

中国航空网络时序特征分析

牟建红, 黄 格, 吕 欣

(国防科技大学系统工程学院 长沙 410073)

【摘要】该文着重研究时序信息对中国航空网络拓扑结构和传播行为的影响。将网络分割为24个时间窗口, 研究拓扑结构在序列中的变化特性; 考虑网络中事件发生的时间间隔, 建立时序网络上的SIR传播模型以及基于时间特性的节点分类方法。结果表明, 度-集聚系数和介数-集聚系数在8:00时出现正相关向负相关的转变; 节点上航班出港行为的时间爆发性存在明显差异, 但整体上可降低网络的传播速度和感染范围。该研究成果对进一步认识航空网络及受航空网络影响的疾病传播等动力学过程具有重要意义。

关 键 词 航空网络; SIR传播模型; 时序网络; 时间窗口; 拓扑结构

中图分类号 O231.1 **文献标志码** A **doi:**10.3969/j.issn.1001-0548.2018.03.022

Analysis of the Temporal Characteristics of Chinese Aviation Network

MOU Jian-hong, HUANG Ge, and LÜ Xin

(College of Systems Engineering, National University of Defense Technology Changsha 410073)

Abstract This paper focuses on the influence of time series on the topology and spreading modes in Chinese aviation network. The network is divided into 24 time windows to study the characteristics of the topological structure in the sequence. Considering the time interval of events in this network, the SIR (susceptible-infective-removed) propagation model and the node classification method based on time characteristics are established. The results show that the correlations among the degree and clustering coefficient, betweenness and clustering coefficient have changed from positive to negative at 8:00am. And there are obvious differences among nodes in terms of the burstiness of exporting time related to flights, but it has reduced the propagation speed and the scope of infection. The results are of great significance for the further understanding of the aviation network and its dynamic process, such as the propagation of diseases.

Key words aviation network; SIR propagation model; temporal network; time window; topological structure

随着科技的发展和社会的进步, 人们空乘出行的需求逐步增长。我国民用机场数量从改革开放初期的69个到2006年增至136个, 目前仅大陆就拥有183个民用机场、28家航空企业。同时, 航空网络不再以单一机场为辐射中心, 形成了以北京、上海、广州和西安等大城市为枢纽的多中心网络^[1]。航空网络不断扩展, 网络对象交互行为异常复杂。这对认识航空运输发展的规律, 改善网络结构和提高网络运行的效率都是一个巨大挑战。复杂网络理论的兴起和后续研究的深入, 为认知航空网络提供了重要的手段^[2-3]。

国外航空网络领域着重运用复杂网络理论研究世界或国家范围内航空网络的拓扑结构、网络中的疾病传播, 同时也在寻找更加符合实际的网络模型。

文献[4]对2002年世界航空网络的研究结论表明, 航空网络具有小世界现象。文献[5]对印度的航空网络进行研究, 并将其与世界范围的航空网络进行比对, 定性描述了发展中国家不同于发达国家的航空网络演变过程。文献[6]基于有效距离的概念将复杂的时空网络简化为简单的同质流传播网络, 并成功运用到2009年H1N1和2003年SARS传染病病源的发现上。文献[7]提出随机交互网络模型(random interacting network, RAIN), 研究申根国家(Schengen)和美国航空网络之间的交互关系。文献[8]首次命名了交通流驱动的加权网络演化模型, 力求最大限度地贴近实际网络。

国内航空网络研究方向主要涉及城市体系、航线连接和网络结构分析等。文献[9]利用统计物理的

收稿日期: 2016-12-25; 修回日期: 2017-02-05

基金项目: 国家自然科学基金优秀青年基金(71522014); 国家自然科学基金重大项目(71690233); 国家自然科学基金面上项目(71771213)

作者简介: 牟建红(1993-), 男, 主要从事复杂网络理论与应用方面的研究。

方法对包含128个节点和1 165条边的中国航空网络进行实证分析,表明我国航空网络存在明显的小世界现象。文献[10-11]分别研究了1930~2012年和1952~2008年中国航空网络的演变,发现网络的连通性不断提高,呈现出的小世界网络趋势越发明显。文献[1]对我国航空网络的演化进行了分析,发现网络拓扑保持稳定,但节点之间的相对重要性发生变化且流量呈现出明显的季节性。文献[12]运用随机的演员模型分析城市体系和航空网络之间的协同发展,发现两者存在内在的偏好附属和外在的经济和区域影响以及共同促进发展的效果。文献[13]使用反重力场模型中的粒子群方法对中国航空网络进行实证研究,表明城市节点的吸引力与人口规模有正相关关系,且大城市的主导地位在下降,航空网络向更均衡方向发展。

