation experiments. KZ and XL wrote the paper. GC and ZS reviewed and edited the manuscript. All authors read and approved the final manuscript.

Authors' information

Kaisa Zhang received the B.Sc. degree in communication engineering from Hebei University of Science and Technology in 2014. He received MEng degree in communication engineering from the University of York in 2016. He is currently working towards the Ph.D. degree in information and communications engineering, Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications Beijing, China; His main studying area is big data analysis and processing, self-optimizing and machine learning in communication network.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

References

1. Z. Wang, S. Zhang, in *2016 IEEE 14th Intl Conf on Dependable, Autonomic & Secure Computing, 14th Intl Conf on Pervasive Intelligence & Computing, 2nd Intl Conf on Big Data Intelligence & Computing & Cyber Science & Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*. CDR based temporal-spatial analysis of anomalous mobile users. p. 710–714 (2016). https://ieeexplore.ieee.org/document/7588924.
2. S.C. Nair, M.S. Elayidom, S. Gopalan, 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS). *Impact of CDR Data Analysis Using Big Data Technologies for the Public: an Analysis*. p. 1–6 (2017). https://ieeexplore.ieee.org/document/8014651.
3. S. Jiang, J. Ferreira, M.C. Gonzalez, Activity-based human mobility patterns inferred from mobile phone data: a case study of Singapore. IEEE Trans. Big Data **3**(2), 208–219 (2017).
4. E. Siggiridou, D. Kugiumtzis, Granger causality in multivariate time series using a time-ordered restricted vector autoregressive model[J]. IEEE Trans. Signal Process. **64**(7), 1759–1773 (2016).
5. C. Hiemstra, J. Jones, Testing for linear and nonlinear Granger causality in the stock price-volume relation. J. Financ. **49**(5), 1639 (1994).
6. A.C. Lozano, H. Li, in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France*. Spatial-temporal causal modeling for climate change attribution. (2009). https://www.researchgate.net/publication/221654446_Spatial-temporal_causal_modeling_for_climate_change_attribution.
7. A. Roebroeck, E. Formisano, R. Goebel, Mapping directed influence over the brain using Granger causality and fMRI. Neuroimage **25**(1), 230–242 (2005).
8. Y. Chen, G. Rangarajan, J. Feng, M. Ding, Analyzing multiple nonlinear time series with extended Granger causality. Phys. Lett. A. **324**(1), 26–35 (2004).
9. Y.T. Tsai, Y.R. Zeng, Y.S. Chang, in *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*. Air pollution forecasting using RNN with LSTM. P. 1074–1079 (2018). https://ieeexplore.ieee.org/document/8512020.
10. C.W. Huang, C.T. Chiang, Q. Li, in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. A study of deep learning networks on mobile traffic forecasting. p. 1–6 (2017). https://ieeexplore.ieee.org/document/8292737.
11. E. Thuillier, L. Moalic, S. Lamrous, A. Caminada, Clustering weekly patterns of human mobility through mobile phone data. IEEE Trans. Mob. Comput. **17**(4), 817–830 (2018).
12. (2014). [Online]. Available: https://dandelion.eu. Accessed Nov 2018.
13. J. Reades, F. Calabrese, A. Sevtsuk, C. Ratti, Cellular census: explorations in urban data collection. IEEE Pervasive Comput. **6**(3), 30–38 (2007).
14. M. Panda, S.P. Padhy, Traffic analysis and optimization of gsm network. IJSCI Int J Comput Sci Issues **1**, 28–31 (2011).
15. X. Zuo, Y. Zhang, Detection and analysis of urban area hotspots based on cell phone traffic. JCP **7**(7), 1753–1760 (2012).
16. A. Seth, Granger causality. Scholarpedia **2**(7), 1667 (2007).
17. C.W.J. Granger, Econometrica **37**, 424 (1969).
18. F. Xue, G. Fang, X. Yue, E. Zhao, S. Brauth, Y. Tang, Resting-state brain networks revealed by Granger causal connectivity in frogs. Neuroscience, **334**, 332–340 (2016).
19. M. Dhamala, G. Rangarajan, M. Ding, Estimating Granger causality from Fourier and wavelet transforms of time series data. Phys. Rev. Lett. **100**(1), 018701 (2008).
20. D. Yang, H. Chen, Y. Song, Z. Gong, in *2017 IEEE International Conference on Big Knowledge (ICBK)*. Granger causality for multivariate time series classification. p. 103–110 (2017). https://ieeexplore.ieee.org/document/8023401/.
21. Z. Cao, Y. Zhu, Z. Sun, M. Wang, Y. Zheng, P. Xiong, J. Hou, L. Tian, *Improving Prediction Accuracy in LSTM Network Model for Aircraft Testing Flight Data*. p. 7–12. (IEEE International Conference on Smart Cloud (SmartCloud), 2018). https://ieeexplore.ieee.org/document/8513708.
22. Y. Gu, S. Liu, L. He, L. Wang, *Research on Failure Prediction Using DBN and LSTM Neural Network. 2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)* (2018).
23. P. Vidnerová, R. Neruda, in *Federated Conference Computer Science Information Systems*. Evolving KERAS architectures for sensor data analysis (2017).
24. C. Thirumalai, C. Raju, G. Saikrishna, M. Senthilkumar, in *International Conference on Trends in Electronics and Informatics ICEI*. Analysis of global warming in India over maximum temperature using Pearson and Machine learning (2017).
25. R.S. Tsay, *Multivariate Time Series Analysis and Its Applications[M]// Analysis of Financial Time Series, Second Edition* (2000).