Temporal gravity model for important node identification in temporal networks

Bi, Jialin; Jin, Ji; Qu, Cunquan; Zhan, Xiuxiu; Wang, Guanghui; Yan, Guiying

DOI
10.1016/j.chaos.2021.110934

Publication date
2021

Document Version
Final published version

Published in
Chaos, Solitons and Fractals

Chaos, Solitons and Fractals 147 (2021) 110934

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Temporal gravity model for important node identification in temporal networks

Jialin Bi[a], Ji Jin[a], Cunquan Qu[a,b,*], Xiuxiu Zhan[c], Guanghui Wang[a,b], Guiying Yan[d,e]

[a] School of Mathematics, Shandong University, Jinan 250100, PR China
[b] Data Science Institute, Shandong University, Jinan 250100, PR China
[c] Delft University of Technology, Intelligent Systems, Delft 2600GA, the Netherlands
[d] Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, PR China
[e] University of Chinese Academy of Sciences, Beijing 100049, PR China

ARTICLE INFO

Article history:
Received 12 December 2020
Revised 20 March 2021
Accepted 29 March 2021

Keywords:
Temporal networks
Temporal gravity model
Important node
Centrality

ABSTRACT

Identifying important nodes in networks is essential to analysing their structure and understanding their dynamical processes. In addition, myriad real systems are time-varying and can be represented as temporal networks. Motivated by classic gravity in physics, we propose a temporal gravity model to identify important nodes in temporal networks. In gravity, the attraction between two objects depends on their masses and distance. For the temporal network, we treat basic node properties (e.g., static and temporal properties) as the mass and temporal characteristics (i.e., fastest arrival distance and temporal shortest distance) as the distance. Experimental results on 10 real datasets show that the temporal gravity model outperforms baseline methods in quantifying the structural influence of nodes. When using the temporal shortest distance as the distance between two nodes, the proposed model is more robust and more accurately determines the node spreading influence than baseline methods. Furthermore, when using the temporal information to quantify the mass of each node, we found that a novel robust metric can be used to accurately determine the node influence regarding both network structure and information spreading.

1. Introduction

Network science is increasingly important in numerous fields, including physics, biology, finance, and social sciences. In fact, many real systems can be suitably represented as complex networks [1,2].

The nodes in a network may exhibit varying connectivity and represent different dynamical processes, such as epidemic spreading, information diffusion, and opinion formation. If we remove a node from a network such that the network collapses into disconnected components, this node is important in terms of network connectivity. We call this type of influence the structural influence. On the other hand, a node can be the seed of information (epidemic) spreading and cause wide circulation through the network. Such a node is influential in terms of spreading, so we call this type of influence the spreading influence. For any type of influence, we call the corresponding node an important node.

Because important nodes influence network behaviour, they should be identified [3–6]. If the connections between nodes are fixed, they establish a static network. Various methods for important node identification have been developed for static networks [7], and they can be divided into structural-based centrality methods [8,9] (e.g., degree [10], closeness [11] and betweenness centrality [12]) and iterative-based centrality methods (e.g., PageRank [13], HITS [14], and SALSA [15]). Inspired by the concept of gravity, Ma et al. [16] proposed two gravity models, namely, gravity centrality and extended gravity centrality, to identify influential spreaders on static networks by considering both neighbourhood information and path information. Likewise, Li et al. [17] proposed a local gravity model that relies on a truncation radius. However, these methods are restricted to static networks.

In practice, many systems are time-varying [18–21], and the time order has been shown to substantially influence the network structure and information spreading [19,22]. Connections appearing in a complex system can be represented by a temporal network [23,24]. In temporal networks, however, identifying important nodes is much more challenging than in static networks. In fact, a node may play different roles over time in a temporal network [22], so its importance varies over time. For example, an in-

* Corresponding author at: School of Mathematics, Shandong University, 27 Shanda Nanlu, Jinan 250100, PR China.
E-mail addresses: jialinbi@mail.sdu.edu.cn (J. Bi), mincjinji@gmail.com (J. Jin), cqqu@sdu.edu.cn (C. Qu), x.zhan@tudelft.nl (X. Zhan), ghwang@sdu.edu.cn (G. Wang), yangy@amt.ac.cn (G. Yan).

https://doi.org/10.1016/j.chaos.2021.110934
0960-0779/

dividual may be very active and post several messages and information on a social network during a given year and then become inactive the next year, interrupting all information spread. Thus, to identify important nodes in temporal networks, we should consider structural properties and time-dependent information.

Most metrics for temporal networks extend those for static networks [25,26]. Hence, various methods either integrate a temporal network into a static one or segment a temporal network into a series of static snapshots over time. For instance, a temporal network considering centrality metrics can be analysed as follows. First, the temporal network is divided into several snapshots at a given time resolution. Each snapshot is viewed as a static network. The centrality score per node and snapshot is then obtained. The overall centrality score of a node is obtained as the average of the scores across snapshots [27,28]. Although such methods allow better identification of important nodes compared to static centrality methods, they may lose temporal information, such as the time order of contacts.

We propose a temporal gravity model to identify important nodes in temporal networks. Two main elements in universal gravitation are the masses of objects and the distance between them. Our main assumption is that the centrality of a node depends on its gravitation to nearby nodes, which is determined from the temporal distance. Hence, nearby nodes should be close to target nodes in both structure and time. In addition, we use node properties, such as static centrality metrics and their extension to temporal networks, to represent mass and the temporal distance between nodes to represent distance. The temporal distance between nodes captures both the structure and time order of contacts. Specifically, we use two definitions for temporal distance between two nodes: the fastest arrival distance and the temporal shortest distance. We use the temporal gravity model to identify important nodes with structural influence and spreading influence in 10 temporal networks. We use the network efficiency to determine the reference structural influence and the susceptible–infected–recovered (SIR) model to describe spreading in the temporal networks. The node spreading capacity, which is the range of spreading caused by a node, determines the reference spreading influence. Next, we obtain the Kendall correlation between the node reference influence and the importance score obtained from a centrality metric. A higher correlation coefficient indicates higher performance of the centrality method to identify important nodes. Experimental results demonstrate that the temporal gravity model considerably outperforms state-of-the-art centrality methods for important node identification.

The remainder of this paper is organised as follows. In Section 2, we describe the representation of a temporal network and the definition of the temporal distance between nodes. We briefly describe the static and temporal centrality metrics, which correspond to node mass in the proposed temporal gravity model, along with baseline metrics. In Section 3, we detail the proposed temporal gravity model based on baseline centrality metrics. In Section 4, we report experimental results obtained from the temporal gravity model and baseline metrics on real temporal networks. We also introduce a novel metric, the time degree, that represents the mass of the proposed temporal gravity model, improves robustness, and reduces computational complexity. In Section 5, we further analyse the performance of the proposed metrics on synthetic temporal networks. In Section 6, we draw conclusions from our study.

2. Preliminaries

In this section, we present basic concepts about temporal networks, including their representation, temporal paths, and distance. We then briefly describe benchmark centrality metrics. Centrality metrics and the temporal distance are the bases of the proposed temporal gravity model for important node identification.

2.1. Basic notations and definitions

Let $G^T = (V, E^T)$ represent a temporal network on time interval $[1, T]$. The network consists of a set V of $N = |V|$ nodes and a set of temporal events E^T. Each event $e \in E^T$ is given by a three-tuple (v_i, v_j, t), denoting that node v_i and node v_j make contact at time t. At each $t \in [1, T]$, the adjacency matrix is A_t, where $A_t(i, j) = 1$ represents that nodes v_i and v_j are connected at time t, and $A_t(i, j) = 0$ represents no connection between the nodes.

We can generate networks at various time scales according to the time resolution of network data. For example, email exchange datasets are usually collected in seconds. By setting the time resolution to 1 h, we can represent hourly data on the connection between two users within that period. We denote the time resolution as Δt. Temporal network G^T with $n = T/\Delta t$ snapshots can represent a dataset. The network snapshots are given by $G_1, G_2, \ldots G_n$. If Δt is small, the temporal network has several snapshots. If $\Delta t = T$, we obtain the corresponding static network of G^T, denoted as $G = (V, E)$. A pair of nodes v_i, v_j is connected by a link $(v_i, v_j) \in E$ if the nodes have at least one contact in G^T. The adjacency matrix of G is denoted as A, where $A(i, j) = 1$ if nodes v_i and v_j are connected, and $A(i, j) = 0$ otherwise. Each snapshot of G^T can be considered as a static network within the period corresponding to Δt.

Let us consider the example of a temporal network shown in Fig. 1. Fig. 1(b) shows a temporal network with five nodes and $T = 4$ time steps. By setting $\Delta t = 1$, the temporal network contains four snapshots, G_1, G_2, G_3, and G_4. Fig. 1(a) shows the corresponding aggregated static network G.

2.1.1. Temporal path

Given temporal network $G^T = (V, E^T)$ with n snapshots, a temporal path is a node sequence $P =< v_1, v_2, \ldots, v_k, v_{k+1} >$, where event $(v_i, v_{i+1}, t_i) \in E^T$ is the ith temporal event on P for $1 \le i \le k$ and $t_i \le t_{i+1}$. Hence, t_1 is the initial time of P, denoted as $t_{start}(P)$, and t_k is the final time of P, denoted as $t_{end}(P)$. We define the *temporal path length* of P as $l(P) = t_{end}(P) - t_{start}(P) + 1$. Given a time interval $[t_a, t_b]$, v_i is the initial node, and v_j is the final node $\forall v_i, v_j \in V$. Let $\mathbf{P}(v_i, v_j, [t_a, t_b]) = \{P | P$ be a temporal path from v_i to v_j, such that $t_{start}(P) \ge t_a$ and $t_{end}(P) \le t_b\}$. We consider two different definitions of temporal paths: the fastest arrival path and the temporal shortest path [29]. These paths are representative in temporal networks.

Fastest arrival path [29] The fastest arrival path between initial node v_i and final node v_j is the temporal path with the minimum duration counted from $t = 1$. Hence, the fastest arrival path is the path from initial node v_i to final node v_j with the minimum elapsed time over a period. Therefore, $P \in \mathbf{P}(v_i, v_j, [t_a, t_b])$ is the fastest arrival path if $t_{end}(P) = \min\{t_{end}(P') | P' \in \mathbf{P}(v_i, v_j, [t_a, t_b])\}$. The fastest arrival distance, $\varphi(v_i, v_j)$, between nodes v_i and v_j is the path length of the corresponding fastest arrival path.

Temporal shortest path [29] The temporal shortest path between initial node v_i and final node v_j is the path for which the overall time needed to communicate is the shortest. In other words, $P \in \mathbf{P}(v_i, v_j, [t_a, t_b])$ is the temporal shortest path if $l(P) = \min\{l(P') | P' \in \mathbf{P}(v_i, v_j, [t_a, t_b])\}$. The temporal shortest distance, $\theta(v_i, v_j)$, between nodes v_i and v_j is the path length of the corresponding temporal shortest path.

Let us illustrate the calculation of the temporal paths as shown in Fig. 1(c) and (d). The fastest path from node 1 to 4 is $P_1 =< 1, 2, 3, 4 >$. The fastest arrival distance, $l(P_1)$, between nodes 1 and 4 is $\varphi(1, 4) = 3$. On the other hand, the temporal shortest path

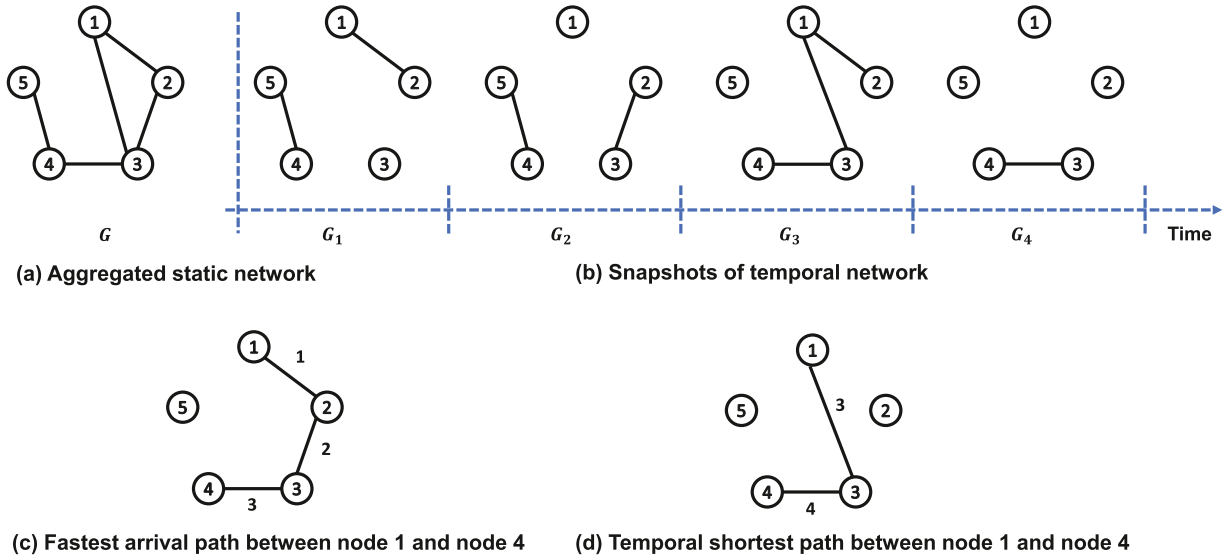

Fig. 1. Temporal network $G^T = (V, E^T)$ with five nodes and $T = 4$ time steps. (a) Aggregated static network G. (b) Static snapshots G_1, G_2, G_3, and G_4. (c) Fastest arrival path from initial node 1 to final node 4. (d) Temporal shortest path from initial node 1 to final node 4. The corresponding fastest arrival distance and temporal shortest distance are 3 and 2, respectively.