然而,上述研究均集中于讨论航空网络的静态结构,忽视了航空网络中边与边之间的时序关系,即相同机场内相邻两次航班起飞的时间间隔。航空网络作为一个典型的时序网络^[14],其边的状态具有明显的间歇性,事件在网络上的传播也具有显著的时序性。研究者^[5, 15-17]多将时序网络中任意时间窗口内存在的边都视为对应静态网络中的有效边,扩大了边的存在时间,从而过度强调网络中节点的连接关系,低估了节点之间的传播距离。当网络拓扑结构的重要性远不及网络中的时序信息时,这种缺陷将导致无法正确分析网络中的传播和最短路径发现问题。本文利用航班时刻表中航班的起降时间作为边的时序标签,通过划分时间窗口^[18]的方法分析网络元素(节点和边)、拓扑统计量的变化,从而刻画网络结构的变化;同时,本文深入研究整个网络航班起飞时间的爆发性及节点传播所需的等待时间对时序网络上传播特性的影响,并根据节点呈现出来的爆发性对其进行有效分类。

1 数据来源

本文使用复杂网络理论对中国航空网络进行动态拓扑结构分析。网络中的边数据来自OK旅行网^[24],该网站提供国内所有民航公司预售票航班信息,同时提供了大量的机场分布和航班计划信息,本文提取了2014年183个机场的进离港航班时刻表,包括航班号、航空公司、起飞机场、起飞时间、目的机场、降落时间、班期以及机型,蕴含了全国1 627条航线和14 268个计划航线,由国内包括南方、厦门、国航等在内的28家航空公司运行经营。网络中节点数

据集源于互联网上已有的机场信息,包括机场代号、经纬度和所属省市。网络中边的方向反映了机场的进出港。数据预处理完成了航线端点和机场信息的映射,保留了航线上的时间属性。该数据集包含了2014年国内最新最全的航班计划表信息,是已知研究中包含航班数、机场数最多的数据集,同时本文也首次以航班时刻表为基础对航空网络进行实证分析,探索网络中的时序特征。

2 网络拓扑静态特征

首先移除网络时间信息,假设边连接在整个周期有效,将航空网络视为节点数取183,边数为1 627,边权为航线上航班数量的无向赋权静态网络,分析网络静态拓扑参数,并与已有研究进行比较。

中国航空网络拓扑结构特征变化比较如表1所示,表中 a 表示分布第一阶段指数, b 表示分布第二阶段指数。可以看出中国航空网络随着规模的扩大,其连接更加紧密,呈现出来的无标度特性和小世界现象更加明显。本文最新的数据中包含了部分新机场,这些机场只配置了少量与部分重要节点相连的航班,集聚系数较高;节点的增加也缩短了网络中航班的平均中转次数。

表1 中国航空网络拓扑结构特征变化比较

方法	节点/边	度分布	平均度	集聚系数	平均路径长度
文献[16]	121/1 378	$k=20-0.53^a-2.05^b$	11.38	0.75	2.26
文献[19]	161/1 185	$k=29-0.40^a-2.16^b$	14.72	0.70	2.14
本文	183/1 627	$k=30-0.56^a-2.41^b$	17.8	0.73	2.06

