Study on Delay Propagation Relations Among Airports Based on Transfer Entropy

YINHONG XIAO¹,²,³, YAOSHUAI ZHAO²,³, GE WU²,³, AND YIZHEN JING¹,³
¹School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China
²Travelsky Technology Limited, Beijing 101318, China
³Key Laboratory of Intelligent Passenger Service of Civil Aviation, CAAC, Beijing 101318, China
Corresponding author: Yizhen Jing (yzjing@bjtu.edu.cn)

ABSTRACT To understand the mechanism of delay propagation from the perspective of multiple airports, constructing delay propagation relation (DPR) networks among airports is a novel analysis method. The latest method is to use transfer entropy to mine the delay propagation relation among airports. However, the transfer entropy will produce bias due to estimating high-dimensional conditional mutual information (CMI) in the delay propagation scenario. In this paper, we propose the low-dimensional approximation of CMI for transfer entropy (LTE) to address the above issue. By applying this improved algorithm, the delay propagation relation among airports can be more accurately explored and a more accurate DPR network can be obtained. For this network, we apply the complex network theory and its related indicators to provide systematic analysis about delay propagation. The results of case analysis show that in the delay propagation interactions among airports, large airports always receive delays and small airports propagate delays outward. Meanwhile, delays propagate more efficiently in the aviation system of smaller airlines. These results can provide some theoretical supports for making measures to reduce delay propagation.

INDEX TERMS Delay propagation relation, transfer entropy, flight delay, causality analysis, complex network.

I. INTRODUCTION

With the rapid development of the civil aviation industry, the volume of airline business grows at a high speed, and the problem of flight delay has become increasingly prominent. When upstream flights are delayed, downstream flights that share resources, such as aircraft or crews, with them can also be delayed with a high probability [1], [2]. If there is no reasonable way to understand the mechanism of delay propagation, the impact of delays cannot be handled properly, and will not only increase the operating costs of airlines and airports but also endanger the safety of passengers, aircrafts and, airports [3], [4]. Therefore, the study of flight delay propagation is of significant importance.

So far, there have been many studies in the field of flight delay propagation. Beatty et al. [5] put forward the concept of delay multiplier to quantify delay propagation through the analysis of flight status tables of airlines. Based on this idea, Wu and Law [6] proposed two new delay multiplier metrics to better explain historical flight delays. In addition, there are other studies such as [7], [8], and [9]. These studies are different at the level of details involved, but in general, they all focus on the impact of delays at hub airports. The current transportation system of civil aviation is a complex system, so the study of this system from the perspective of the complex network can grasp its characteristics more comprehensively. Fleurquin et al. [10] defined the indicators to measure the degree of network congestion, which describes the mode of delay propagation in the American air transportation network. The results point out that the connection mechanism between passengers and aircrew is the main internal cause of delay propagation. Campanelli et al. [11] use two different agent-based models to simulate the delay propagation and assess the effect of disruptions in the air traffic networks (US and ECAC areas). In the networks constructed in these works, nodes represent airports and links represent real air routes. Reference [12]–[14], and [15] also construct such networks at different geographical resolution scales. Such physical networks can describe the explicit interaction among airports, but cannot represent the real process of delay propagation. At present, only a few studies try to model the delay propagation problem as causality analysis with delay time.
series. Based on Granger causality test [16], Zanin et al. [17] and Du et al. [18] found the dynamic process of delay propagation through flight delay time series and constructed the delay causality network among airports. However, due to the high complexity of the current civil aviation transportation scenario, flight delay time series has the characteristic of non-linear, and the Granger causality test is an established model that can only deal with linear time series [19]. To solve this issue, Zhang et al. proposed using transfer entropy [20] which can deal with nonlinear time series to quantify the effect of delay propagation and verified that the quantitative results based on transfer entropy is of great help to the downstream prediction task. However, due to the high-dimensional characteristics of the delay time series in the current scene, it is hard to estimate accurately the high-dimensional condition mutual information (CMI) that is a key part in the calculation of transfer entropy, which eventually leads to the deviation of transfer entropy. If we can solve this problem and find more accurate delay propagation relations, it is of great significance not only to understand the delay propagation mechanism, but also to the downstream prediction task. Therefore, this paper proposes the low-dimensional approximation of CMI for transfer entropy (LTE), which is more suitable for high-dimensional and nonlinear delay time series. This improved algorithm can discover the real dynamic process of delay propagation (i.e., delay propagation relation) among airports more accurately and help us construct realistic delay propagation relation (DPR) network. Then, applying the complex network theory and its related indicators, we analyze the delay propagation characteristics of the DPR network from the global and local perspectives, and propose to compare DPR networks of different airlines to analyze the similarities and differences of airlines.

II. METHOD
This section explains in detail how to use the LTE to build a DPR network. The whole process takes time series of airport delay as input and outputs the DPR network, in which nodes correspond to airports and links between pairs of them correspond to delay propagation relations. Based on the DPR network, the complex network theory and its related indicators can be used to analyze the delay propagation mechanism from the perspective of multiple airports.

A. THE PROCESS OF CONSTRUCTING DPR NETWORK
1) TIME SERIES CONSTRUCTION
The total arrival delays per unit time window obtained from the historical status data of flight operation can depict the state of airport delays within the same time window. In order to capture the delay propagation relation among airports, we utilize arrival delay time series to infer the propagation relation between airports. In the arrival delay time series, each point represents the average delay time of flights in the corresponding time window. Two time series of a pair of airports are the input of the subsequent algorithm.