Table 1
Fastest arrival distance between nodes in the temporal network shown in Fig. 1. Node pairs without fastest arrival paths are denoted by ∞.

Node	1	2	3	4	5
1	0	1	2	3	∞
2	1	0	2	3	∞
3	3	2	0	3	∞
4	∞	∞	3	0	1
5	∞	∞	3	1	0

Table 2
Temporal shortest distance between nodes in the temporal network shown in Fig. 1. Node pairs without temporal shortest paths are denoted by ∞.

Node	1	2	3	4	5
1	0	1	1	2	∞
2	1	0	1	2	∞
3	1	1	0	1	∞
4	∞	∞	1	0	1
5	∞	∞	2	1	0

from node 1 to node 4 is $P_2 =< 1, 3, 4 >$, with the temporal shortest distance being $\theta(1, 4) = 2$. The fastest arrival distance and temporal shortest distance of every pair of nodes in Fig. 1 are listed in Tables 1 and 2, respectively.

2.2. Benchmark centrality metrics

We briefly introduce existing centrality metrics to identify important nodes.

Degree centrality. The degree of a node is defined on static network G as the number of its neighbours [10]. The degree centrality (DC) of a node is the proportion of nodes it is connected to. A higher degree implies greater importance of a node. The DC is defined as

$$DC(i) = \frac{k_i}{N-1}, \tag{1}$$

where k_i is the degree of node v_i, and N is the number of nodes in the network.

Closeness centrality. The closeness centrality (CC) [11] measures the distance between nodes. In practice, the CC determines the speed of communication of a node with all other nodes in a network. For a disconnected network, the CC is calculated as the sum of the reciprocal of the shortest distances from a given node v_i to all other nodes in static network G:

$$CC(i) = \frac{1}{N-1}\sum_{i\neq j}\frac{1}{d_{ij}}, \tag{2}$$

where d_{ij} is the shortest distance between v_i and v_j in G.

Betweenness centrality. The betweenness centrality (BC) [12] measures the number of shortest paths passing through node v_i in static network G:

$$BC(i) = \sum_{s\neq i\neq t}\frac{\sigma_{st}^{i}}{\sigma_{st}}, \tag{3}$$

where σ_{st} is the number of shortest paths between nodes v_s and v_t, and σ_{st}^{i} is the number of shortest paths between nodes v_s and v_t through node v_i.

PageRank. The PageRank (PR) [13] was introduced by Google to measure the importance of webpages from their hyperlink network structure. The PR measures the importance of a webpage (node) considering its neighbours and the number of pages (nodes) linked to each neighbour defined on static network G. Explicitly, PR is defined as

$$PR(i)^t = \sum_{j=1}^{N}(a_{ij}\frac{PR(j)^{t-1}}{k_j^{out}}), \tag{4}$$

where k_j^{out} is the out degree of node v_j, and a_{ij} represents the connection between nodes v_i and v_j. $PR(i)^t$ is the PR value of node v_i at time step t. After several iterations, the PR value gradually converges and stabilises. We use $PR(i)$ to represent the final PR value of node v_i.

Gravity model. The gravity centrality [16] of a node v_i is defined on static network G as

$$g(i) = \sum_{v_j\in\phi_i}\frac{ks_i ks_j}{d_{ij}^2}, \tag{5}$$

Table 3
Centrality scores per node obtained from temporal snapshots (row Averaged) and aggregated static network (row Aggregated). The temporal network and corresponding static network are show in Fig. 1.

Node	Type	PageRank	Degree	Closeness	Betweenness
1	Averaged	0.163	0.188	0.203	0.083
	Aggregated	0.192	0.500	0.571	0.000
2	Averaged	0.178	0.188	0.219	0.000
	Aggregated	0.192	0.500	0.571	0.000
3	Averaged	0.250	0.250	0.266	0.083
	Aggregated	0.283	0.750	0.800	0.667
4	Averaged	0.265	0.250	0.281	0.000
	Aggregated	0.213	0.500	0.667	0.500
5	Averaged	0.145	0.125	0.125	0.000
	Aggregated	0.120	0.500	0.444	0.000

where ks_i is the k-shell index of node v_i, ϕ_i is the neighbour set of node v_i, and d_{ij} is the shortest distance between nodes v_i and v_j.

Local gravity centrality The local gravity centrality [17] of node v_i in static network G is defined as

$$g_R(i) = \sum_{d_{ij} \le R, i \ne j} \frac{k_i k_j}{d_{ij}^2}, \tag{6}$$

where k_i is the degree of node v_i, and d_{ij} is the shortest distance between nodes v_i and v_j. We use R to truncate the contribution of high-order nodes on the centrality score of target node v_i. Specifically, $d_{ij} \le R$ indicates that nodes within distance R to v_i contribute to centrality score $g_R(i)$. This truncation represents a trade-off in node centrality between the local and global network structures.

2.3. Centrality metrics on temporal networks

Because each snapshot of a temporal network can be viewed as a static network, we can compute the centrality score per node in each snapshot. The centrality of a node in the temporal network can then be defined as the average centrality score across snapshots [27,30]. Consider the DC as an example. We can compute the DC of node v_i per snapshot to obtain an n-dimensional sequence of DC scores. The DC of node v_i in the temporal network is the average of this sequence. For a temporal network with n snapshots $G_1, G_2, \ldots, G_n$, we can analogously compute the four centrality scores, namely, PR, DC, CC, and BC, per node. For clarity, we denote the corresponding average centralities of a temporal network divided into snapshots as PR^m, DC^m, CC^m, and BC^m. We also compute the centrality scores on the aggregated network and denote them as PR^s, DC^s, CC^s, and BC^s to indicate the centrality scores of aggregated static network G.

The centrality scores per node of the temporal network shown in Fig. 1 are listed in Table 3, where 'Averaged' and 'Aggregated' indicate the corresponding centrality scores. The centrality scores for the metrics obtained from the temporal network by snapshot averaging considerably differ from those obtained from the static network.

3. Temporal gravity model for node ranking

The classical law of gravitation establishes a relation containing the mass of objects in the numerator and the distance between two objects in the denominator. Gravity models, such as the k-shell-based gravity model and local gravity model, have been used to identify important nodes in static networks. Inspired by the concept of gravity and existing gravity models for static networks, we propose a temporal gravity model to identify important nodes in temporal networks. In the temporal gravity model, the importance of a node depends on both its temporal distance to other nodes and its structural properties.

In the proposed *temporal gravity model*, we use the node properties as the mass and the distance between two nodes on a temporal network as the distance. Thus, the node importance of v_i is defined as follows:

$$TG(i) = \sum_{d_{ij} \le R, i \ne j} \frac{M_i M_j}{d_{ij}^2}, \tag{7}$$

where TG denotes the temporal gravity model, M_i represents the node properties of v_i, d_{ij} is the temporal distance between nodes v_i and v_j, and R is a truncation radius.

As node properties, we use baseline centrality metrics PR^s, DC^s, CC^s, BC^s, PR^m, DC^m, CC^m, and BC^m to represent mass M_i of node v_i. For the temporal distance (denominator of Eq. (7)), we consider either the fastest arrival distance (*TG-fad* model) or the temporal shortest distance (*TG-std* model).

We denote the temporal gravity model as function $TG(x, y)$, where $x \in \{fad, std\}$ and $y \in \{PR^m, DC^m, CC^m, BC^m, PR^s, DC^s, CC^s, BC^s\}$. For example, if PR^m represents the mass and *FAD* represents the distance in Eq. (7), the corresponding model is denoted as $TG(fad, PR^m)$.

4. Evaluation of temporal gravity model on real temporal networks

We use the centrality metrics described in Sections 2.2 and 2.3 as baseline metrics to evaluate the temporal gravity model. We use the network efficiency and SIR spreading model on temporal networks for performance evaluation. The network efficiency aims to determine the role of a node in information exchange, whereas the SIR spreading model aims to evaluate the spreading capacity of a node.

The node importance score obtained from different centrality metrics and the performance evaluation methods—namely, network efficiency and SIR spreading model—are compared by using Kendall correlation coefficient τ.[1] A high Kendall correlation coefficient τ indicates that the centrality metric suitably identifies important nodes in a temporal network.

We first define the network efficiency on temporal networks and present the results of identifying structural influence nodes in temporal networks by using centrality metrics and the temporal gravity model. We then compare the temporal gravity model and the baseline centrality metrics on the identification of important nodes during SIR spreading in temporal networks.

4.1. Network efficiency

For network efficiency [32], we assume that information in a network is transmitted only through the temporal shortest paths. The efficiency measures the quality of information exchange over a network. We define the network efficiency of temporal network G^T as follows:

$$\varepsilon\left(G^T\right) = \frac{1}{N(N-1)} \sum_{v_i \ne v_j \in G^T} \frac{1}{d_{ij}}, \tag{8}$$

[1] Kendall's tau [31] is a measure of the correlation strength between two sequences. A larger tau indicates higher similarity between the sequences. Consider two sequences with N elements, $X = (x_1, x_2, \ldots, x_N)$ and $Y = (y_1, y_2, \ldots, y_N)$. Any pair of two-tuples (x_i, y_i) and $(x_j, y_j)(i \ne j)$ is concordant if both $x_i > x_j$ and $y_i > y_j$ or both $x_i < x_j$ and $y_i < y_j$. The pair is discordant if $x_i > x_j$ and $y_i < y_j$ or $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ or $y_i = y_j$, the pair is neither concordant nor discordant. Kendall's tau of two sequences X and Y can be calculated as

$$\tau = \frac{1}{N(N-1)} \sum_{i \ne j} sgn(x_i - x_j) sgn(y_i - y_j).$$

Table 4
Kendall correlation coefficient τ between node centrality score obtained from NE_{fad} and centrality metrics for 10 empirical networks. The highest τ of each network is highlighted in bold and with an asterisk. The highest τ of each network obtained from baseline metrics is highlighted in bold.

	HS2011	HS2012	HS2013	WP	HC	PS	HT2009	Infections	SFHH	DNC
$TG(fad, PR^m)$	0.73206	0.72502	0.63329	0.68323	0.88180	**0.62505***	0.79741	**0.83589***	**0.82366***	**0.70045***
$TG(fad, PR^s)$	0.73206	**0.75233***	**0.67115***	0.74104	0.87459	0.49604	0.78729	0.79583	0.77739	0.62305
$TG(fad, DC^m)$	0.71632	0.69758	0.62841	0.72193	0.78883	0.57745	0.72408	0.69442	0.75628	0.63410
$TG(fad, DC^s)$	0.71606	0.73979	0.66849	**0.75108***	0.85225	0.46723	0.77655	0.71594	0.74235	0.60856
$TG(fad, CC^m)$	0.73333	0.69944	0.63108	0.73053	0.81189	0.52711	0.75411	0.73038	0.79462	0.68422
$TG(fad, CC^s)$	**0.74375***	0.72539	0.64233	0.68849	**0.91063***	0.59473	**0.80468***	0.83317	0.80798	0.63972
$TG(fad, BC^m)$	0.69973	0.65347	0.61902	0.66444	0.59495	0.46867	0.62263	0.54512	0.62040	0.33123
$TG(fad, BC^s)$	0.63378	0.62702	0.61839	0.72241	0.75063	0.37945	0.66498	0.50254	0.62421	0.34447
PR^m	0.41029	0.38423	0.37716	0.30244	0.30523	0.33336	0.33028	0.07638	0.42005	0.37230
DC^m	0.51515	0.40689	0.41068	0.31342	0.32924	**0.36287**	0.35987	0.09910	0.41199	**0.53762**
CC^m	**0.53041**	**0.41043**	0.41014	0.32011	0.30306	0.26374	**0.41056**	0.16953	**0.45180**	0.50284
BC^m	0.51660	0.34806	0.37426	0.27800	0.27712	0.28027	0.27497	**0.19491**	0.33410	0.32488
PR^s	0.46210	0.34364	0.40947	0.39943	0.37946	0.22883	0.36315	0.13729	0.36444	0.17510
DC^s	0.47604	0.35421	**0.41230**	**0.41010**	0.38060	0.22845	0.36582	0.11060	0.36446	0.37313
CC^s	0.44594	0.30643	0.39309	0.40781	**0.38671**	0.22781	0.36740	0.17343	0.36301	0.37683
BC^s	0.43975	0.27940	0.38872	0.37458	0.37225	0.23398	0.35240	0.17433	0.35885	0.33477
max of all metrics	0.74375	0.75233	0.67115	0.75108	0.91063	0.62505	0.80468	0.83589	0.82366	0.70045
max of baseline metrics	0.53041	0.41043	0.41230	0.41010	0.38671	0.36287	0.41056	0.19491	0.45180	0.53762

where d_{ij} is the temporal distance between nodes v_i and v_j. The temporal distance between two nodes can be defined by either the fastest arrival distance or the temporal shortest distance.