3 网络拓扑时序特征

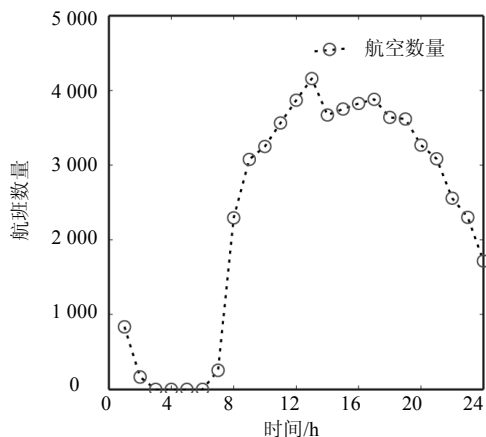
中国航空时序网络由183个机场节点以及14 268条带时间标签的边组成。一个时间周期内,航班的飞行情况因时而异,导致不同时刻同一机场起飞的航班数量有别,边的权重或数量有所增减,导致不同时刻的网络拥有不同的拓扑结构。至此,航空网络可视为一个基于航班时刻表的周期性变化的动态网络。动态航空网络中的边是带有一个或多个时间标签的航线,且这些时间区间可能重叠。过去对此类问题的处理常常忽视航线上的时间标签,将整个周期内动态网络中的所有边都视为静态边,以进行静态拓扑分析。本文将单个周期划分为长度为1小时的24个时间窗口,将出现在该时间窗口的航线视为有效边,包括之前起飞但还未降落的航班和在该时间段内起飞的航班,网络中节点为有效边上的节点集合。由此,动态网络被表示成静态网络序

列,对每个时间窗口中的静态网络进行分析,研究网络拓扑特性以及相关关系的时序变化,从而达到分析动态网络拓扑结构特征的目的。

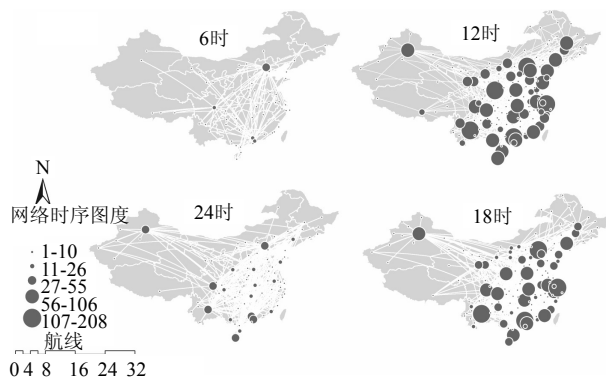
3.1 网络元素(点和边)时序特征

航空网络中的时间信息主要是指边的时间标签,即航班的起降时间。本节对24个时间窗口内空中的飞机数量进行统计,分析不同时间段内我国空域中在飞航班数量的变化情况。从图1a中可见,网络中航班数量从6:00开始,并在6:00—10:00间急剧上升,随后直到18:00一直处于顶峰时期,接着逐渐下降,并在次日3:00达到0且持续到6:00。航班运行主要集中在10:00—24:00,占比达88.37%,且在10:00—18:00之间达到流量高峰,占比为59.22%。航班数量体现出来的人类出行规律与其他交通工具略有不同^[20],高峰时期持续时间较长且不出现在人们上下班时刻。

节点度随时间的变化在一定程度上反应了机场运营量的时序特征,机场所在城市的人口密度和经济发展程度确定了机场运营时间和持续的时长。图1b表明在6:00—7:00之间,网络节点地理分布稀疏、连接较弱,出行人数较少,且出发地多是经济发达城市。8:00以后大量小型机场出现以满足人们逐渐增长的出行需求,致使节点在12:00—13:00之间分布密集且航线覆盖面广,并且每个机场的运量都较大。18:00时网络的运输压力较12:00时有所缓解,运营时间较短的机场在18:00之前停止运营,此类机场多处于偏远地区或机场空间密度较高的局部地区,如阿勒泰机场和庆阳机场等。偏远地区人口少,出行量低且时间比较集中;机场空间密度较高地区的小型机场是为缓解该地区主要机场在高峰时期的运行压力。由于部分航班降落时间较晚,目的地机场持续等待,所以末期(24:00)机场数量大于初期(6:00)开始运行时。



a. 航班数量时序特征



b. 网络节点和边的时序特征

图1 航空网络组成元素时序特征

3.2 网络属性时序特征

拓扑统计量是网络拓扑特性的定量描述,本节通过研究网络统计量、分布函数和相关性的时序特征,分析网络的演化过程。

3.2.1 拓扑统计量

反应整体网络特征的统计量包括网络平均度、平均介数和平均集聚系数以及相应的分布函数。图2a表明网络平均度随时间的变化较平均介数和平均集聚系数平缓,且达到稳定(指标值不随时间波动)的时间较长;平均介数和平均集聚系数同时在15:00达到稳定,认为此时网络已饱和,再增加航班数量对两者并无影响。节点的出现有时间间隔且部分节点出现时间较晚导致平均度不可能出现激增。同时,航班数量在6:00—11:00之间激增且在13:00出现高峰,已基本构成“无标度”结构,后续节点和边的数量变化对平均介数和平均集聚系数影响并不大。