Specifically, the first step is to aggregate the arrival delays of each flight according to the airport within a one-hour time window which is regarded as the unit time in this paper. The second step is to average arrival delays according to the number of flights within the same windows. The specific calculation method is as follows:

\[
D^i(d, h) = \frac{\sum_{j=1}^{N_i} (T_{act}^{i,j} - T_{sch}^{i,j})}{N_i(d, h)} \quad 1 \leq h \leq 24 \quad (1)
\]

where \(T_{act}^{i,j}\) and \(T_{sch}^{i,j}\) represents the actual arrival time and the scheduled arrival time of the flight \(j\) whose scheduled arrival time is within the hour \(h\) of the day \(d\) at airport \(i\) and their difference indicates arrival delays of the flight \(j\); \(N_i(d, h)\) represents the number of flights whose scheduled arrival time is within the hour \(h\) of the day \(d\) at airport \(i\). It should be noted that this paper only pay attention to whether the scheduled arrival time of the flight \(j\) is within the hour \(h\), regardless of the actual arrival time of the flight \(j\). Then, the set of arrival delay time series of the airport \(i\) can be expressed as follows:

\[
S'(start, end) = \{D^i(start, 1), \ldots, D^i(start, 24), \ldots, D^i(end, 1), \ldots, D^i(end, 24)\} \quad (2)
\]

where \(start\) and \(end\) represent the starting and ending dates of a period of time, respectively. The length of the time series is \(T = (end - start + 1) \times 24\).

2) RELATION DISCOVERY
After all input time series are obtained, the delay propagation relations among airports based on the LTE. Before going into the details of our proposed method, we first show the traditional transfer entropy method.

Transfer entropy has been widely used to evaluate the flow of information within nonlinear dynamic systems. Let \(K\) subsystems \((X, Y, Z_1, \ldots, Z_{K-2})\) constitute an overall dynamical system \([x_t, y_t, z_{1,t}, \ldots, z_{K-2,t}]_{t=1}^N\). Suppose that the driving subsystem is \(X\) and the target subsystem is \(Y\). In other words, the current value of variable \(Y\) is affected by the past of variable \(X, Z = [Z_1, \ldots, Z_{K-2}]\) represent the remaining subsystems. The transfer entropy from subsystem \(X\) to \(Y\) on the condition of \(Z\) is defined as the form of conditional mutual information or the difference of two conditional entropies.

\[
T_{X \rightarrow Y | Z} = I\left(Y_n; X_n | Y_n^{-}, Z_n^{-}\right) - H\left(Y_n | X_n^{-}, Y_n^{-}, Z_n^{-}\right) \quad (3)
\]

\[
\begin{align*}
X_n^- = [X_{n-1}, X_{n-2}, \ldots, X_{n-L}], & \quad Y_n^- = [Y_{n-1}, Y_{n-2}, \ldots, Y_{n-L}], \\
Z_n^- = [Z_{n-1}, Z_{n-2}, \ldots, Z_{n-L}] & \text{are sets of variables that describe past states. The lags of } X, Y \text{ and } Z \text{ are sought within a range given by a maximum lag } L \text{ for each variable. The transfer entropy quantifies the information provided by the past of } X \text{ about the present state of } Y. \\

\end{align*}
\]

The key to compute transfer entropy \(T_{X \rightarrow Y | Z}\) is to construct mixed embedding vectors \(V_n \in M = \ldots \).
The process of construction is described as follows:

1) An empty embedding vector $V_n^{(0)} = \emptyset$ is initialized.
2) At the first iteration $k = 1$, the embedding vector $V_n^{(1)} \in M$ is selected most related to $Y_n$:

$$W_n^{(1)} = \arg\max_{W_n \in M} I(Y_n; W_n) \quad (4)$$

Then the embedding vector is $V_n^{(1)} = W_n^{(1)}$.

3) At the iteration $k > 1$, the mixed embedding vector is augmented by the component $W_n^{(k)}$ of $M$, giving most information about $Y_n$ additionally to the information already contained in $V_n^{(k-1)} = [W_n^{(1)}, \ldots, W_n^{(k-1)}]$. $W_n^{(k)}$ will be tested by a criterion through computing the CMI:

$$W_n^{(k)} = \arg\max_{W_n \in M \setminus V_n^{(k-1)}} I(Y_n; W_n|V_n^{(k-1)}) \quad (5)$$

4) The selection loop is terminated when $W_n^{(k)}$ does not fulfill the significance test and we have the mixed embedding vector $V_n = V_n^{(k-1)}$, which is composed of $V_n = [V_n^X, V_n^Y, V_n^Z]$, where $V_n^X$, $V_n^Y$ and $V_n^Z$ denote the components of $V_n$ belonging respectively to $X$, $Y$, and $Z$.

5) The transfer entropy is calculated as:

$$T_{X \rightarrow Y|Z} = H(Y_n|V_n^Y, V_n^Z) - H(Y_n|V_n) \quad (6)$$

The construction process of the mixed embedded vector can be seen as a typical feature selection process. The constructed mixed embedding vector maximizes the amount of information to predict the future state of target variable and minimizes redundancy within the set of selected variables. It is necessary to estimate the exact value of the CMI and select the variable with the largest amount of information to embed at each iteration. However, as the dimension of the conditional embedding vector $V_n$ increases, the problem of curse of dimensionality arises from the sparsity of the available data of an increasing volume of state space that makes the estimation of entropy rates gradually decrease towards zero. Hence, the problem of the curse of dimensionality makes the CMI estimation inaccurate. To overcome this problem, we present a low-dimensional approximation paradigm for TE estimation in the following sections.