Removing a node from a temporal network may decrease the network efficiency if the network becomes disconnected. Therefore, the efficiency after node removal can reflect the importance of nodes in temporal networks [28]. A larger efficiency reduction indicates a higher node importance in terms of structural influence. Let $G^T \setminus v_i$ denote the temporal network after removing node v_i and all contacts associated with it. The difference between the network efficiency of G^T and that of $G^T \setminus v_i$ is defined as the importance score of node v_i regarding network efficiency:

$$NE(v_i) = \varepsilon(G^T) - \varepsilon(G^T \setminus v_i). \tag{9}$$

The node efficiencies based on fastest arrival distance and temporal shortest distance are denoted as NE_{fad} and NE_{std}, respectively.

We analyse the performance of the temporal gravity model and the baseline centrality metrics to identify important nodes regarding network efficiency by using the fastest arrival distance as the distance in the temporal gravity model.

In the temporal gravity model, we first use the baseline centrality metrics to represent mass. Taking PR centrality as an example, temporal gravity model $TG(fad, PR^m)$ integrates the PR score of nodes within a certain temporal distance as the centrality score of the target node. PR centrality PR^m can also be independently used as a centrality metric. We report the results from 10 empirical temporal networks (see Appendix Table A1). Fig. 2 shows that the temporal gravity model improves important node identification compared with the use of baseline centrality metrics PR^m, DC^m, CC^m, and BC^m for the temporal network snapshots and PR^s, DC^s, CC^s, and BC^s for the aggregated static networks. The temporal gravity model based on different baseline centrality metrics is more effective than the simple use of the corresponding baseline centrality metrics for important node identification in all evaluated networks.

Table 4 lists the Kendall correlation coefficient τ between the node centrality scores derived by the corresponding centrality metrics and the node efficiency based on fastest arrival distance NE_{fad}. In general, the temporal gravity model better identifies important nodes than baseline centrality metrics $PR^m, DC^m, CC^m, BC^m, PR^s, DC^s, CC^s$, and BC^s. Except for temporal gravity models $TG(fad, BC^m)$ and $TG(fad, BC^s)$ in network *DNC*, the Kendall correlation coefficients of the temporal gravity models are higher than those obtained from the baseline centrality metrics. Specifically, the highest τ of the temporal gravity model is 85.47% higher on average than the highest τ obtained from the baseline centrality metrics across the 10 empirical networks. Remarkably, τ of network *Infectious* increases from 0.19491 to 0.83589 from the best baseline metric to the best of temporal gravity model, representing an improvement of 328.85%.

Fig. 2 and Table 4 show the results considering the fastest arrival distance. Using the temporal gravity model based on the shortest distance to estimate NE_{std} is similar to using the model based on the fastest arrival distance to estimate NE_{fad}. Compared with using only the baseline centrality metrics (Fig. 3), the corresponding temporal gravity model improves the important node identification, except when compared with the aggregated DC in networks HT2009 and HS2011. Even under this condition, the performance of the proposed $TG(std, DC^s)$ is comparable with that of aggregated DC DC^s.

The results of important node identification based on the temporal shortest distance are listed in Table 5, which shows that the proposed model can suitably identify important nodes. The highest τ of the temporal gravity model is 5.79% higher on average than the highest τ obtained from the baseline centrality metrics across the 10 empirical networks. In network DNC, τ increases from 0.65626 to 0.78852 from the best baseline metric to the best temporal gravity model, representing an improvement of 20.15%.

4.2. Performance evaluation based on spreading capacity

A node is important if it originates information that spreads to a large population. We call such a node a seed node given its high spreading capacity. We evaluate the performance of the proposed temporal gravity model and baseline metrics for important node identification in temporal networks in terms of the spreading capacity. We use the SIR model to simulate information spreading in temporal networks [28,33]. In the SIR spreading model, a node can be susceptible (S), infected (I), or recovered (R), as illustrated in Fig. 4. A susceptible node can become infected after contact with an infected node with probability β. An infected node can become recovered with probability μ. The spreading process follows the flow of temporal networks. In the reported experiments, the infection and recovery probabilities remain fixed to $\beta = 0.1$ and $\mu = 0.01$, respectively.

In a temporal network with n snapshots $G_1, G_2, \ldots, G_n$, a node can appear in multiple snapshots. Therefore, if we choose a seed node, we should consider the time to start spreading. For node v_i, we assume the node appears at time $T_{v_i} = \{t^1_{v_i}, t^2_{v_i}, \ldots, t^m_{v_i}\}$. We take every time step $t^j_{v_i} \in T_{v_i}$ as the starting time of the spread-

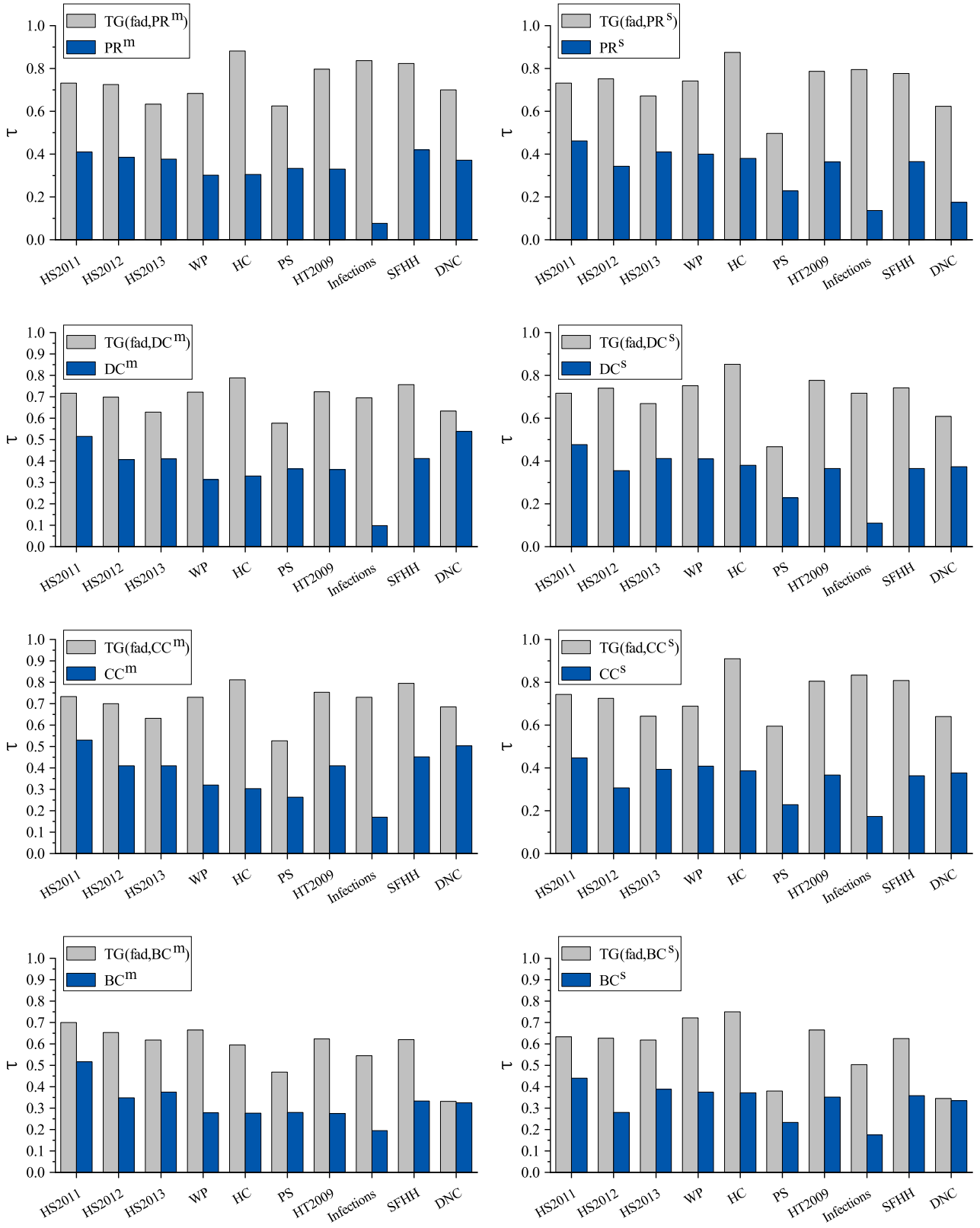

Fig. 2. Kendall correlation coefficient τ between the node importance score obtained from NE_{fad} and centrality metrics. The correlation histograms between NE_{fad} and the temporal gravity model/baseline centrality metrics are shown in grey/blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 3. Kendall correlation coefficient τ between the node importance score obtained from NE_{std} and centrality metrics. The correlation histograms between NE_{std} and the temporal gravity model/baseline centrality metrics are shown in grey/blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
Kendall correlation coefficient τ between node centrality score obtained from NE_{std} and centrality metrics for 10 empirical networks. The highest τ of each network is highlighted in bold and with an asterisk. The highest τ of each network obtained from baseline metrics is highlighted in bold.

	HS2011	HS2012	HS2013	WP	HC	PS	HT2009	Infections	SFHH	DNC
$TG(std, PR^m)$	0.83238	0.81142	0.80938	**0.87864***	0.82270	0.86413	0.86504	0.67578	0.83927	0.73534
$TG(std, PR^s)$	0.83365	0.83439	0.81674	0.86670	0.87459	0.91646	0.92225	0.67077	**0.86971***	0.66210
$TG(std, DC^m)$	0.77397	0.74811	0.71640	0.76684	0.84649	0.81907	0.83850	0.64911	0.85884	0.77186
$TG(std, DC^s)$	0.80800	0.82297	0.79362	0.85714	0.86883	0.90954	0.91561	0.69403	0.86551	0.62968
$TG(std, CC^m)$	0.73003	0.71471	0.72199	0.76684	0.83063	0.83375	0.83123	0.72909	0.81936	**0.78852***
$TG(std, CC^s)$	**0.83721***	**0.85102***	**0.84184***	0.85714	**0.90126***	**0.92943***	**0.93268***	**0.78484***	0.86917	0.76614
$TG(std, BC^m)$	0.76830	0.70524	0.76131	0.71031	0.64108	0.64137	0.79709	0.67989	0.80382	0.35997
$TG(std, BC^s)$	0.80952	0.75353	0.77982	0.80889	0.82991	0.88094	0.90076	0.53887	0.83619	0.37066
PR^m	0.60000	0.64184	0.60432	0.62255	0.65766	0.66051	0.73925	0.35974	0.65145	0.25222
DC^m	0.74402	0.70797	0.65638	0.66571	0.80846	0.73185	0.79285	0.54285	0.83115	0.53272
CC^m	0.66425	0.64544	0.56811	0.66699	0.71748	0.73410	0.75190	**0.66781**	0.66695	**0.65626**
BC^m	0.69516	0.63571	0.70018	0.62986	0.57477	0.53719	0.75474	0.58768	0.75721	0.35106
PR^s	0.81689	0.78994	**0.77404**	0.83421	0.83423	0.90686	0.92004	0.52716	0.85573	0.14369
DC^s	**0.82287**	**0.80619**	0.77157	**0.85886**	0.84518	0.91203	0.92330	0.62512	**0.86124**	0.41227
CC^s	0.80315	0.75970	0.73217	0.76988	**0.85556**	**0.91556**	**0.92542**	0.62666	0.85401	0.42378
BC^s	0.75594	0.68674	0.73460	0.78261	0.77802	0.83670	0.87389	0.47615	0.80663	0.35896
max of all metrics	0.83721	0.85102	0.84184	0.87864	0.90126	0.92943	0.93268	0.78484	0.86971	0.78852
max of baseline metrics	0.82287	0.80619	0.77404	0.85886	0.85556	0.91556	0.92542	0.66781	0.86124	0.65626

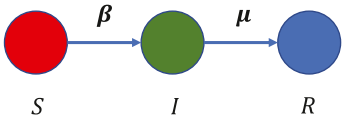

Fig. 4. Diagram of SIR spreading model.

ing. Let node v_i be the seed and $t^j_{v_i}$ be the starting time. We run SIR spreading until the end of the temporal network to obtain the final spreading range, $R^j_{v_i}$, which describes the infected and recovered nodes after spreading. For each $t^j_{v_i}$ as the starting time, spreading proceeds over 100 trials to obtain the average spreading range, $\overline{R^j_{v_i}}$. In addition, for each seed node v_i, spreading proceeds starting at time $t^j_{v_i}$ over 100 trials. Therefore, the spreading ranges are $R(v_i) = \{\overline{R^1_{v_i}}, \overline{R^2_{v_i}}, \ldots, \overline{R^m_{v_i}}\}$. We use two definitions for the node spreading capacity. First, the average spreading capacity of node v_i can be given by the average spreading range, $R_{mean}(v_i)$, over set $R(v_i)$. Second, the normalised spreading capacity can be given by $R_{norm}(v_i) = \frac{1}{m}\sum_{j=1}^{m} \frac{\overline{R^j_{v_i}}}{n - t^j_{v_i} + 1}$. A large R_{mean} or R_{norm} implies that the node has a high spreading capacity.