网络的平均度、平均介数和平均集聚系数易受极值影响,未能充分反映网络特征的变化,为此,可通分析统计量分布函数随时间的变化来呈现。图2b表明拓扑统计量分布函数的参数值随时间变化逐渐趋于稳定,度分布的两阶段参数和介数分布参数都在短时间内攀升并逐渐趋于稳定。网络在6:00时,只有少部分的节点之间存在微弱的联系(图1b),度分布没有呈现出明显的两阶段现象,导致第二阶段参数出现突高。集聚系数的分布参数则经历了先降低后增长的过程。大量孤立节点(集聚系数为0)的存在使其分布函数斜率较大;下降过程中连接孤立节点的边不断出现,使得孤立节点不断减少,在11:00达到0;增长过程中机场之间的往来逐步增强,网络趋于紧密。

综上,网络拓扑结构呈现出明显的激增和稳定现象,主要受早期网络中航班数量激增的影响。

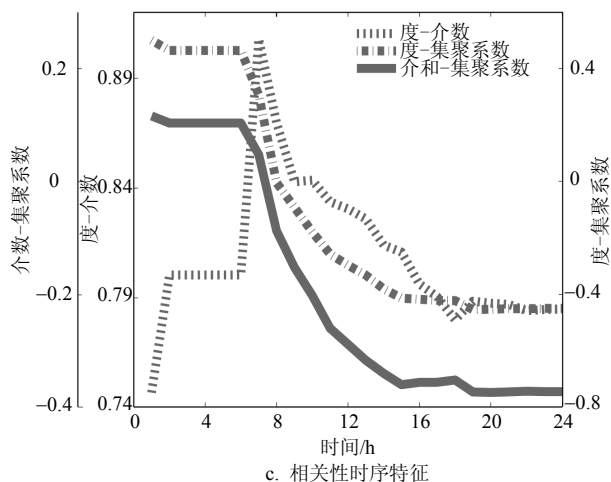
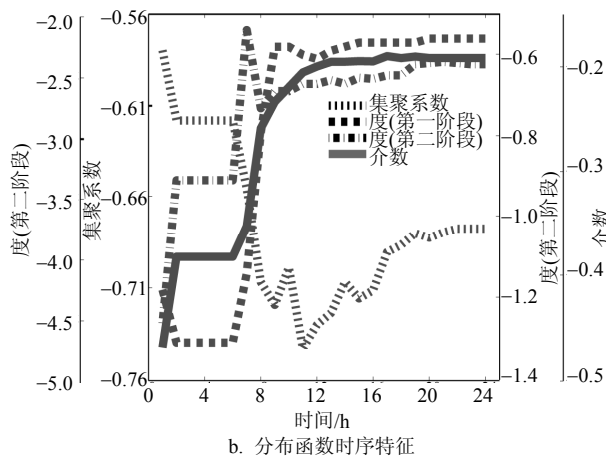
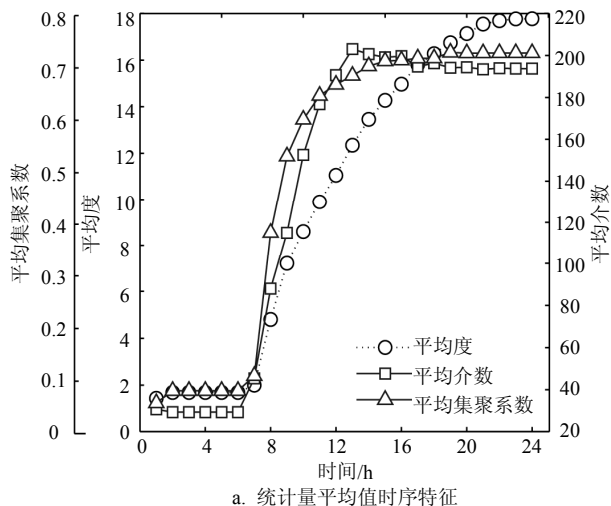