Estimating high-dimensional CMI is a long-term challenge in the field of feature selection. There are many well-accepted techniques in published literature, such as the joint mutual information (JMI) criterion, the minimal-redundancy-maximal-relevance (MRMR) criterion, the double input symmetrical relevance (DISR) criterion. Brown et al. emphasize that lots of feature selection heuristics are all approximate iterative maximisers of the conditional likelihood, which can be interpreted in a unifying framework of conditional likelihood maximisation under certain assumptions of independence. Consequently, the methods are summarized as a parameterized general criterion:

$$J(W_k) = I(W_k; Y_n) - \beta \sum_{W_j \in V} I(W_k; W_j) + \gamma \sum_{W_j \in V} I(W_k; W_j|Y_n) \quad (7)$$

where the difference between different criteria depends on $\beta$ and $\gamma$. For example, the JMI criterion is obtained with $\beta = \gamma = 1/|V|$ and the MRMR criterion is under the case of $\beta = 1/|V|, \gamma = 0$. Higher-order feature interactions are considered in the RelaxMRMR which is a principle method for modifying feature selection criteria by relaxing some assumptions of conditional independence.

Then, we present the LTE based on nonuniform embedding, which uses a low-dimensional approximation CMI criterion for embedding:

1) An empty embedding vector $V_n^{(0)} = \emptyset$ is initialized.
2) At the first iteration $k = 1$, the embedding vector $V_n^{(1)} \in M$ is selected most related to $Y_n$:

$$W_n^{(1)} = \arg\max_{W_n \in M} I(Y_n; W_n) \quad (8)$$

Then the embedding vector $V_n^{(1)} = W_n^{(1)}$.

3) At the iteration $k > 1$, the mixed embedding vector is augmented by the component $W_n^{(k)}$ of $M$, giving most information about $Y_n$ additionally to the information already contained in $V_n^{(k-1)} = [W_n^{(1)}, \ldots, W_n^{(k-1)}]$. $W_n^{(k)}$ will be tested by a criterion through computing the CMI:

$$J(W_k) = I(W_k; Y_n) - \beta \sum_{W_j \in V} I(W_k; W_j) + \gamma \sum_{W_j \in V} I(W_k; W_j|Y_n) \quad (9)$$

4) The selection loop is terminated when $W_n^{(k)}$ does not fulfill the significance test and we have the mixed embedding vector $V_n = V_n^{(k-1)}$, which is composed of $V_n = [V_n^X, V_n^Y, V_n^Z]$, where $V_n^X$, $V_n^Y$ and $V_n^Z$ denote the components of $V_n$ belonging respectively to $X$, $Y$, and $Z$.

5) The transfer entropy is calculated as:

$$T_{X \rightarrow Y|Z} = H(Y_n|V_n^Y, V_n^Z) - H(Y_n|V_n) \quad (10)$$

Suppose $X$ and $Y$ represent the average arrival delay time series of airport $A_{apt}$ and $B_{apt}$ respectively. $Z_1, \ldots, Z_{K-2}$ denote the average arrival delay time series of other airports. The transfer entropy can be calculated in the above way. If the transfer entropy is zero, we can conclude that there is no delay propagation relation between airport $B_{apt}$ and airport $A_{apt}$. However, if the transfer entropy is greater than zero, we need the significance test before drawing conclusion. First, we establish a null hypothesis that there is no delay propagation relation between airport $B_{apt}$ and airport $A_{apt}$.
Next, we artificially construct $S$ time series whose statistical properties and sequence length are the same as the time series $Y_t$. Then, we calculate the transfer entropy between each artificially constructed time series and $X_t$, respectively and combine them with the original transfer entropy into a incremental set $D$. The p-value of the test is $1-(i-0.326)/(i+1+0.348)$, where $i$ presents the index position of the original transfer entropy in the set $D$. Finally, if p-value is greater than the significance level $\alpha$, we conclude that there is no delay propagation relation between airport $B_{apt}$ and airport $A_{apt}$; if p-value is less than the significance level $\alpha$, we conclude that there is a delay propagation relation from airport $A_{apt}$ to airport $B_{apt}$.

3) NETWORK CONSTRUCTION

Based on the above delay propagation relation discovery method, the adjacent matrix $A = \{(a_{ij})\} \epsilon \{0, 1\}^{N \times N}$ is obtained after traversing all the airport pairs. $a_{ij} = 1$ indicates that airport $i$ will propagate the delay to airport $j$; that is, there is a directed link from airport $i$ to airport $j$. $a_{ij} = 0$ indicates that airport $i$ will not propagate the delay to airport $j$; that is, there is no directed link from airport $i$ to airport $j$. A DPR network is constructed in which nodes correspond to airports and links between pairs of them correspond to delay propagation relation.

B. NETWORK ANALYSIS INDICATORS

The complex network theory and its related indicators are important means to analyze complex systems in the real world from multiple perspectives. The current civil aviation transportation system includes many airports and airlines. Thus, the scene of delay propagation is extremely complex. In this paper, we use the complex network as a tool to analyze the delay propagation relations from the perspective of multiple airports. The network indicators used in the analysis are described below.

In a DPR network, $N$ represents the number of nodes which are associated to airports; $E$ represents the number of links which are associated to delay propagation relation. $A = \{(a_{ij})\} \epsilon \{0, 1\}^{N \times N}$ is the adjacency matrix of the network.

1) DEGREE

The degree of an airport indicates the number of airports that have delay propagation relations with it. The in-degree and out-degree of airport $i$ are $d^i_{in} = \sum_{j=1}^{N} a_{ij}$ and $d^i_{out} = \sum_{j=1}^{N} a_{ji}$ respectively, which represents the airport $i$ would be affected by delayed propagation of $d^i_{in}$ airports and will propagate delays to $d^i_{out}$ airports. The average degree of the network is $< d >= E/N$, which indicates each airport has delayed propagation relations with other $E/N$ airports on average.

2) LINK DENSITY

Link density denotes the proportion of active links to the total number of potential links, i.e., the proportion of airport pairs with delayed propagation relations to all airport pairs.