We evaluate the temporal gravity model and baseline centrality metrics for identifying nodes with high spreading capacity in temporal networks. The temporal gravity model considers the temporal shortest distance. The real temporal capacity of a node is defined by R_{mean} or R_{norm}. Taking the temporal gravity model as an example, we determine the performance using the Kendall correlation coefficient. First, we compute the importance score of every node using the temporal gravity model to obtain a list of centrality scores per node. Then, spreading proceeds to determine a list of spreading capacities per node. Kendall correlation coefficient τ is computed between the lists of centrality scores and spreading capacities. A high value of τ indicates that the evaluated centrality metric can suitably identify important nodes.

The important node identification results are listed in Tables 6 and A2 for spreading capacities R_{norm} and R_{mean}, respectively. The spreading influence in the experiment is the normalised (or averaged) result of the nodes at different starting times. Therefore, the gravity model considering the shortest path

Table 6
Kendall correlation coefficient τ between real node spreading capacity R_{norm} and node centrality score obtained from centrality metrics for 10 empirical networks. The highest τ of each network is highlighted in bold and with an asterisk. The highest τ of each network obtained from baseline metrics is highlighted in bold.

	HS2011	HS2012	HS2013	WP	HC	PS	HT2009	Infections	SFHH	DNC
$TG(std, PR^m)$	0.59924	0.62359	0.60481	0.38939	0.54955	0.47882	0.70796	0.54029	0.61821	**0.65801***
$TG(std, PR^s)$	0.57283	0.56760	0.58575	0.30769	0.55604	0.46655	0.67636	0.53137	0.58379	0.59151
$TG(std, DC^m)$	**0.65714***	**0.66096***	**0.69359***	**0.42714***	0.58486	**0.51620***	**0.74210***	**0.57452***	**0.62799***	0.62994
$TG(std, DC^s)$	0.58425	0.59243	0.61126	0.29861	0.55315	0.46552	0.67573	0.54868	0.58192	0.61340
$TG(std, CC^m)$	0.62717	0.64444	0.59149	0.41376	0.50559	0.46408	0.67794	0.49205	0.59157	0.61191
$TG(std, CC^s)$	0.57917	0.57480	0.57854	0.31199	0.55027	0.46518	0.68173	0.48330	0.59718	0.65043
$TG(std, BC^m)$	0.51381	0.58320	0.47167	0.33186	0.50775	0.40146	0.69216	0.37630	0.54656	0.40272
$TG(std, BC^m)$	0.46057	0.43633	0.44639	0.30148	0.52937	0.43658	0.65518	0.29761	0.54942	0.38472
PR^m	0.41867	0.57666	0.47059	**0.42379**	0.48108	0.43918	0.66688	0.36852	0.52167	0.24533
DC^m	**0.63226**	**0.63544**	**0.66220**	0.41901	**0.60029***	**0.51156**	**0.73904**	0.55896	**0.61955**	0.45733
CC^m	0.59416	0.59814	0.49674	0.41567	0.42559	0.44062	0.64096	0.48547	0.51259	**0.47610**
BC^m	0.43812	0.52671	0.42033	0.33058	0.49045	0.35181	0.68173	0.35857	0.51717	0.39356
PR^s	0.53752	0.51173	0.53172	0.31151	0.55243	0.45551	0.66941	0.48628	0.56537	0.16401
DC^s	0.56493	0.56181	0.58323	0.31907	0.55945	0.46196	0.67751	**0.56921**	0.57046	0.41924
CC^s	0.49522	0.49158	0.45245	0.30237	0.55753	0.45885	0.67698	0.32204	0.56927	0.46352
BC^s	0.41816	0.38331	0.40110	0.30387	0.50198	0.41868	0.63401	0.24377	0.52846	0.37387
max for all	0.65714	0.66096	0.69359	0.42714	0.60029	0.51620	0.74210	0.57452	0.62799	0.65801
max for based methods	0.63226	0.63544	0.66220	0.42379	0.60029	0.51156	0.73904	0.56921	0.61955	0.47610

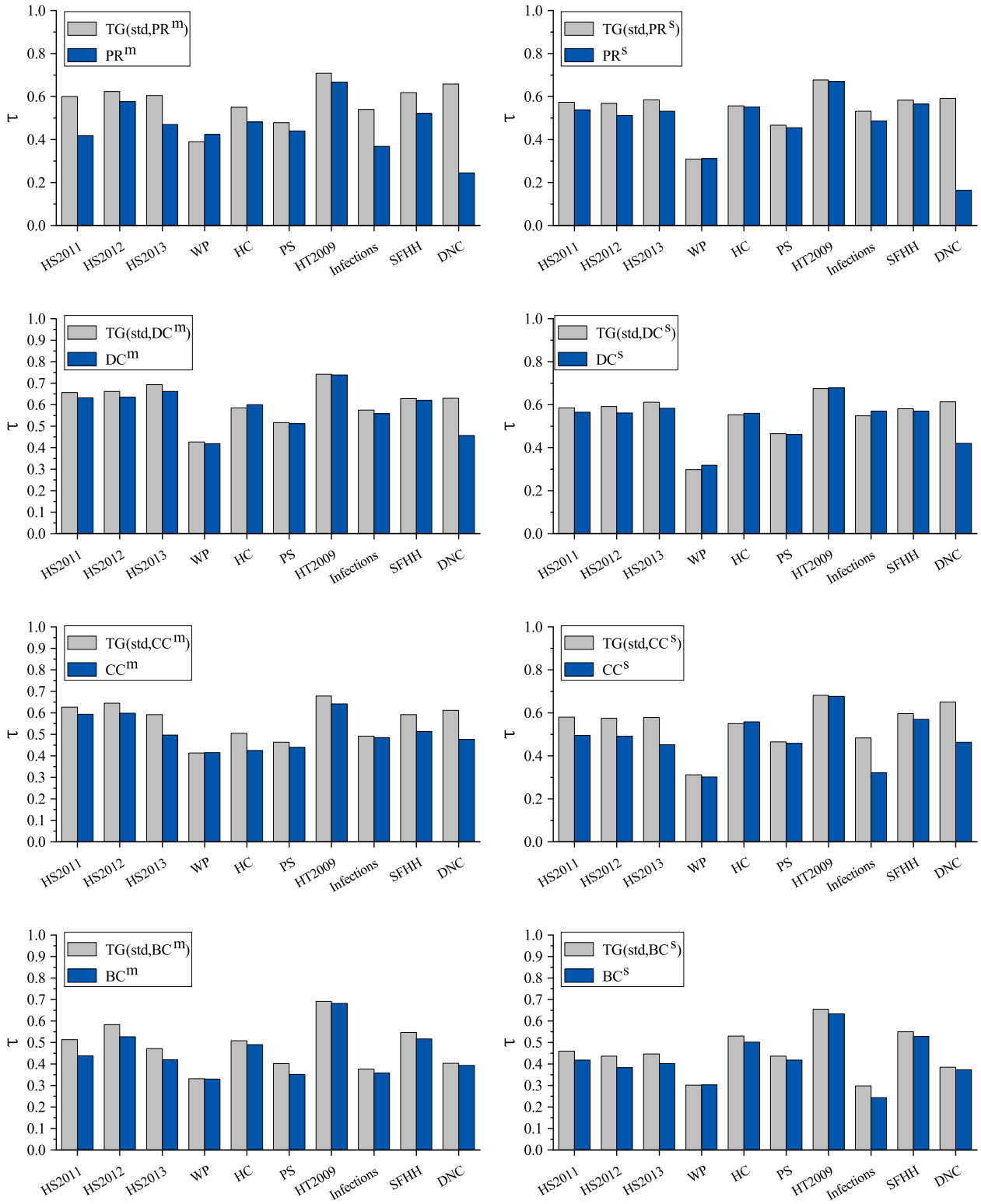

Fig. 5. Kendall correlation coefficient τ between real node spreading capacity R_{norm} and the node centrality score obtained from centrality metrics. The correlation histograms between R_{norm} and the temporal gravity model/baseline centrality metrics are shown in grey/blue. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

is more suitable for this type of influence. To reduce the effect of noise from neighbouring nodes, we set the truncation radius of temporal gravity model *TG-std* to $R = 5$ (see Fig. A1). Fig. 5 shows that *TG-std* based on the baseline centrality metrics outperforms the use of baseline centrality metrics to identify important nodes. In fact, temporal gravity model *TG-std* presents high-performance important node identification in most temporal networks. Even for some datasets for which *TG-std* fails to provide the highest performance, its Kendall correlation coefficients remain high.

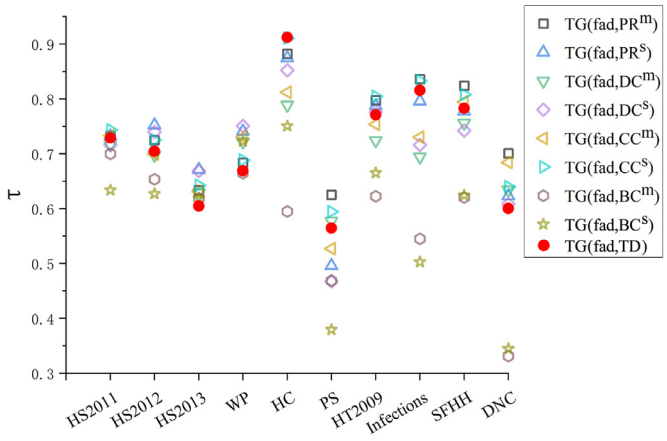

Fig. 6. Kendall correlation coefficient τ between the node centrality score obtained from NE_{fad} and temporal gravity models for 10 temporal empirical networks.

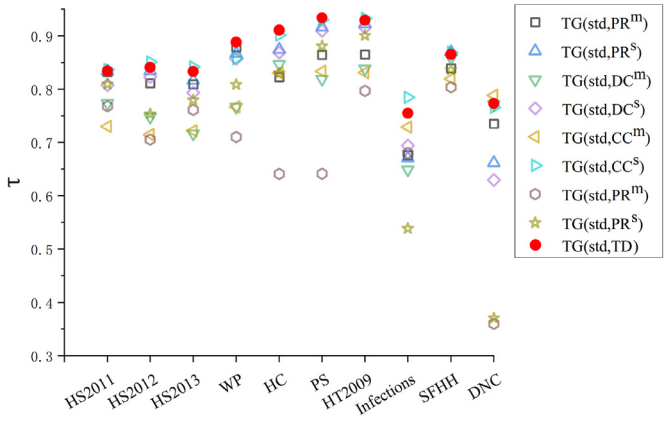

Fig. 7. Kendall correlation coefficient τ between the node centrality score obtained from NE_{std} and temporal gravity models for 10 empirical networks.

4.3. Efficient node centrality metric

By considering temporal information to quantify the mass, we propose a node degree for temporal networks, called the time degree, which can be used as the mass of a node in Eq. (7). The time degree is defined below.

- Time degree
  For a temporal network with n snapshots, $G_1, G_2, \ldots, and\ G_n$, the degree centrality values of node v_i for the snapshots are given by $DC(1), DC(2), \ldots, and\ DC(n)$, respectively. We define the time degree of node v_i as

  $$TD(i) = e^{DC(1)} + e^{DC(2)} + \ldots + e^{DC(n)}.$$

The temporal gravity model can consider both the structural properties and time information to obtain an overall improvement in important node identification. Considering the influence of the degree of a node at different snapshots, we obtain the time degree. Among the temporal gravity models, the one based on time degree TD has relatively stable performance on the NE_{fad} prediction, as shown in Fig. 6.

Among the temporal gravity models, the one based on the time degree, $TG(std, TD)$, also has relatively stable and robust performance, as shown in Fig. 7. $TG(std, TD)$ is more effective than the baseline centrality metrics in the 10 networks. In general, when measuring the structural influence of a temporal network, we can use the temporal gravity model based on the time degree to reduce both the selection burden of indicators and computational complexity.

