Article

Title

Firstname Lastname [1,†,‡], Firstname Lastname [1,‡] and Firstname Lastname [2,*]

[1] Affiliation 1; e-mail@e-mail.com
[2] Affiliation 2; e-mail@e-mail.com
* Correspondence: e-mail@e-mail.com; Tel.: (optional; include country code; if there are multiple corresponding authors, add author initials) +xx-xxxx-xxx-xxxx (F.L.)
† Current address: Affiliation 3
‡ These authors contributed equally to this work.

Version December 27, 2020 submitted to Journal Not Specified

Abstract: A single paragraph of about 200 words maximum. For research articles, abstracts should give a pertinent overview of the work. We strongly encourage authors to use the following style of structured abstracts, but without headings: (1) Background: Place the question addressed in a broad context and highlight the purpose of the study; (2) Methods: Describe briefly the main methods or treatments applied; (3) Results: Summarize the article's main findings; and (4) Conclusion: Indicate the main conclusions or interpretations. The abstract should be an objective representation of the article, it must not contain results which are not presented and substantiated in the main text and should not exaggerate the main conclusions.

Keywords: keyword 1; keyword 2; keyword 3

1. Introduction

The introduction should briefly place the study in a broad context and highlight why it is important. It should define the purpose of the work and its significance. The current state of the research field should be reviewed carefully and key publications cited. Please highlight controversial and diverging hypotheses when necessary. Finally, briefly mention the main aim of the work and highlight the principal conclusions. As far as possible, please keep the introduction comprehensible to scientists outside your particular field of research. Citing a journal paper [**?**]. And now citing a book reference [**?**]. Please use the command [**?**] for the following MDPI journals, which use author-date citation: Administrative Sciences, Arts, Econometrics, Economies, Genealogy, Humanities, IJFS, JRFM, Languages, Laws, Religions, Risks, Social Sciences.

2. Main result

In airlines, the spatial distribution of aircraft stations is a non-Euclidean structure, that is, the number of stations around each station is uncertain, and even if two stations are adjacent, they may not actually communicate with each other, resulting in no spatial relationship between their traffic. Therefore, traditional convolutional neural network (CNN) cannot accurately obtain their spatial information. At this point, multiple sites can be abstracted into A graph (see Figure **??**). Features are extracted from the original input data to obtain the result of feature mapping of multiple channels. The intercommunication relationship between each site is represented by adjacency matrix A.

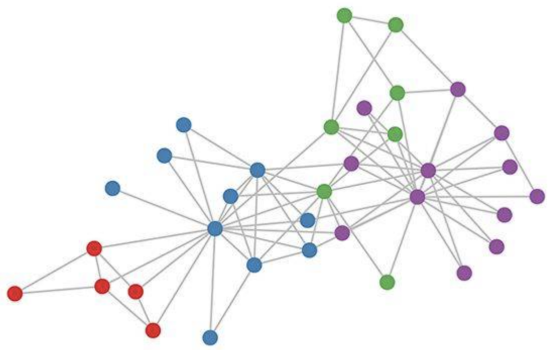

Figure 1. The distribution of the stations.

2.1. The definition of the problem

The problem of airline passenger flow prediction can be described as follows: the historical flow data of each station $X_{t-s}, X_{t-s+1}, X_{t-2}, \cdots, X_{t-1}$ (s is the time step) can be used to predict the flow X_t of the next period. The formula is described as

$$X_t = F([X_{t-s}, X_{t-s+1}, \cdots, X_{t-2}, X_{t-1}]) \tag{1}$$

where, X is the site characteristics at each time step, and F is a nonlinear function.

In the actual traffic system, the road network is regarded as A directed graph $G = (q, V, A)$. Each sensor in the road network is regarded as a node vi V and its value Q R is a scalar. V RN and N is the number of sensors. The flow relationship between nodes consists of adjacency matrix A that is, the element Aij in A represents the connection relationship between node VI and vJ.

2.2. The description of the GCN

When dealing with the structure of the graph, it is necessary to obtain its Laplace matrix L first, which is generally defined in the following ways:

$$L = D^{-\frac{1}{2}}(D - A)D^{-\frac{1}{2}} = I_N - D^{-\frac{1}{2}}AD^{-\frac{1}{2}} \tag{2}$$

Where, $I_N$is the identity matrix of NN; Degree matrix D is defined as $D_{ii} = \sum_i A^i_{ij}$. Decompose the eigenvalue of L to get $L = \mathrm{L} = \mathrm{U}\Lambda\mathrm{U}^{\mathrm{T}}$. Λ is made up of L eigenvalues of diagonal matrix. $\mathrm{U} = \{\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_\mathrm{N}\}$ is composed of the eigenvector L, and it is an orthonormal basis for R^N.