It is defined as follows:

\[
L = \frac{\sum_{i \neq j} a_{ij}}{N^2} = \frac{E}{N^2} \quad (11)
\]

The higher the link density $L$ is, the denser the network is, and the more likely the delay is to propagate in the network.

3) ASSORTATIVITY COEFFICIENT

Assortativity coefficient [21] is used to examine whether vertices with similar degrees tend to be connected. The coefficient is calculated as follows:

\[
S = \frac{1}{E} \sum_{j>i} \frac{1}{2} (d_i + d_j) a_{ij} \quad (12)
\]

where $d_i$ represents the degree of airport $i$ including the in-degree and out-degree.

The value of $S$ is positive, which indicates that airports tend to propagate delays to airports with similar degrees and the network is called assortative. Otherwise, airports tend to propagate delays to airports that differ greatly from its degrees and the network is called disassortative.

4) CLUSTERING COEFFICIENT

For a specific node in the network, its local clustering coefficient $C(i)$ [22] reflects the degree of clustering (complete subgraph) of the nodes connected with it. The calculation formula is as follows:

\[
C(i) = \frac{T(i)}{d^i_{out} (d^i_{out} - 1) - 2d^i_{in}} \quad (13)
\]

where $T(i)$ is the number of directed triangles through airport $i$, $d^i_{out}$ is the sum of in degree and out degree of airport $i$, and $d^i_{in}$ is the reciprocal degree of airport $i$. After obtaining the local clustering coefficient of each vertex in the network, the average clustering coefficients of the whole network can be calculated. The calculation formula is as follows:

\[
\overline{C} = \frac{1}{N} \sum_{i=1}^{N} C(i) \quad (14)
\]

The higher the average clustering coefficient $\overline{C}$ is, the higher the clustering degree of airport nodes in the network is, and it is easier to form a local area in which delays propagate.

5) MODULARITY

Modularity [23] is used to measure the strength of dividing a network into multiple communities. Community [24] refers to dividing the delay propagation relation network into several sub-regions, in which the internal airports of each sub-region have dense delay propagation links, but the links with airports outside the sub-region are sparse. The modularity $Q$ is defined as:

\[
Q = \frac{1}{E} \sum_{ij} \left( a_{ij} - \frac{d^i_{out} d^j_{out}}{E} \right) \delta (c_i, c_j) \quad (15)
\]
where if the airport \( i \) and \( j \) are in the same community, the \( \delta \) function outputs 1, otherwise 0. The higher the modularity \( Q \) is, the higher the regionalization degree of delay propagation is.

6) EFFICIENCY

The efficiency reflects how easily delay propagates between two airports, in other words, how many intermediate nodes must delays pass through to reach the target node, which is defined as follows:

\[
E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{e_{ij}}
\]  

(16)

where \( e_{ij} \) is the distance between airport \( i \) and airport \( j \). The higher the efficiency is, the faster delays can propagate to the whole network.

7) RANDOM NETWORK

Network indicators need to be compared to illustrate whether it is high or low. Random network refers to the use of the ER random graph model [25] to generate networks with the same number of nodes and the same number of links as the original network. After obtaining more random networks, the average value of corresponding indicators is obtained, which can be used for the comparison of indicators.

8) NETWORK MOTIF

Network motif [26] reflects the local relation pattern among any three airports. The Z-score method can be used to evaluate the significance level of each motif. The calculation formula is as follows:

\[
Z(i) = \frac{U_i - \langle U'_i \rangle}{\sqrt{\langle (U'_i)^2 \rangle - \langle U'_i \rangle^2}}
\]  

(17)

where \( U_i \) is the number of motif \( i \) found, \( U'_i \) is the count of the same motif but in a random network \( r \), and \( \langle \cdot \rangle \) denotes the average value in several random networks. These random networks need not only the same number of nodes and links but also the same in-degree and out-degree of each node as the original network.

III. RESULTS

In this section, we first give the simulation experimental results based on artificial nonlinear time series. The results prove that the proposed algorithm can discover causal relations from multiple types of nonlinear time series accurately. Then, based on the real historical status data of flight operation, we use the proposed algorithm to construct DPR networks and analyze the delay propagation from three perspectives by the complex network theory, which are the basic characteristics, the local characteristics, and the layered characteristics of the networks.

A. SIMULATION EXPERIMENT

Due to the complexity of the current civil aviation transportation scenario, flight delay time series has the characteristic of nonlinearity. Specific proof can be found in subsection C. Since we cannot achieve the causality label from real-world delay propagation data, in this subsection we verify the effectiveness of our method on simulated data comparing with the other two representative methods. One is the Granger causality (GC) test [16], [27] developed by the economy Nobel Prizewinner Clive Granger. The other is the Transfer entropy (TE) [20]. To ensure the robustness and reliability of simulation experiments, we randomly generate 100 implementations for each simulation system. The causality intensity obtained is the average of 100 implementations.

1) PERFORMANCE EVALUATION

Sensitivity(SE), specificity (SP), and F1 score are used to evaluate the performance of different methods. The concrete formulas are as follows:

\[
\text{SE} = \frac{TP}{TP + FN}
\]  

(18)

\[
\text{SP} = \frac{TN}{TN + FP}
\]  

(19)

\[
\text{F1 score} = \frac{2TP}{2TP + FP + FN}
\]  

(20)

where TP (true positive) represents the number of actual positives that are correctly identified by methods; FN (false negative) represents the number of actual positives that are incorrectly rejected by methods; TN (true negative) represents the number of actual negatives that are correctly rejected by methods; FP (false positive) represents the number of actual negatives that are incorrectly identified by methods.