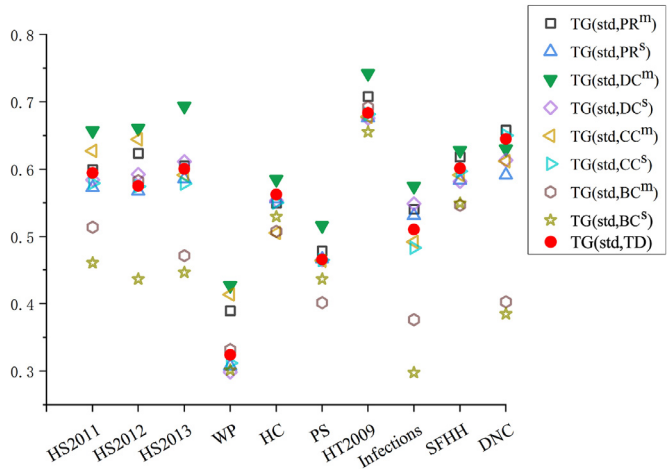

Fig. 8. Kendall correlation coefficient τ between the node centrality score obtained from R_{norm} and temporal gravity models for 10 empirical networks.

The spreading influence has a high correlation with the node degree, and a higher degree implies a faster spreading. The average degree can better reflect the overall state over time than the aggregated degree. Compared with the proposed *TG-std* methods, $TG(std, DC^m)$ more accurately predicts the node spreading influence (see Fig. 8), and $TG(std, TD)$ also provides stable performance. When measuring the spreading influence of temporal networks, we can use $TG(std, TD)$ or $TG(std, DC^m)$ to reduce the selection burden of indicators.

5. Performance analysis in activity-driven network models

Based on the analysis in Section 4, the temporal gravity model achieves an overall improvement compared with the baseline centrality metrics. For structural influence, the temporal gravity model based on time degree TD shows a steady improvement. For spreading influence, $TG(std, DC^m)$ and $TG(std, TD)$ are better predictors. Thus, we can use the temporal gravity model with TD or DC^m to reduce the selection burden of indicators. We further analyse the performance of the proposed metrics on synthetic temporal networks. To analyse the type of data for which the proposed method is more suitable, our empirical analysis naturally leads to the use of the activity-driven model [34] to generate the synthetic temporal networks.

The activity-driven network model considers N nodes (aggregated) and gives each node a fixed activity probability per unit time $a_i = \eta x_i$, which is defined as the probability to create edges with other nodes. Here, x_i (bounded in the interval $\epsilon \le x_i \le 1$) is the activity potential of node v_i, which is distributed according to power law distribution $F(x) \sim x^{-\gamma}$, and η is a rescaling factor that determines the average number of nodes per unit time, $\eta\langle x\rangle N$. A network evolves according to the following steps: (i) Each node has a fixed activity potential probability x_i. At the beginning of each time snapshot t, the N nodes of network G_t are disconnected initially. (ii) Through random seeding, each node becomes an active node with probability $a_i \triangle t$ and connects with m other nodes randomly. (iii) At the next snapshot $t + \triangle t$, all edges are removed, and step 2 is repeated.

The average degree per unit time of the temporal network is $\langle k\rangle_t = 2m\eta\langle k\rangle$. The values of link number m influence the average degree $\langle k\rangle_t$ and the network density. By changing the value of m, we can obtain experimental networks with different average degrees. Most social networks are heterogeneous. Without loss of generally, we fix the parameters $\epsilon = 10^{-3}$, $\triangle t = 1$, $\eta = 10$ and $\gamma = 2.5$. In order to better evaluate our method, we set random

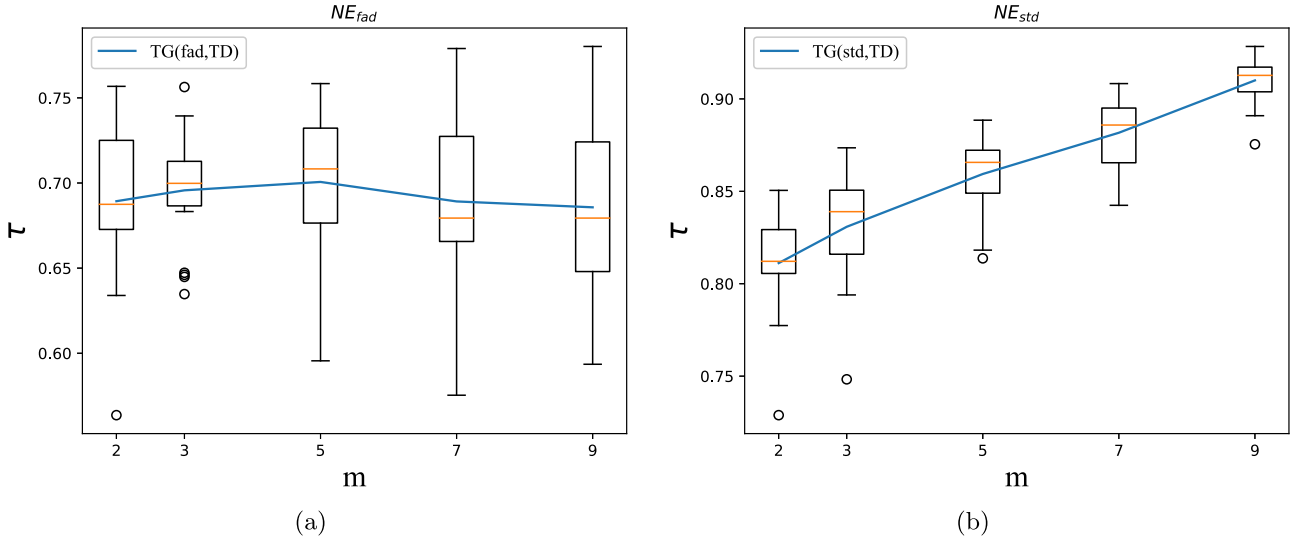

Fig. 9. Measures of correlations between NE_{fad} (or NE_{std}) and temporal gravity models under different parameter link number m in the empirical networks. We fix $N = 100$, $T = 80$, $\eta = 10$, $F(x) \sim x^{-\gamma}$ with $\gamma = 2.5$, and $\epsilon \leq x \leq 1$ with $\epsilon = 10^{-3}$. Under each m, we set random seeds to generate 20 temporal networks, and the coefficient result is the corresponding box. The blue line is the mean value under each m. (a) shows the different Kendall correlation coefficients τ between NE_{fad} and our method $TG(fad, TD)$ on the parameters m. (b) shows the different Kendall correlation coefficients τ between NE_{std} and our method $TG(std, TD)$ on the parameters m. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

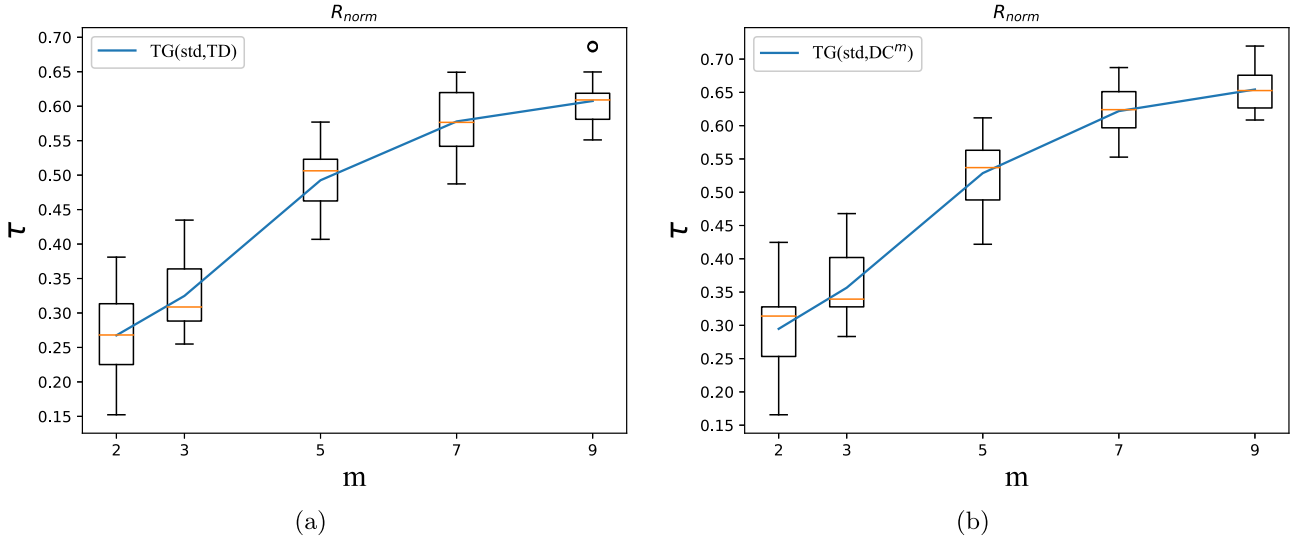

Fig. 10. Measures of correlations between R_{norm} and temporal gravity models under different parameter link number m in the empirical networks. We fix $N = 100$, $T = 80$, $\eta = 10$, $F(x) \sim x^{-\gamma}$ with $\gamma = 2.5$, and $\epsilon \leq x \leq 1$ with $\epsilon = 10^{-3}$. Under each m, we set random seeds to generate 20 temporal networks, and the coefficient result is the corresponding box. The blue line is the mean value under each m. (a) and (b) show the different Kendall correlation coefficients τ between R_{norm} and our method $TG(std, TD)$ (or $TG(std, DC^m)$) on the parameters m. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

seeds to generate 20 networks with the same average degree $\langle k \rangle_t$ under each link number m.

We examine how the parameter m influences our method identification quality. For network efficiency, the effect of the parameter m on the value of correlation τ is plotted in Fig. 9. As shown in Fig. 9(a), the value of correlation τ obtained by comparing NE_{fad} and $TG(fad, TD)$ is stable when m is increased. Fig. 9(a) indicates that m has little effect on important node identification in terms of network efficiency NE_{fad}. However, in Fig. 9(b), we can see that the τ obtained by comparing NE_{std} and $TG(std, TD)$ increases as m increases in general. The mean value of τ increases to 0.91000. Fig. 9 suggests that our method can also identify the structural influential nodes on synthetic networks. For NE_{std}, our method is more suitable for denser networks.

In terms of spreading influence, the evolution of the correlation τ between R_{norm} and temporal gravity models with increasing m is plotted in Fig. 10. The correlation τ increases greatly with increasing m. Therefore, the increase in the correlation when m is large implies that our model is more suitable for denser networks.

6. Discussion and conclusions

In practice, most complex systems are dynamic and time varying. To preserve the temporal information of systems, we can represent them as temporal networks. Although many centrality metrics have been proposed for static networks, important node identification in temporal networks remains an open question.

The law of gravitation is a simple, elegant, and representative formula to estimate the strength of interaction between objects by considering the inherent influence of the objects and their distance. Inspired by the concept of gravity and existing gravity models for static networks, we propose a temporal gravity model to

identify important nodes in temporal networks. The temporal gravity model leverages both neighbourhood information and temporal information. In addition, it provides a mathematical and computational framework that can use different node properties to represent the node masses and different temporal distances as distance analogous in gravity. To determine the node properties, we use baseline centrality metrics and the proposed time degree. For the temporal distance, we consider either the fastest arrival distance or the temporal shortest distance.

The Kendall correlation measures the ranking correlation between two variables. Hence, we use Kendall correlation coefficient τ between the real influence of a node on either the network connectivity or spreading and the importance score obtained from different centrality metrics to evaluate the metric performance on empirical temporal networks. To enhance the demonstration results, we also use the Spearman correlation ρ[2] to quantify the performance (see Appendix B). In addition, we consider the network efficiency based on the temporal shortest path, NE_{std}, and that based on the fastest arrival path, NE_{fad}. Regarding NE_{fad}, $TG-fad$ outperforms the baseline centrality metrics across different networks. Specifically, the Kendall correlation coefficient τ of the temporal gravity model increases by 85.51% on average compared with the highest τ obtained from baseline centrality metrics across the 10 empirical networks. Regarding NE_{std}, the $TG-std$ methods provide an overall improvement. For structural influence, the proposed temporal gravity model based on time degree TD shows a steady improvement over the baseline centrality metrics. When measuring the structural influence of temporal networks, we can thus use the temporal gravity model with TD to reduce the selection burden of indicators.