The spectral convolution theory in the graph structure has been supplemented and perfected in the paper. The convolution operation of convolution kernel G and input signal X in the time domain can be converted into the inner product form in the frequency domain.

$$g^* x = U\left(U^{\mathrm{T}} g\right) \odot \left(U^{\mathrm{T}} \boldsymbol{x}\right)\right) = U_{g_\theta}(A) U^{\mathrm{T}} x \tag{3}$$

where $g_s(\Lambda) = U^T g = \mathrm{diag}(\theta), \theta \in R^N$, $\odot$ represents the hadamar product, $U^T g$means mapping g to the frequency domain space based on U. Due to g_θ high computational complexity, so using hierarchical linear model constraints and the chebyshev polynomial to approximate calculation. In this paper, The simplified first order polynomial form of $g * x$ is adopted.

$$g^* x = U_{g_\theta} U_x^{\mathrm{T}} \approx \theta\left(I_N + D^{-1/2} A D^{-1/2}\right) x \tag{4}$$

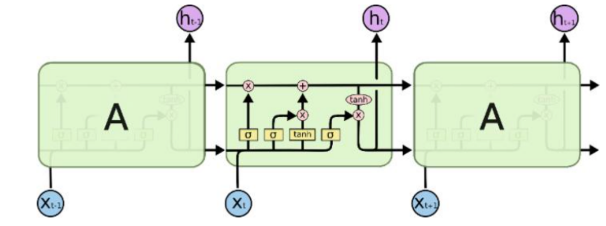

Figure 2. The structure of LSTM.

where $\widetilde{D}^{-1/2}\widetilde{A}\widetilde{D}^{-1/2} = I_N + D^{-1/2}AD^{-1/2} \quad \tilde{A} = I_N + A, D = \sum_i \widetilde{A}_{ij}$, Therefore, the output of layer L is

$$H^{(l)} = \sigma\left(\widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}H^{(l-1)}W^{(l)}\right. \tag{5}$$

where δ is the activation function, $\widetilde{W}^{(l)} = \theta^{(l-1)}W^{(l)} \quad \boldsymbol{\theta}^{(l-1)} \in \boldsymbol{R}^{c^{(1-l)}\times F^{(l-1)}}, \boldsymbol{W}^{(l)} \in \boldsymbol{R}^{F^{(l-1)}\times c^{(l)}}, C^{(L-1)}$ is the output dimension of the $(L-1)$ layer, and $F^{(L-1)}$ is the characteristic vector size of each dimension. Therefore

$$H^{(l)} = \sigma\left(\widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}}H^{(l-1)}\widetilde{W}^{(l)}\right) \tag{6}$$

At present, there is no effective measurement method for the calculation of adjacency matrix A. Most scholars use heuristic methods, that is, based on the Euclidean distance or Markov distance between sensors to determine the element value corresponding to the adjacency matrix. However, these methods all require manual calculation of the distance relationship between the sensors in advance. In this paper, the data-driven method is adopted to calculate the adjacency matrix, and $A = D$, then the formula can be written as:

$$H^{(l)} = \sigma / \widetilde{A}H^{(l-1)}\widetilde{W}^{(l)}\Big) \tag{7}$$

The element value of matrix A is learned from the sample data, that is, the matrix is composed of trainable parameters. The data-driven approach is more realistic than the heuristic approach. Therefore, the L layer of the convolutional neural network is constructed in accordance with Formula **??**. It should be noted that the initial A maki is the same for each layer of the convolutional network, and the parameters are updated only when the error is propagated backwards.

2.3. the import of LSTM

It is found that Recurrent Neural networks (RNN) are widely used in sequential data such as natural language and image processing, which have a significant effect. Since then, various types of circulating neural networks have been proposed. Aiming at the problem of air passenger flow prediction, this paper introduces the long and short term memory network, which can extract the characteristic information of the input sequence and find its internal relation, so as to improve the prediction accuracy of the model.

The LSTM network structure used in this paper is shown as follows The network model mainly accepts three inputs: X, H and C represent the current state, hidden layer state and cell state, respectively.