图2 航空网络拓扑结构时序特征

3.2.2 统计量相关性

节点统计量之间的相关性反映航空网络中节点偏好连接导致的结构上的相关程度。如图2c所示, 度-介数呈现出的正相关性先增后降, 最低值高达0.74, 6:00—7:00之间网络节点较少且多为重要机

场, 连接边较多且多为有效的最短连接, 故此时度-介数相关性高达0.9, 而后随着小型机场和大量边的出现, 相关性逐渐减小。

不同于静态的拓扑研究结果^[16-17], 度-集聚系数和介数-集聚系数存在正相关向负相关的过渡。6:00—7:00之间, 网络中节点和边的数量较少且多是重要节点之间的连接, 度大的节点的邻居节点之间相互连接的概率更大, 呈现出度-集聚系数正相关性; 随后节点间实际边的增长明显慢于节点增加导致的可能连接边数, 从而正相关性削弱, 负相关性出现并逐渐增强, 稳定在-0.4左右。

4 网络传播时序特征

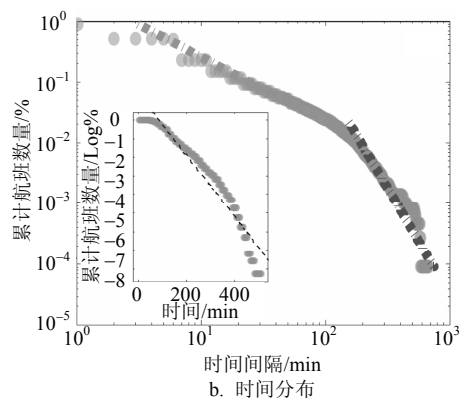
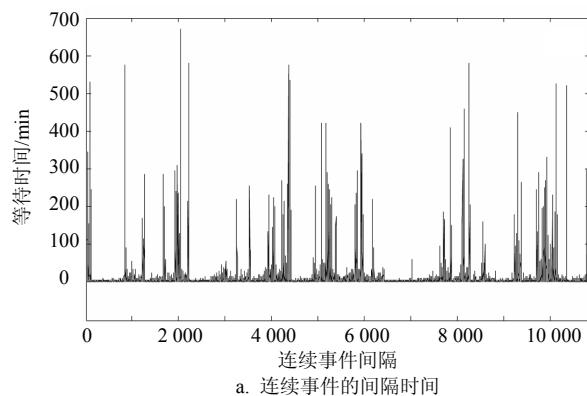
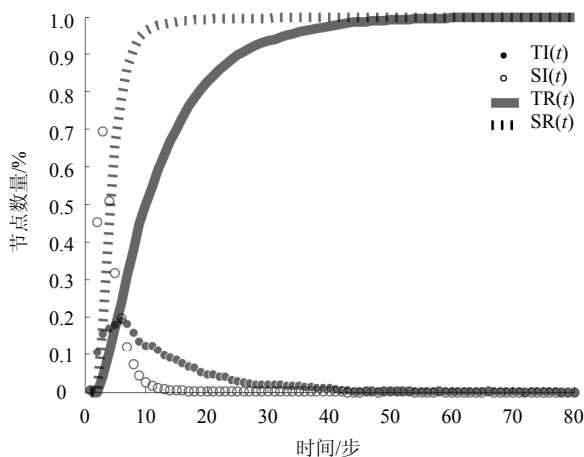


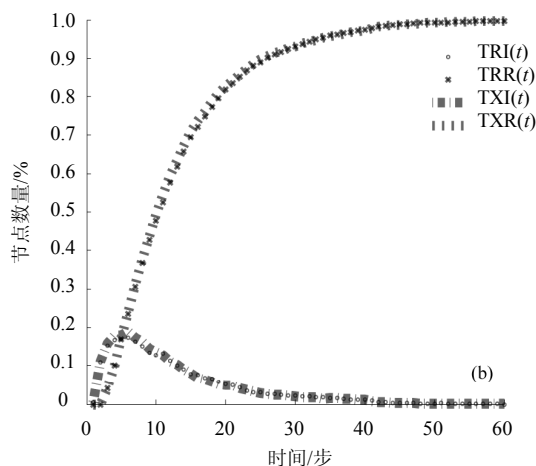
图3 航空系统中航班行为时间间隔

时序网络与静态网络的不同之处主要体现在节点上的等待时间和边上的完成时间。等待时间指同一节点连续两个事件发生的时间间隔, 完成时间指被感染节点将疾病传递给其邻接节点的时间差。静态网络普遍认为疾病在节点和边上的传播只受网络拓扑结构的影响, 不存在任何时间上的限制。大多数真实网络并非如此。图3a展示了中国航空网络中相同机场起飞事件连续发生的间隔时间, 图3b表明中国航空网络机场中航班的起飞时间间隔服从两阶