Regarding the spreading influence, we simulate the SIR spreading model on the empirical temporal networks. We use the normalised and average spreading ranges to represent the node spreading capacity. The $TG-std$ methods using the baseline centrality metrics (average scores) outperform the baseline centrality metrics. $TG(std, DC^m)$ and $TG(std, TD)$ are better predictors for node spreading influence. We also use the activity-driven model to generate synthetic temporal networks for further analysis. As network density changes, our method is stable for NE_{fad}. For NE_{std} and R_{norm}, our method is more suitable for denser networks. Overall, the temporal gravity model provides robust performance for important node identification across networks. The temporal gravity model using baseline centrality metrics to represent mass outperforms the corresponding baseline centrality metrics, and the model has superior robustness for important node identification in temporal networks. Considering the suitability of the gravity model for important node identification in both static and temporal networks, we will extend the model to other types of networks, such as multi-layer networks [36,37] and bipartite networks [2]. Moreover, the spreading influence of a node can vary if we consider different dynamical processes. Thus, we can further explore important node identification related to different spreading models, such as susceptible–infected–susceptible [6] and coevolution spreading processes [38], for temporal networks.

[2] Spearman's rank correlation [35] coefficient ρ is a nonparametric measure of rank correlation. The closer ρ to 1, the stronger the association between the two ranks. Consider two sequences with N elements. Start by ranking the two sequences. Data ranking can be achieved by assigning the ranks '1' to the largest number, '2' to the second-largest number, and so forth. Calculate the difference between ranks denoted as 'd'. The coefficient ρ can be depicted in the formula

$$\rho = 1 - \frac{6\sum d^2}{N(N^2-1)}.$$

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Jialin Bi: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Writing - review & editing, Visualization. **Ji Jin:** Software, Validation, Formal analysis, Investigation, Data curation. **Cunquan Qu:** Methodology, Writing - original draft, Writing - review & editing, Supervision, Funding acquisition. **Xiuxiu Zhan:** Conceptualization, Methodology, Investigation, Writing - review & editing. **Guanghui Wang:** Resources, Writing - review & editing, Funding acquisition. **Guiying Yan:** Methodology, Writing - review & editing.

Acknowledgements

The authors were supported in part by National Natural Science Foundation of China under Grants 11631014, 11871311, 12001324, in part by China Postdoctoral Science Foundation under Grants 2019TQ0188, 2019M662315, in part by Shandong University multidisciplinary research and innovation team of young scholars under Grand 2020QNQT017, in part by Taishan Scholars Program Foundation of Shandong Province, China. We thank the SocioPatterns collaboration (http://www.sociopatterns.org) and the KONECT (http://konect.uni-koblenz.de/networks) for providing the datasets. The scientific calculations in this paper have been done on the HPC Cloud Platform of Shandong University.

Appendix A. Empirical networks

We evaluate the performance of the temporal gravity model on the following empirical temporal network datasets. Some of the detailed properties of the networks are shown in Table A1.

- High school 2011 (2012,2013) dynamic contact networks [39,40] (HS2011, HS2012, HS2013). These datasets correspond to the contacts and friendship relations between students in a high school in Marseilles, France.
- Workplace (WP) [41]. This dataset contains contacts between employees in an office building in France from June 24 to July 3, 2013.
- Hospital contract (HC) [42]. This dataset contains contacts between patients, contacts between patients and health-care workers (HCWs), and contacts between HCWs in a hospital ward in Lyon, France, from December 6 to December 10, 2010.

Table A1
Property description of the empirical networks. The number of nodes N, the number of snapshots n, the total number of contacts $|C|$, and the number of links $|E|$ in aggregated static network G are shown.

Network	N	n	$\|C\|$	$\|E\|$
HS2011	126	76	28,561	1710
HS2012	180	203	45,047	2239
HS2013	327	101	188,508	5818
WP	92	275	9827	755
HC	75	97	32,424	1139
PS	242	65	125,773	8317
HT2009	113	118	20,818	2196
Infectious	410	79	17,298	2765
SFHH	403	64	70,261	9889
DNC	1760	71	38,484	5428

Table A2
Kendall correlation coefficient τ between R_{mean} and temporal network methods for 10 empirical networks. The highest τ of each network is highlighted in bold and with an asterisk. The highest τ of each network obtained from centrality metrics is highlighted in bold. The SIR model parameters are $\beta = 0.1$ and $\mu = 0.01$. For each node at one occurrence, the spreading range is the average across 100 independent trials.

	HS2011	HS2012	HS2013	WP	HC	PS	HT2009	Infections	SFHH	DNC
$TG(std, TD)$	0.63911	0.51421	0.47716	0.50024	0.38667	0.26072	0.49905	0.67884	0.44923	**0.57545***
$TG(std, PR^m)$	0.58324	0.54364	0.47205	0.55996	0.38234	0.28830	0.50569	0.60043	0.46883	0.56827
$TG(std, PR^s)$	0.60356	0.48939	0.46147	0.48782	0.39604	0.26710	0.48546	0.63600	0.43293	0.47618
$TG(std, DC^m)$	0.65486	**0.58150***	**0.54140***	0.63497	**0.40613***	0.28452	0.52465	0.67760	0.47096	0.54119
$TG(std, DC^s)$	0.63276	0.50726	0.47881	0.48065	0.39892	0.26100	0.48609	0.66925	0.43066	0.49881
$TG(std, CC^m)$	**0.68737***	0.57418	0.48740	0.63306	0.37441	0.25188	**0.53445***	**0.68208***	**0.47380***	0.52520
$TG(std, CC^s)$	0.63429	0.49460	0.45772	0.47874	0.39171	0.26024	0.49210	0.67898	0.44513	0.56842
$TG(std, BC^m)$	0.52118	0.52410	0.35925	0.51105	0.32036	0.20805	0.47977	0.59671	0.41199	0.35615
$TG(std, BC^s)$	0.46489	0.36655	0.35056	0.41949	0.38523	0.29131	0.47440	0.39009	0.41302	0.34179
PR^m	0.37625	0.49472	0.34519	0.59245	0.34847	**0.29296***	0.45828	0.29588	0.41575	0.30063
DC^m	0.57574	**0.57435**	**0.51316**	0.63260	**0.40513**	0.29065	0.51590	0.60771	**0.46870**	**0.43246**
CC^m	**0.61879**	0.55742	0.42230	**0.64262***	0.33766	0.24145	**0.52339**	**0.64752**	0.44884	0.35385
BC^m	0.43279	0.48946	0.31272	0.48835	0.31027	0.18597	0.46492	0.51072	0.39357	0.34816
PR^s	0.51390	0.44022	0.41712	0.50215	0.39243	0.27348	0.47724	0.50993	0.42064	0.15792
DC^s	0.57084	0.47402	0.45397	0.50700	0.39369	0.26684	0.48043	0.62944	0.42280	0.36675
CC^s	0.54144	0.39203	0.35006	0.40685	0.38962	0.26424	0.48079	0.51059	0.42174	0.37786
BC^s	0.41994	0.30460	0.31135	0.41328	0.38955	0.28795	0.45575	0.31524	0.40204	0.33171
max of all metrics	0.68737	0.58150	0.54140	0.64262	0.40613	0.29296	0.53445	0.68208	0.47380	0.57545
max of baseline metrics	0.61879	0.57435	0.51316	0.64262	0.40513	0.29296	0.52339	0.64752	0.46870	0.43246

Table A3
Spearman correlation coefficient ρ between node centrality score obtained from NE_{fad} and centrality metrics for 10 empirical networks. The highest ρ of each network is highlighted in bold and with an asterisk. The highest ρ of each network obtained from baseline metrics is highlighted in bold.

	HS2011	HS2012	HS2013	WP	HC	PS	HT2009	Infections	SFHH	DNC
$TG(fad, TD)$	0.89210	0.85391	0.78106	0.83655	**0.98020***	0.74414	0.84907	0.91782	0.91859	0.76046
$TG(fad, PR^m)$	0.89477	0.86688	0.80890	0.84742	0.97425	**0.80148***	0.86698	**0.92729***	**0.94350***	0.78780
$TG(fad, PR^s)$	0.89597	**0.88765***	**0.84395***	0.89430	0.97229	0.67005	**0.86778***	0.91846	0.92399	0.76071
$TG(fad, DC^m)$	0.88242	0.85916	0.80747	0.88680	0.93542	0.75832	0.83399	0.85746	0.91290	**0.84307***
$TG(fad, DC^s)$	0.88457	0.88114	0.84204	**0.90155***	0.96512	0.63562	0.86418	0.87202	0.90150	0.77276
$TG(fad, CC^m)$	0.88894	0.85621	0.80523	0.88703	0.94703	0.70498	0.85199	0.87931	0.93370	0.83320
$TG(fad, CC^s)$	**0.90326***	0.86671	0.81462	0.85344	0.97994	0.77464	0.86571	0.92466	0.93350	0.80276
$TG(fad, BC^m)$	0.87852	0.82275	0.80359	0.85278	0.79661	0.64223	0.76799	0.71367	0.80801	0.41180
$TG(fad, BC^s)$	0.82626	0.80814	0.80551	0.89520	0.91084	0.53600	0.80259	0.66977	0.81105	0.42559
PR^m	0.56408	0.53428	0.53317	0.43988	0.43713	0.48051	0.46716	0.10137	0.58231	0.52376
DC^m	0.70179	0.56574	**0.57988**	0.45075	0.46463	**0.51896**	0.50617	0.15626	0.56972	0.66341
CC^m	**0.71955**	**0.56895**	0.57955	0.46400	0.41775	0.37823	**0.56863**	0.25018	**0.63461**	**0.66878**
BC^m	0.69036	0.49843	0.52836	0.40935	0.40492	0.40192	0.39139	**0.27832**	0.47304	0.40347
PR^s	0.63744	0.48947	0.57461	0.55676	0.51562	0.33650	0.49800	0.20387	0.51336	0.25229
DC^s	0.65079	0.49785	0.57251	0.56031	0.51250	0.33458	0.49896	0.16761	0.51141	0.46669
CC^s	0.60741	0.43467	0.55125	**0.57205**	**0.51925**	0.33371	0.50269	0.23409	0.51175	0.51012
BC^s	0.60748	0.40673	0.54798	0.53564	0.51166	0.34457	0.49094	0.25647	0.50766	0.41213
max of all metrics	0.90326	0.88765	0.84395	0.90155	0.98020	0.80148	0.86778	0.92729	0.94350	0.84307
max of baseline metrics	0.71955	0.56895	0.57988	0.57205	0.51925	0.51896	0.56863	0.27832	0.63461	0.66878

- Primary school (PS) [43]. This dataset contains contacts between the children and teachers used in the study published in BMC Infectious Diseases 2014.
- Hypertext2009 (HT2009) [44]. This network contains contacts between the attendees of the ACM Hypertext 2009 conference.
- Infectious [44]. This network contains contacts between people during the exhibition INFECTIOUS: STAY AWAY in 2009 at the Science Gallery in Dublin.
- SFHH conference (SFHH) [45]. This dataset contains contacts between participants in the 2009 SFHH conference in Nice, France.
- DNC Email (DNC) [46]. This is the network of emails in the 2016 Democratic National Committee email leak.

We evaluate the temporal gravity model and baseline centrality metrics for identification of nodes with high spreading capacity in temporal networks.

The effect of the truncation radius on the temporal gravity model is illustrated in Fig. A1.

The results considering average spreading range R_{mean} are listed in Table A2.

Appendix B. Evaluation of temporal gravity model based on spearman correlations

We also use the Spearman correlations to quantify the performance. The results based on real data are shown in Table A3, A4, A5 and A6, and the results of experimental data are shown in Fig. A2 and A3. In terms of network efficiency, the effect of the parameter m on the value of Spearman correlation coefficient ρ is plotted in Fig. A2. As shown in Fig. A2(a), the value of ρ obtained by comparing NE_{fad} and $TG(fad, TD)$ is stable when m is increased. The average value of coefficient ρ is 0.85980. In Fig. A2(b), we can see that the ρ obtained by comparing NE_{std} and $TG(std, TD)$ increases as m increases in general. The mean value of ρ increases to 0.98583. In terms of spreading influence, the evolution of the correlation ρ between R_{norm} and temporal gravity models with increasing m is plotted in Fig. A3. The correlation ρ increases with increasing m.