LSTM mainly realizes the management of long and short term memory through three gating units. The first step in the LSTM is to determine what information needs to be thrown out of the cell state. This decision is made by a sigmoID layer called the "Forget Gate." Input x and h, output a number between 0 and 1. 1 means "keep the value completely", while 0 means "throw the value away completely". The formula of forgetting gate is as follows:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{8}$$

where W_f and b_f are the parameters to be learned, and σ is the sigmod activation function.

The second step is to determine what information we need to store in the Cell State. There are two parts to this question. First, a sigmod layer calls the "Input Gate" to determine which data needs to be updated. Then, a tanh layer creates a vector C1 for the new candidate values, which can be added to the state.

$$\begin{aligned} i_t &= \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \\ \tilde{C}_t &= \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \end{aligned} \tag{9}$$

After deciding what needs to be forgotten and what needs to be added, the old cell state Ct-1 can be updated to the new cell state Ct.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{10}$$

Finally, we need to decide what to export. This output is based on our cell state, but will be a filtered part of the value. First, we run an output gate to determine which part of the cell state we are going to output. Then we put the cell state into the tanh (pressing the value between -1 and 1), and finally we multiply it by the output of the output gate.

$$\begin{aligned} o_t &= \sigma\left(W_o\left[h_{t-1}, x_t\right] + b_o\right) \\ h_t &= o_t * \tanh\left(C_t\right) \end{aligned} \tag{11}$$

2.4. The description of algorithm

The network structure based on GCN-LSTM model proposed in this paper is shown in the figure. The prediction model is composed of multiple layers of GCN and an LSTM model. Among them, GCN model is mainly used to extract the features of graph networks, while LSTM network model is mainly used to solve the long-term and short-term dependence between data.

In this paper, features of different graph networks are extracted through multiple GCN, and the extracted features are transferred into LSTM, and the timing characteristics of sequence data are analyzed through LSTM, and a more accurate prediction value is given through multi-layer full connection layer.

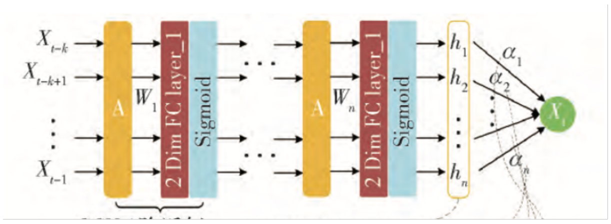

Figure 3. Overall structure of GCN-LSTM model (original features of air passenger flow data are extracted by using first-order approximate GCN, and output features are analyzed by LSTM for long-term and short-term sequence characteristics, and then predicted values are obtained).

3. Patents

This section is not mandatory, but may be added if there are patents resulting from the work reported in this manuscript.

Author Contributions: For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used "Conceptualization, X.X. and Y.Y.; methodology, X.X.; software, X.X.; validation, X.X., Y.Y. and Z.Z.; formal analysis, X.X.; investigation, X.X.; resources, X.X.; data curation, X.X.; writing–original draft preparation, X.X.; writing–review and editing, X.X.; visualization, X.X.; supervision, X.X.; project administration, X.X.; funding acquisition, Y.Y. All authors have read and agreed to the published version of the manuscript.", please turn to the CRediT taxonomy for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

Funding: Please add: "This research received no external funding" or "This research was funded by NAME OF FUNDER grant number XXX." and and "The APC was funded by XXX". Check carefully that the details given are accurate and use the standard spelling of funding agency names at https://search.crossref.org/funding, any errors may affect your future funding.

Acknowledgments: In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

Conflicts of Interest: Declare conflicts of interest or state "The authors declare no conflict of interest." Authors must identify and declare any personal circumstances or interest that may be perceived as inappropriately influencing the representation or interpretation of reported research results. Any role of the funders in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript, or in the decision to publish the results must be declared in this section. If there is no role, please state "The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results".

Abbreviations

The following abbreviations are used in this manuscript:

MDPI	Multidisciplinary Digital Publishing Institute
DOAJ	Directory of open access journals
TLA	Three letter acronym
LD	linear dichroism

Appendix A

Appendix A.1

The appendix is an optional section that can contain details and data supplemental to the main text. For example, explanations of experimental details that would disrupt the flow of the main text, but nonetheless remain crucial to understanding and reproducing the research shown; figures of replicates for experiments of which representative data is shown in the main text can be added here if brief, or as Supplementary data. Mathematical proofs of results not central to the paper can be added as an appendix.

Appendix B

All appendix sections must be cited in the main text. In the appendixes, Figures, Tables, etc. should be labeled starting with 'A', e.g., Figure A1, Figure A2, etc.

Sample Availability: Samples of the compounds ...... are available from the authors.