2) NONLINEAR MULTIVARIATE SYSTEM

The first simulation system is a nonlinear VAR process of order 1 in three variables NLVAR,3(1) with unit Gaussian noise as shown in (21).

\[
\begin{align*}
&x_{1,n} = 0.4 \ast x_{1,n} + 3.4 \ast x_{1,n-1} \ast (1 - x^2_{1,n-1}) \\
&\ast e^{-x^2_{1,n-1}} + u_{1,n} \\
&x_{2,n} = 0.4 \ast x_{2,n} + 3.4 \ast x_{2,n-1} \ast (1 - x^2_{2,n-1}) \\
&\ast e^{-x^2_{2,n-1}} + 0.5 \ast x_{1,n-1} \ast x_{2,n-1} + u_{2,n} \\
&x_{3,n} = 0.4 \ast x_{3,n} + 3.4 \ast x_{3,n-1} \ast (1 - x^2_{3,n-1}) \\
&\ast e^{-x^2_{3,n-1}} + 0.3 \ast x_{2,n-1} + 0.5 \ast x^2_{1,n-1} + u_{3,n}
\end{align*}
\]  

(21)

According to the above equations, the true causal relations of this simulation system are \( X_1 \rightarrow X_2 \), \( X_1 \rightarrow X_3 \) and \( X_2 \rightarrow X_3 \), which are represented as the directed network in Fig. 1.

Fig. 2 shows that the matrix representations of the detected average causality intensity for the NLVAR,3(1) with the time series length \( N = 1024 \) measured by different methods. The true causal relations in the NLVAR,3(1) are represented as the matrix elements (1,2), (1,3) and (2,3) in Fig. 2a. The GC test only detects the element (2,3) with high causality intensity, but the elements (1,2) and (1,3) are almost neglected. Although the TE finds three correct elements, there are some redundant relations. The LTE proposed in this
paper accurately detects all the correct elements with higher intensity and fewer redundancy than the TE. The sensitivity, specificity, and F1 scores obtained from 100 implementations of the NLVAR3(1) with different time series lengths are listed in Table 1. It is obvious that the three indicators of the LTE are the highest.

\[
x_{i,t} = \begin{cases} 
1.4 - x_{i,t-1}^2 + 0.3x_{i,t-2} & \text{for } i = 1, K \\
1.4 - (0.5C (x_{i-1,t-1} + x_{i+1,t-1}) + (1-C)x_{i,t-1})^2 +0.3x_{i,t-2} & \text{for } j = 2, \ldots, K - 1 
\end{cases} 
\]

(22)

### TABLE 1. Sensitivity, specificity, and F1 score of different methods for the NLVAR3(1).

| Methods | N  | SE  | SP  | F1 score |
|---------|----|-----|-----|----------|
| GC      | 512| 0.373| 0.945| 0.503    |
| TE      | 1.000| 0.958| 0.960|          |
| LTE     | 1.000| 0.968|       |          |
| GC      | 1.000| 0.978| 0.945| 0.499    |
| TE      | 1.000| 0.942| 0.956|          |
| LTE     | 1.000| 0.958|       |          |

3) COUPLED HÉNON MAPS

The second simulation system is K(K = 6) coupled Hénon maps, defined as (22).

According to the above equations, the true causal relations of this simulation system are \(X_1 \rightarrow X_2, X_2 \rightarrow X_3, X_3 \rightarrow X_4, X_4 \rightarrow X_5, X_5 \rightarrow X_6, X_6 \rightarrow X_5, X_5 \rightarrow X_4, X_4 \rightarrow X_3\) and \(X_3 \rightarrow X_2\), which are represented as the directed network in Fig. 3.

Fig. 4 shows that the matrix representations of the detected average causality intensity for the coupled Hénon maps with the time series length \(N = 1024\) measured by different methods. The true causal relations in the coupled Hénon maps are represented as the matrix elements (1,2), (2,3), (3,4), (4,5), (5,4), (4,3) and (3,2) in Fig 4(a). The GC test detects redundant causal relations. Like the previous simulation system, the LTE detects all the correct elements with higher intensity than the TE and contains no redundancy. The sensitivity, specificity, and F1 scores obtained from 100 implementations of the coupled Hénon maps with different time series lengths are listed in Table 2. It is obvious that the three indicators of the LTE are the highest.

### TABLE 2. Sensitivity, specificity, and F1 score of different methods for the Coupled Hénon Maps.

| Methods | N  | SE  | SP  | F1 score |
|---------|----|-----|-----|----------|
| GC      | 512| 0.995| 0.809| 0.747    |
| TE      | 1.000| 1.000| 1.000|          |
| LTE     | 1.000| 1.000|       |          |
| GC      | 1.000| 0.733| 0.100| 0.681    |
| TE      | 1.000| 1.000| 1.000|          |
| LTE     | 1.000| 1.000|       |          |

Based on the comparison of the above results, it is evident that the LTE is better at performing causality analysis on a variety of nonlinear systems than the other two methods.

### B. REAL DATASET DESCRIPTION

The dataset which contains flight status data of all departure and arrival airports in China from 1 to 31 December 2016, including more than 300,000 flights, more than 200 airports, and more than 30 airlines used in this paper is provided by China Civil Aviation Information Network Co., Ltd. For each flight, available information includes airline IATA code,
flight number, aircraft registration number, scheduled departure time, scheduled arrival time, actual departure time, actual arrival time, departure airport IATA code, and arrival airport IATA code.

C. DATA PREPROCESSING AND ANALYSIS

Different analysis perspectives correspond to different networks constructed by different data, such as delay time series of different objects or different lengths. For the basic and local characteristics, we need to calculate the set of monthly arrival-delay time series \( \{S_i(31)\}_{i=1}^N \) for \( N \) airports, and then obtain a DPR network for the whole month based on the proposed algorithm. For the layered characteristics, we construct a monthly network for each airline individually.