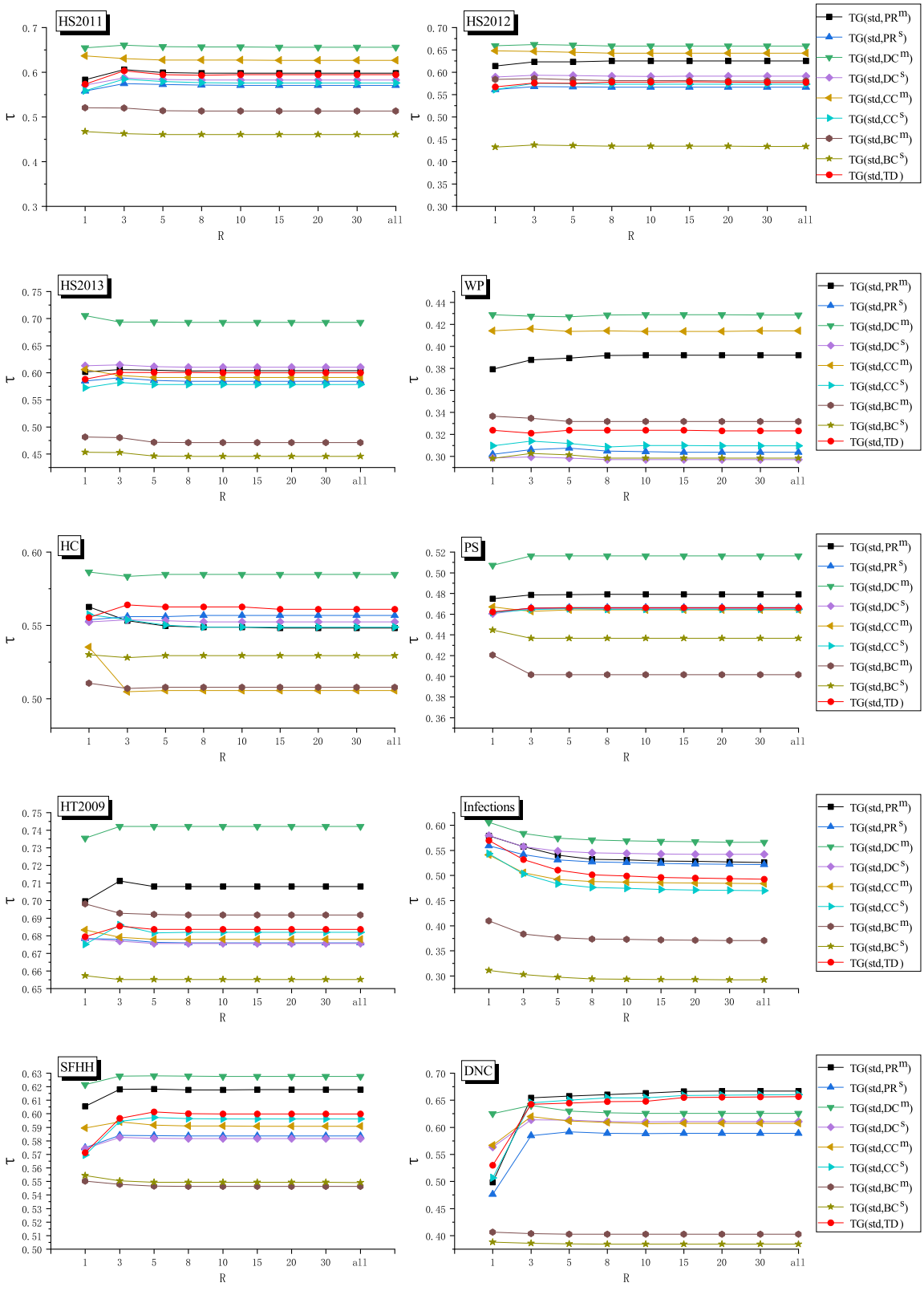

Fig. A1. Kendall correlation coefficient τ between node centrality score obtained from R_{norm} and temporal gravity models with truncation radius from 1 to the maximum for 10 empirical networks. A stable value can be reached at a truncation radius of 5.

Table A4
Spearman correlation coefficient ρ between node centrality score obtained from NE_{std} and centrality metrics for 10 empirical networks. The highest ρ of each network is highlighted in bold and with an asterisk. The highest ρ of each network obtained from baseline metrics is highlighted in bold.

	HS2011	HS2012	HS2013	WP	HC	PS	HT2009	Infections	SFHH	DNC
$TG(std, TD)$	0.95891	0.95280	0.95765	**0.98029***	**0.98723***	**0.99022***	0.99045	0.91283	0.97386	0.89259
$TG(std, PR^m)$	0.95560	0.93474	0.94720	0.97656	0.95055	0.97241	0.97002	0.85468	0.96508	0.89206
$TG(std, PR^s)$	**0.96255***	0.95048	0.95021	0.97121	0.97323	0.98724	0.98970	0.85094	**0.97602***	0.78065
$TG(std, DC^m)$	0.92091	0.89756	0.88468	0.92032	0.96555	0.95175	0.96092	0.83969	0.97343	0.86079
$TG(std, DC^s)$	0.94899	0.94450	0.93676	0.96746	0.97183	0.98569	0.98846	0.87344	0.97439	0.80556
$TG(std, CC^m)$	0.88984	0.87643	0.88925	0.91865	0.95550	0.95961	0.95114	0.89438	0.95287	**0.90310***
$TG(std, CC^s)$	0.96097	**0.95646***	**0.96090***	0.96827	0.98336	0.98930	**0.99178***	**0.92978***	0.97526	0.89262
$TG(std, BC^m)$	0.92644	0.87271	0.91996	0.88777	0.83087	0.80704	0.94358	0.86082	0.94942	0.44549
$TG(std, BC^s)$	0.94708	0.91035	0.93202	0.94493	0.95220	0.98016	0.98430	0.72301	0.96267	0.45722
PR^m	0.79327	0.82617	0.78940	0.81581	0.79414	0.84995	0.90297	0.50032	0.84173	0.36366
DC^m	0.90348	0.87322	0.83897	0.84660	0.94584	0.89724	0.94041	0.73322	0.96103	0.64780
CC^m	0.84560	0.82142	0.75641	0.84279	0.88731	0.90705	0.90375	**0.85255**	0.85205	**0.81203**
BC^m	0.87242	0.81801	0.87529	0.82003	0.76267	0.70620	0.91891	0.78334	0.92240	0.43366
PR^s	0.95508	0.93315	**0.92780**	0.96113	0.95565	0.98538	0.98895	0.70667	0.97116	0.21815
DC^s	**0.95521**	**0.93510**	0.92112	**0.96444**	0.95968	0.98517	0.98874	0.80480	**0.97137**	0.50834
CC^s	0.94639	0.90941	0.90031	0.92037	**0.96465**	**0.98642**	**0.98930**	0.81744	0.97019	0.56723
BC^s	0.91523	0.86490	0.90212	0.92562	0.90993	0.96304	0.97665	0.64494	0.94899	0.44082
max of all metrics	0.96255	0.95646	0.96090	0.98029	0.98723	0.99022	0.99178	0.92978	0.97602	0.90310
max of baseline metrics	0.95521	0.93510	0.92780	0.96444	0.96465	0.98642	0.98930	0.85255	0.97137	0.81203

Table A5
Spearman correlation coefficient ρ between R_{norm} and temporal network methods for 10 empirical networks. The highest ρ of each network is highlighted in bold and with an asterisk. The highest ρ of each network obtained from centrality metrics is highlighted in bold. The SIR model parameters are $\beta = 0.1$ and $\mu = 0.01$. For each node at one occurrence, the spreading range is the average across 100 independent trials.

	HS2011	HS2012	HS2013	WP	HC	PS	HT2009	Infections	SFHH	DNC
$TG(std, TD)$	0.75275	0.75858	0.79315	0.44725	0.72498	0.64320	0.86375	0.67608	0.78621	0.83366
$TG(std, PR^m)$	0.74707	0.79454	0.79046	0.52347	0.71260	0.65411	0.87700	0.71876	0.79572	**0.84327***
$TG(std, PR^s)$	0.73838	0.75659	0.78075	0.42976	0.72407	0.64529	0.85842	0.71158	0.76795	0.77537
$TG(std, DC^m)$	**0.79282***	**0.82724***	**0.86617***	0.56889	0.76048	**0.69212***	**0.89937***	0.74084	**0.80681***	0.80949
$TG(std, DC^s)$	0.74815	0.77998	0.80557	0.41720	0.71929	0.64523	0.85699	0.71787	0.76690	0.79411
$TG(std, CC^m)$	0.76427	0.80596	0.76868	0.55278	0.68267	0.63552	0.84393	0.64515	0.76611	0.79887
$TG(std, CC^s)$	0.73952	0.75948	0.77210	0.43355	0.71750	0.64246	0.86184	0.63591	0.78224	0.83871
$TG(std, BC^m)$	0.68202	0.77290	0.65366	0.47191	0.70014	0.56248	0.87399	0.53029	0.73246	0.50906
$TG(std, BC^s)$	0.62748	0.61338	0.62312	0.42853	0.70993	0.61418	0.84415	0.42789	0.73161	0.49103
PR^m	0.56806	0.74539	0.64129	**0.58669***	0.65693	0.60246	0.84664	0.53855	0.69838	0.36150
BC^m	**0.77887**	**0.79989**	**0.84029**	0.56789	**0.77582***	**0.68489**	**0.89842**	0.73886	**0.79624**	0.58578
CC^m	0.73435	0.76380	0.67038	0.55404	0.59844	0.60333	0.81214	0.64917	0.68593	**0.64193**
BC^m	0.59851	0.71063	0.58999	0.47349	0.68262	0.49983	0.86296	0.51256	0.70359	0.49835
PR^s	0.71220	0.69939	0.72485	0.43321	0.72760	0.63505	0.85355	0.67410	0.75029	0.24999
DC^s	0.73525	0.74757	0.77203	0.43680	0.72770	0.63897	0.85576	**0.74373***	0.75312	0.52820
CC^s	0.66744	0.67295	0.62946	0.42352	0.72806	0.63704	0.85622	0.45170	0.75280	0.62938
BC^s	0.58110	0.54489	0.56924	0.43432	0.68558	0.59233	0.82926	0.35562	0.70912	0.47612
max of all metrics	0.79282	0.82724	0.86617	0.58669	0.77582	0.69212	0.89937	0.74373	0.80681	0.84327
max of baseline metrics	0.77887	0.79989	0.84029	0.58669	0.77582	0.68489	0.89842	0.74373	0.79624	0.64193

Table A6
Spearman correlation coefficient ρ between R_{mean} and temporal network methods for 10 empirical networks. The highest ρ of each network is highlighted in bold and with an asterisk. The highest ρ of each network obtained from centrality metrics is highlighted in bold. The SIR model parameters are $\beta = 0.1$ and $\mu = 0.01$. For each node at one occurrence, the spreading range is the average across 100 independent trials.

	HS2011	HS2012	HS2013	WP	HC	PS	HT2009	Infections	SFHH	DNC
$TG(std, TD)$	0.82404	0.68747	0.65958	0.68694	0.52976	0.33982	0.64676	0.86972	0.62829	0.72014
$TG(std, PR^m)$	0.76787	0.71877	0.64925	0.74912	0.53388	0.36463	0.65013	0.78898	0.64654	**0.72328***
$TG(std, PR^s)$	0.79257	0.66851	0.64172	0.66682	0.55021	0.35113	0.63442	0.83029	0.60733	0.64354
$TG(std, DC^m)$	0.83403	**0.75789***	**0.72489***	0.82035	**0.55269***	0.35061	**0.67426***	0.86847	**0.65143***	0.70729
$TG(std, DC^s)$	0.82069	0.68680	0.66187	0.65839	0.54982	0.34261	0.63291	0.86582	0.60501	0.66159
$TG(std, CC^m)$	**0.86166***	0.75060	0.65361	0.81479	0.50805	0.31376	0.67354	**0.87366***	0.64324	0.66670
$TG(std, CC^s)$	0.81677	0.67151	0.63561	0.65720	0.53733	0.33909	0.64192	0.86884	0.62264	0.71785
$TG(std, BC^m)$	0.69650	0.72059	0.51645	0.70939	0.47616	0.28217	0.63920	0.79522	0.58143	0.45087
$TG(std, BC^s)$	0.63711	0.52322	0.50173	0.60205	0.54259	0.39004	0.61885	0.55170	0.58107	0.43712
PR^m	0.52689	0.66865	0.49622	0.77615	0.49770	0.37503	0.60425	0.43545	0.58047	0.44111
DC^m	0.76074	**0.74803**	**0.69619**	0.81048	0.54875	0.35594	**0.66522**	0.80598	**0.64637**	**0.55556**
CC^m	**0.79954**	0.73225	0.57539	**0.82162***	0.45841	0.29600	0.65850	**0.84363**	0.60858	0.47836
BC^m	0.59769	0.67991	0.45558	0.68365	0.46629	0.25988	0.61969	0.70299	0.55927	0.44210
PR^s	0.70364	0.61446	0.58890	0.67110	0.55004	0.36500	0.62125	0.69861	0.59293	0.23725
DC^s	0.75635	0.64735	0.62842	0.66989	0.54520	0.35275	0.62295	0.81857	0.59333	0.46380
CC^s	0.72393	0.55265	0.49989	0.57252	0.54024	0.34696	0.62510	0.70874	0.59235	0.51640
BC^s	0.57444	0.43997	0.45161	0.59287	**0.55115**	**0.39155***	0.60043	0.45254	0.56507	0.42429
max of all metrics	0.86166	0.75789	0.72489	0.82162	0.55269	0.39155	0.67426	0.87366	0.65143	0.72328
max of baseline metrics	0.79954	0.74803	0.69619	0.82162	0.55115	0.39155	0.66522	0.84363	0.64637	0.55556