To verify the nonlinearity of arrival delay time series, we take the set of monthly arrival delay time series \( \{S_i(31)\}_{i=1}^N \) for \( N \) airports as an example and calculate the max Lyapunov exponent which is an important parameter for studying chaos of the monthly arrival delay time series for each airport in December 2016, as shown in Fig. 5. If the maximum Lyapunov exponent is greater than 0, it can be determined that the system is chaotic [28]. The maximum value of all the max Lyapunov exponent is about 0.15, which proves that the arrival delay time series is chaotic, i.e., nonlinear.

D. BASIC CHARACTERISTICS OF DPR NETWORK

To grasp the characteristics of delay propagation among airports, we analyze the basic characteristics of the DPR network for the whole month. Fig. 6 shows that there are 205 airport nodes and 537 directed links representing the delay propagation relations. The size of nodes is proportional to the number of flights, and the color is related to the degree of nodes (from yellow to red, the degree increases gradually).

In Fig. 6, it is noteworthy that the degree of red nodes is high, but the size is small. On the contrary, the size of some yellow nodes is large, but the degree is low. Fig. 7 further shows that there is an inverse relation between the number of flights and the degree of nodes. The degree of airport nodes with a larger number of flights is extremely small, which is consistent with the situation observed in other aviation transportation systems [17], [29].

Based on the above analysis, it can be concluded that in the daily delay propagation interaction among airports, airports with a small number of flights frequently appear. The possible reason for this phenomenon is that with the increase of the number of flights at the airport, the impact of the delay of single flight on airports will be gradually “diluted”. Airports with a larger number of flights have better measures and facilities to alleviate delays, but airports with a smaller number of flights lack such conditions. Therefore, airports with a smaller number of flights may cause or even aggravate the propagation of delays in the network.
Table 3 lists the values of indicators in the DPR network and the average values of indicators of 1000 random networks. According to the value of the average degree, airports have delay propagation relations with approximate three other airports on average. The link density indicates this network is sparse. The clustering coefficient, modularity, and efficiency of this network are higher than those average values of the random networks, which is indicative of that the network has a strong aggregation trend, relatively obvious community structure and delays are easier to propagate. The assortativity coefficient of the network is lower than the average value of the random networks. The result suggests that the network is more disassortative. In other words, the nodes with the larger degree tend to connect with the nodes with the smaller degree. Then, taking into account the contrary relation between degree and the number of flights, the delay propagation between airports with a larger number of flights and those with a smaller number of flights is more obvious.

Table 3. The Values of Indicators in the Network.

| Indicators                  | The values of indicators in the DPR network | The average values of indicators in 1000 random networks |
|-----------------------------|---------------------------------------------|--------------------------------------------------------|
| Average degree              | 2.6195                                      | 2.6195                                                 |
| Link density                | 0.0128                                      | 0.0128                                                 |
| Assortativity coefficient   | -0.0905                                     | -0.0059                                                |
| Clustering coefficient      | 0.0199                                      | 0.0123                                                 |
| Modularity                  | 0.6250                                      | 0.4181                                                 |
| Efficiency                  | 0.3871                                      | 0.3235                                                 |

The previous section describes the analysis results of the basic indicators of the monthly DPR network, while this section introduces the analysis results of the local relation patterns of the same network combined with the airport levels. In order to assess the ability of an airport in providing scheduled flight services, CAAC (civil aviation administration of China) has defined the airport level [30], which are 4F, 4E, 4D, 4C, and 3C, respectively, from high to low. Besides, the 4C airports account for about 53% of the total airports in China. In order to get a more universal conclusion and correspond to the above, 4F airports are regarded as large airports, 4E and 4D airports as medium airports, and 4C and 3C airport as small airports. After the three airport levels are assigned to each node in the monthly DPR network, the local relation patterns of delay propagation and the difference of delay propagation at different levels of airports are analyzed by using the network motif. The calculation method of the significance level of the network motif is shown in section II and the number of random networks generated for significance calculation is 100.

We selected 8 motifs with the highest significance level, and the rest of the network motifs with very low significance levels or less than zero. The significance level of each motif in the network is shown in the middle of Fig. 8, and the specific components of 8 motifs are shown around Fig. 8. The significant levels of G’1, G’2 and G’4 are all relatively high. It indicates that delay propagation among small airports occurs frequently. As can be seen from the significance levels of G’3, G’7, and G’8, small airports also often propagate delays to medium airports and large airports, but the significance level is lower than that among small airports. The G’5 and G’6 reflect the same characteristics of medium airports. This relation pattern is consistent with the previous conclusion that there is an inverse relation between the degree and the number of flights. In all showed significant relation patterns, only G’4 shows cascaded propagation of delays and only among small airports. This reflects that delays are likely to be absorbed by larger airports in the process of delay propagation, thus avoiding cascaded propagation of delays, whereas smaller airports are not. The above can be summarized as follows: in the daily delay propagation interaction among airports, large airports always receive delays to alleviate delay propagation and inhibit the occurrence of delay cascaded propagation, while small airports always propagate delays outwards to cause or aggravate delay propagation.

F. LAYERED CHARACTERISTICS OF DPR NETWORK

The civil aviation transportation system is mainly operated by several airlines. In order to analyze the difference of delay propagation among different airlines, this section constructs a monthly DPR network for each airline. This is equivalent to layering the whole DPR network according to airlines. The airlines are sorted according to the number of their flights. Here we only give the results for the top ten airlines and their IATA codes are replaced by A to J. As shown in Fig. 9, airlines A, B, and C are the three major airlines in China, while the scale of other airlines is relatively small.