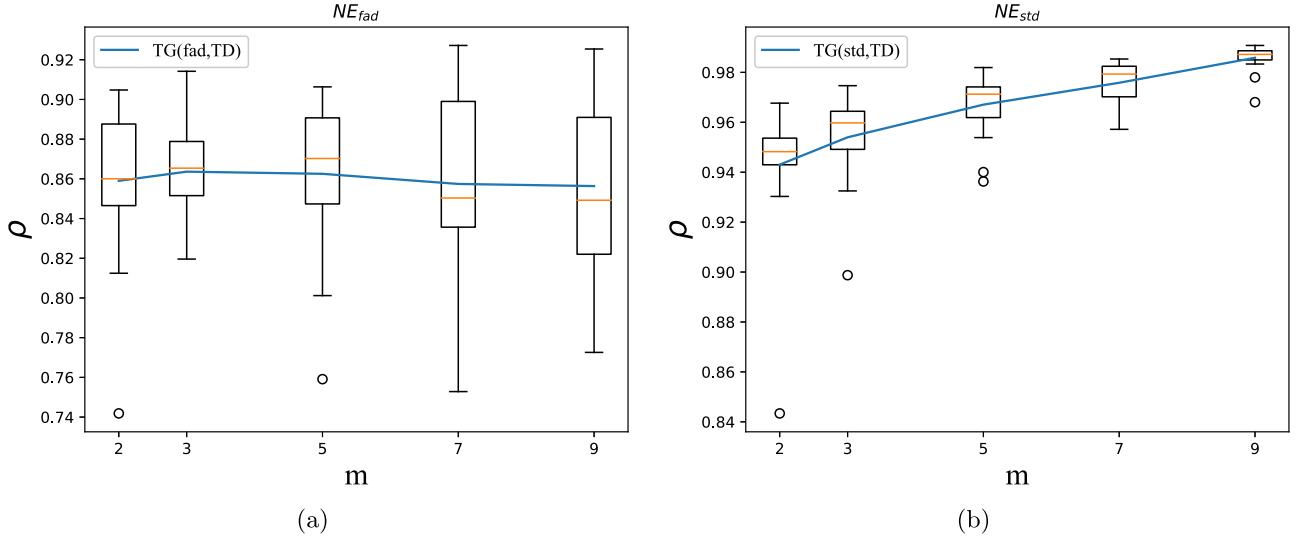

Fig. A2. Measures of correlations between NE_{fad} (or NE_{std}) and temporal gravity models under different parameter link number m in the empirical networks. We fix $N = 100$, $T = 80$, $\eta = 10$, $F(x) \sim x^{-\gamma}$ with $\gamma = 2.5$, and $\epsilon \leq x \leq 1$ with $\epsilon = 10^{-3}$. Under each m, we set random seeds to generate 20 temporal networks, and the coefficient result is the corresponding box. The blue line is the mean value under each m. (a) shows the different Spearman correlation coefficients ρ between NE_{fad} and our method $TG(fad, TD)$ on the parameters m. (b) shows the different Spearman correlation coefficients ρ between NE_{std} and our method $TG(std, TD)$ on the parameters m. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

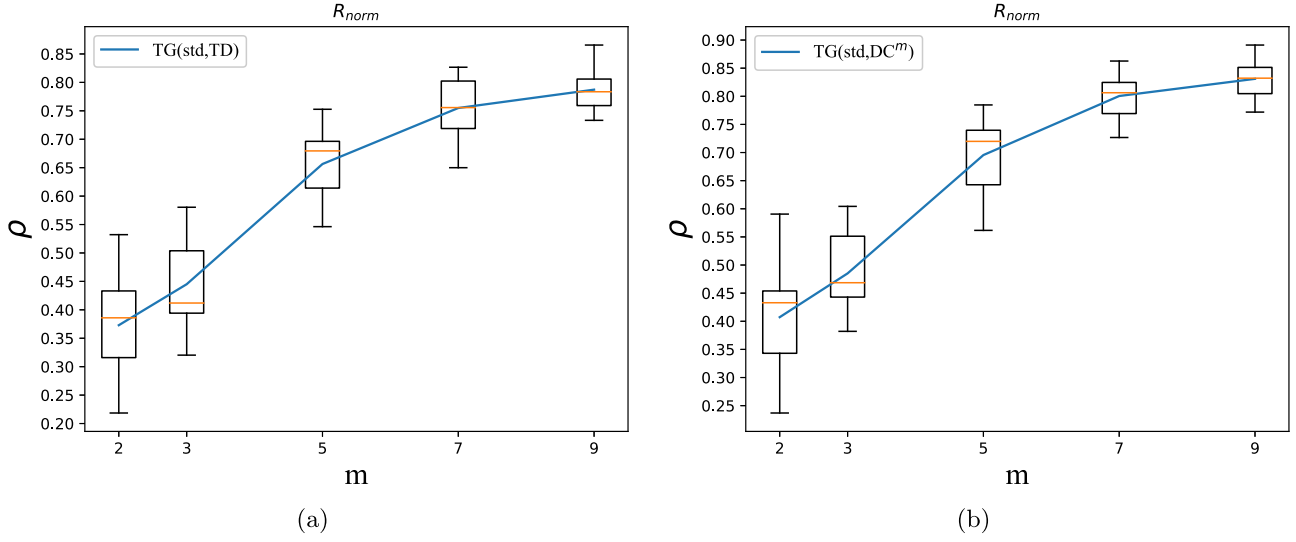

Fig. A3. Measures of correlations between R_{norm} and temporal gravity models under different parameter link number m in the empirical networks. We fix $N = 100$, $T = 80$, $\eta = 10$, $F(x) \sim x^{-\gamma}$ with $\gamma = 2.5$, and $\epsilon \leq x \leq 1$ with $\epsilon = 10^{-3}$. Under each m, we set random seeds to generate 20 temporal networks, and the coefficient result is the corresponding box. The blue line is the mean value under each m. (a) and (b) show the different Spearman correlation coefficients ρ between R_{norm} and our method $TG(std, TD)$ (or $TG(std, DC^m)$) on the parameters m. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

References

[1] Dame N. Statistical mechanics of complex networks. Rev Mod Phys 2002;74(1):xii.
[2] Newman M. Networks. Oxford university press; 2018.
[3] Morone F, Makse HA. Influence maximization in complex networks through optimal percolation. Nature 2015;524(7563):65–8.
[4] Tang J, Sun J, Wang C, Yang Z. Social influence analysis in large-scale networks. In: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining; 2009. p. 807–16.
[5] Zhang X, Zhu J, Wang Q, Zhao H. Identifying influential nodes in complex networks with community structure. Knowl Based Syst 2013;42:74–84.
[6] Zhan X-X, Liu C, Zhou G, Zhang Z-K, Sun G-Q, Zhu JJ, et al. Coupling dynamics of epidemic spreading and information diffusion on complex networks. Appl Math Comput 2018;332:437–48.
[7] Lü L, Chen D, Ren X-L, Zhang Q-M, Zhang Y-C, Zhou T. Vital nodes identification in complex networks. Phys Rep 2016;650:1–63.
[8] Gao Z, Shi Y, Chen S. Measures of node centrality in mobile social networks. Int J Mod Phys C 2015;26(9):1550107.
[9] Qiao T, Shan W, Yu G, Liu C. A novel entropy-based centrality approach for identifying vital nodes in weighted networks. Entropy 2018;20(4):261.
[10] Bonacich P. Factoring and weighting approaches to status scores and clique identification. J Math Sociol 1972;2(1):113–20.
[11] Sabidussi G. The centrality index of a graph. Psychometrika 1966;31(4):581–603.
[12] Freeman LC. A set of measures of centrality based on betweenness.. Sociometry 1977;40(1):35–41.
[13] Brin S, Page L. The anatomy of a large-scale hypertextual web search engine.. Comput Netw 1998;30:107–17.
[14] Kleinberg JM. Authoritative sources in a hyperlinked environment. J ACM 1999;46(5):604–32.
[15] Lempel R, Moran S. Salsa: the stochastic approach for link-structure analysis. ACM Trans Inf Syst 2001;19(2):131–60.
[16] Ma L-l, Ma C, Zhang H-F, Wang B-H. Identifying influential spreaders in complex networks based on gravity formula. Phys A 2016;451:205–12.
[17] Li Z, Ren T, Ma X, Liu S, Zhang Y, Zhou T. Identifying influential spreaders by gravity model. Sci Rep 2019;9(1):8387.

[18] Holme P. Analyzing temporal networks in social media. Proc IEEE 2014;102(12):1922–33.
[19] Takaguchi T, Sato N, Yano K, Masuda N. Importance of individual events in temporal networks. New J Phys 2012;14(9):093003.
[20] Holme P. Temporal network structures controlling disease spreading.. Phys Rev E 2016;94(2):22305.
[21] Scholtes I, Wider N, Pfitzner R, Garas A, Tessone CJ, Schweitzer F. Causality–driven slow-down and speed-up of diffusion in non-Markovian temporal networks. Nat Commun 2014;5(1):5024.
[22] Li A, Cornelius SP, Liu Y-Y, Wang L, Barabási A-L. The fundamental advantages of temporal networks. Science 2017;358(6366):1042–6.
[23] Holme P, Saramäki J. Temporal networks. Phys Rep 2012;519(3):97–125.
[24] Renaud L, Naoki M. A guide to temporal networks, 4. World Scientific; 2016.
[25] Tang J, Musolesi M, Mascolo C, Latora V. Characterising temporal distance and reachability in mobile and online social networks. ACM Spec Interest Group Commun 2010;40(1):118–24.
[26] Grindrod P, Parsons MC, Higham DJ, Estrada E. Communicability across evolving networks.. Phys Rev E 2011;83(4):46120.
[27] Kim H, Anderson R. Temporal node centrality in complex networks.. Phys Rev E 2012;85(2):26107.
[28] Qu C, Zhan X, Wang G, Wu J, Zhang Z-k. Temporal information gathering process for node ranking in time-varying networks. Chaos 2019;29(3):033116.
[29] Wu H, Cheng J, Huang S, Ke Y, Lu Y, Xu Y. Path problems in temporal graphs. In: Very large data bases, 7; 2014. p. 721–32.
[30] Pan RK, Saramäki J. Path lengths, correlations, and centrality in temporal networks. Phys Rev E 2011;84(1):16105.
[31] Kendall MG. A new measure of rank correlation. Biometrika 1938;30(1/2):81–93.
[32] Latora V, Marchiori M. Efficient behavior of small-world networks. Phys Rev Lett 2001;87(19):198701.
[33] Zhan X-X, Hanjalic A, Wang H. Information diffusion backbones in temporal networks. Sci Rep 2019;9(1):1–12.
[34] Perra N, Goncalves B, Pastor-Satorras R, Vespignani A. Activity driven modeling of time varying networks. Sci Rep 2012;2(1):469.
[35] Spearman C. General intelligence objectively determined and measured. Am J Psychol 1904;15(2):201–93.
[36] Kivelä M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA. Multilayer networks. J Complex Netw 2014;2(3):203–71.
[37] Boccaletti S, Bianconi G, Criado R, Del Genio CI, Gómez-Gardenes J, Romance M, et al. The structure and dynamics of multilayer networks. Phys Rep 2014;544(1):1–122.
[38] Wang W, Liu Q-H, Liang J, Hu Y, Zhou T. Coevolution spreading in complex networks. Phys Rep 2019;820:1–51.
[39] Fournet J, Barrat A. Contact patterns among high school students. PLoS One 2014;9(9):e107878.
[40] Mastrandrea R, Fournet J, Barrat A. Contact patterns in a high school: a comparison between data collected using wearable sensors, contact diaries and friendship surveys. PLoS One 2015;10(9):e0136497.
[41] Génois M., Vestergaard C.L., Fournet J., Panisson A., Bonmarin I., Barrat A.. Data on face-to-face contacts in an office building suggests a low-cost vaccination strategy based on community linkers. arXiv preprint arXiv:140970172014.
[42] Vanhems P, Barrat A, Cattuto C, Pinton J-F, Khanafer N, Régis C, et al. Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. PLoS One 2013;8(9):e73970.
[43] Gemmetto V, Barrat A, Cattuto C. Mitigation of infectious disease at school: targeted class closure vs. school closure. BMC Infect Dis 2014;14(1):695.
[44] Isella L, Stehlé J, Barrat A, Cattuto C, Pinton J-F, Van den Broeck W. What's in a crowd? Analysis of face-to-face behavioral networks. J Theor Biol 2011;271(1):166–80.
[45] Génois M, Barrat A. Can co-location be used as a proxy for face-to-face contacts? EPJ Data Sci 2018;7(1):11.
[46] Rossi R, Ahmed N. The network data repository with interactive graph analytics and visualization. In: Proceedings of the AAAI conference on artificial intelligence, 29; 2015.