It is noteworthy that the airline H shown in Fig. 9 has a small number of flights, but the summary of arrival delays is almost the same as that of airline C. The number of nodes and the number of links in the DPR network of each airline shown in Fig. 10 have a similar phenomenon. The number of nodes and links in the network of airline H are almost higher than that of other small airlines. These prove that although the number of flights of airlines H is not very large, its delays...
and delay propagation are more serious than the other small airlines.

Fig. 11 shows the assortativity coefficient and efficiency of the DPR network of each airline. Combining with Fig. 10, the difference between the three major airlines and the other small airlines can be recognized. Due to a large number of navigation cities and flights of the three major airlines, the possible scope of delay propagation is also large, so the

**FIGURE 8.** The significance level of network motifs in the monthly DPR network.

**FIGURE 9.** The number of flights of the top 10 airlines and their summary of arrival delays in one month.

**FIGURE 10.** The number of nodes and links of the monthly DPR network of each airline.
number of links and nodes of the corresponding network will be much larger than the rest of the airlines. In another aspect, the efficiency of delay propagation of the three major airlines is lower than that of other airlines. This means that delays propagate more efficiently in the aviation systems of smaller airlines. For the three major airlines, delay propagation probably is more difficult because of the large scale of the network or the better resources to alleviate delays. As for the assortativity coefficient of Fig. 11, it can be recognized that small airlines have smaller assortativity coefficient than those major airlines and that of the two airlines F and G is still the smallest. This means that nodes with large differences in degrees tend to connect with each other in the network of small airlines. This causes the network to be star-shaped, which also makes delay propagation efficient in the aviation system of small airlines.

IV. CONCLUSION

This paper proposes the low-dimensional approximation of conditional mutual information for transfer entropy which can overcome the drawbacks of traditional causality analysis methods to discover the accurate delay propagation relation among airports from the high-dimensional and nonlinear arrival delay time series and construct the delay propagation relation network. Besides, the mechanism of delay propagation is analyzed from the perspective of multiple airports by using the complex network theory and the results obtained can provide references for the improvement of current flight delays. The analysis results show that small airports always cause or aggravate delay propagation, so the possibility of cascaded propagation of delays occurring in flight chains passing through several small airports is relatively high. Once delays occur in a certain link, managers of subsequent flights should be vigilant and make corresponding preparations and countermeasures. Additionally, another analysis result is that the efficiency of delay propagation of the aviation systems of small airlines is higher than three major airlines, which requires these airlines to do some optimization, such as combining our findings that small airports play an important role in the delay propagation to make more accurate predictions and response measures. The study of using causality analysis to explore delay propagation relation can be extended further. For example, this paper completes the comparison of differences among airlines based on the network of each airline while the interaction among airlines may also deserve analyzed in the future. Furthermore, the results of causality analysis can be taken as prior knowledge of downstream tasks.

REFERENCES

[1] N. Kafle and B. Zou, “Modeling flight delay propagation: A new analytical-econometric approach,” Transp. Res. B, Methodol., vol. 93, pp. 520–542, Nov. 2016.
[2] P. Belobaba, A. Odoni, and C. Barnhart, The Global Airline Industry, Hoboken, NJ, USA: Wiley, 2015.
[3] S. AhmadBeygi, A. Cohn, Y. Guan, and P. Belobaba, “Analysis of the potential for delay propagation in passenger airline networks,” J. Air Transp. Manage., vol. 14, no. 5, pp. 221–236, Sep. 2008.
[4] J. Ferguson, A. Q. Kara, K. Hoffman, and L. Sherry, “Estimating domestic US airline cost of delay based on European model,” Transp. Res. C, Emerg. Technol., vol. 33, pp. 311–323, Aug. 2013.
[5] R. Beatty, R. Hsu, L. Berry, and J. Rome, “Preliminary evaluation of flight delay propagation through an airline schedule,” Air Traffic Control Quart., vol. 7, no. 4, pp. 259–270, Oct. 1999.
[6] C.-L. Wu and K. Law, “Modelling the delay propagation effects of multiple-resource connections in an airline network using a Bayesian network model,” Transp. Res. E, Logistics Transp. Rev., vol. 122, pp. 62–77, Feb. 2019.
[7] L. Schaefler and D. Milliner, “Flight delay propagation analysis with the detailed policy assessment tool,” in Proc. IEEE Int. Conf. Syst., Man Cybern., e-Syst. e-Man Cybern. Cyberspace, vol. 2, Oct. 2001, pp. 1299–1303.
[8] P. T. Wang, L. A. Schaefler, and L. A. Wojick, “Flight connections and their impacts on delay propagation,” in Proc. 22nd Digit. Avionics Syst. Conf. (DASC), vol. 1, Oct. 2003, p. 5.B.4-5.1-9.
[9] L. Lacasa, M. Cea, and M. Zanin, “Jamming transition in air transportation networks,” Phys. A, Stat. Mech. Appl., vol. 388, no. 18, pp. 3948–3954, Sep. 2009.
[10] P. Fleurquin, J. J. Ramasco, and V. M. Eguíluz, “Systemic delay propagation in the US airport network,” Sci. Rep., vol. 3, no. 1, p. 1159, Dec. 2013.
[11] B. Camporelli, P. Fleurquin, A. Arranz, I. Etxebarria, C. Ciruelos, V. M. Eguíluz, and J. J. Ramasco, “Comparing the modeling of delay propagation in the US and European air traffic networks,” J. Air Transp. Manage., vol. 56, pp. 12–18, Sep. 2016.
[12] T. Opsahl, V. Colizza, P. Panzarasa, and J. J. Ramasco, “Prominence and control: The weighted rich-club effect,” Phys. Rev. Lett., vol. 101, no. 16, Oct. 2008, Art. no. 168702.
[13] D. R. Wuehner, S. Roy, and R. M. D’Souza, “Resilience and rewiring of the passenger airline networks in the United States,” Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top., vol. 82, no. 5, 2010, Art. no. 056101.
[14] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, “The architecture of complex weighted networks,” Proc. Nat. Acad. Sci. USA, vol. 101, no. 11, pp. 3747–3752, Mar. 2004.
[15] R. Guimera, S. Mossa, A. Turtschi, and L. A. N. Amaral, “The worldwide air transportation network: Anomalous centrality, community structure, and cities’ global roles,” Proc. Nat. Acad. Sci. USA, vol. 102, no. 22, pp. 7794–7799, May 2005.
[16] C. W. Granger, “Investigating causal relations by econometric models and cross-spectral methods,” Econometrica, vol. 37, no. 3, pp. 424–438, 1969.
[17] M. Zanin, S. Belkoura, and Y. Zhu, “Network analysis of Chinese air transport delay propagation,” Chin. J. Aeronaut., vol. 30, no. 2, pp. 491–499, Apr. 2017.
[18] W.-B. Du, M.-Y. Zhang, Y. Zhang, X.-B. Cao, and J. Zhang, “Delay causality network in air transport systems,” Transp. Res. E, Logistics Transp. Rev., vol. 118, pp. 466–476, Oct. 2018.
[21] M. E. J. Newman, “Assortative mixing in networks,” *Phys. Rev. Lett.*, vol. 89, no. 20, Oct. 2002, Art. no. 208701.

[22] D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” *Nature*, vol. 393, no. 6684, p. 440, 1998.

[23] M. E. J. Newman, “Fast algorithm for detecting community structure in networks,” *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, no. 6, Jun. 2004, Art. no. 066133.

[24] A. Arenas, J. Duch, A. Fernández, and S. Gomez, “Size reduction of complex networks preserving modularity,” *New J. Phys.*, vol. 9, no. 6, p. 176, 2007.

[25] P. Erdos and A. Renyi, “On random graphs I,” *Publicationes Math. Debrecen, Tech. Rep.*, 1959, vol. 6, nos. 290–297, p. 18.

[26] S. S. Shen-Ort, R. Milo, S. Mangan, and U. Alon, “Network motifs in the transcriptional regulation network of Escherichia coli,” *Nature Genet.*, vol. 31, no. 1, pp. 64–68, 2002.

[27] J. F. Geweke, “Measures of conditional linear dependence and feedback between time series,” *J. Amer. Stat. Assoc.*, vol. 79, no. 388, pp. 907–915, Dec. 1984.

[28] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, “Determining Lyapunov exponents from a time series,” *Phys. D, Nonlinear Phenomena*, vol. 16, no. 3, pp. 285–317, Jul. 1985.

[29] A. Cook, G. Tanner, and M. Zanin, “Towards superior air transport performance metrics-imperatives and methods,” *J. Aerosp. Oper.*, vol. 2, nos. 1–2, pp. 3–19, 2013.

[30] CAAC, *Aerodrome Technical Standards*. Accessed: Jul. 31, 2013. [Online]. Available: http://www.caac.gov.cn/XXGK/XXGK/BZGF/HYBZ/201511/P020151103341364055252.pdf

YINHONG XIAO received the bachelor’s degree in industrial electronic technology from Zhejiang University, in 1984, and the master’s degree in engineering management and science from Beihang University, in 1999.

From July 1984 to August 2000, he was with the Civil Aviation Computer Center. From August 2000 to August 2008, he was the Vice President of TravelSky Technology Ltd. From August 2008 to December 2017, he acted as a member of the Standing Committee of the CPC of TravelSky Holding, served as the Board Director, a General Manager, and a member of the Standing Committee of the CPC of TravelSky Technology Ltd. Since December 2017, he entitled as the Board Director and a Deputy Secretary of CPC committee of TravelSky Technology Ltd.

He awarded for science and technology project by National Development and Reform Commission, State Scientific and Technological Commission, and Ministry of Finance of P.R.C., in 1996. He received the Government Special Allowance of the State Council, in 1997. He has twice won the first prize Civil Aviation Science and Technology Progress Award, with triple second and once third prized awarded afterwards.

YAOSHUAI ZHAO received the B.E.C. degree in economics from Xi’an Jiaotong University, China, in 2000, and the M.B.A. degree in management engineering from the Chinese Academy of Sciences, China, in 2010.

Since 2018, he has been a Senior Software Engineer and the Associated Department Manager of the Key Laboratory of Intelligent Passenger Service of Civil Aviation, TravelSky Technology Limited, China. He has led several CAAC research projects, such as Development and Research of Newsky Flight Control System (a Science and Technology Fund Project of CAAC, MHRD2009205, in 2010). He has published some articles in the journal and obtained a number of utility model patents in intelligent passenger service. His awards and honors include the Civil Aviation Science and Technology Award (CCAC), in 2013.

GE WU received the B.S. degree in information and computing science from Shandong University, China, in 2013.

Since 2015, she has been a Software Engineer, a Project Researcher, and a Project Manager of TravelSky Technology Limited, China. She has led and participated in several research projects in different fields, such as intelligent flight optimization in irregular flight scenes, self-identity service of segment market passenger, and civil aviation market forecast. She has obtained a number of utility model patents in intelligent civil aviation passenger service. Her research interest includes development of intelligent civil aviation passenger service using machine learning and deep learning techniques.

YIZHEN JING received the B.E. degree from SWJTU, in 2018. He is currently pursuing the master’s degree with the School of Computer and Information Technology, Beijing Jiaotong University (BJTU). His research interests include time series, data mining, and machine learning.

***