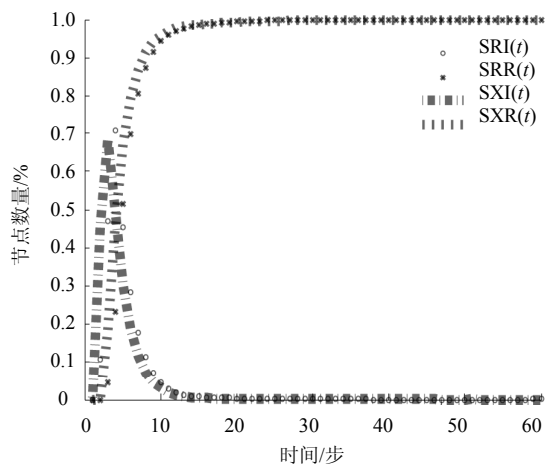
段幂律分布,指数分别为 -0.87 和 -2.97 。这意味着航班的起飞行为存在较明显的爆发现象,即短时间间隔后跟随长时间间隔的频率较大。图3b中的插图表明边上的完成时间服从参数为 -3 的指数分布。



a. 时序对 SIR 传播的影响



b. 感染源对时序网络的影响



c. 感染源对静态网络影响

图4 静态网络和时序网络SIR传播模型对比

网络属性差异将直接导致网络功能的不同。文章分别对静态网络和时序网络进行SIR传播模型的建模,该SIR模型采取接触即感染的传播形式,研究3种状态的人群随时间的变化情况,如图4所示。图中, $TI(t)$ 、 $TR(t)$ 和 $SI(t)$ 、 $SR(t)$ 分别表示时序网络和静态网络在 t 时刻节点感染态和恢复态的比例, $TRI(t)$ 、 $TRR(t)$ 、 $TXI(t)$ 、 $TXR(t)$ 和 $SRI(t)$ 、 $SRR(t)$ 、 $SXI(t)$ 、 $SXR(t)$ 分别表示随机感染和蓄意攻击情况下时序网络和静态网络在 t 时刻节点感染态和恢复态的比例。由图4a可见,由于时序网络中时间依赖性和时序相关性对传播的限制,导致传播稳定的时间较长,感染的范围缩小;同时对比图4b和图4c,发现时序网络对感染源的选择并不敏感,蓄意攻击(重要节点作为感染源)和随机攻击(随机节点作为感染源)对时序网络的影响差异不大,具有较好的鲁棒性;相反,度值大的节点被感染将使静态网络在更短的时间内达到稳定,使其感染范围更广。

5 基于时序特征的节点分类

航空网络的异质性表明,不同节点拥有不同的行为模式。以节点的爆发性衡量节点的行为模式,以“等待时间”的平均值和标准差作为其定量衡量的依据^[21],在平均值和标准差维度下运用基于密度的聚类算法(density-based spatial clustering of applications with noise, DBSCAN)对其进行社区挖掘,如图5a所示,其插图分别表示成都双流(类别1)、大庆萨尔图(类别2)和库车机场(类别3)内航班起飞的时间间隔分布。聚类结果在以GDP差值为距离定义的地图中具有良好的区分度,如图5b所示,表明机场城市GDP是导致机场时序特征差异的主要原因。图5b以GDP值最大的上海为圆心,将其他城市GDP与上海的差值作为半径,并标出机场所属的不同类别。其中万州和黔江以及江北机场同在重庆市,但重要程度各异,时间特性有别,所属类别不同。图5c展示了节点分类在实际空间地图中的分布情况,类别3中机场主要分布在沿海以及省会等经济文化中心,运行时间密集,具有明显的爆发性;类别2中机场分布广泛,主要是东北和中部地区,运行时间较为稀疏,其爆发性较弱;类别1中机场多集中在西部、南部,也出现在类别2和类别3机场附近(机场空间密度较高地区),只在几个固定时间节点运作,呈现一定的周期性趋势。

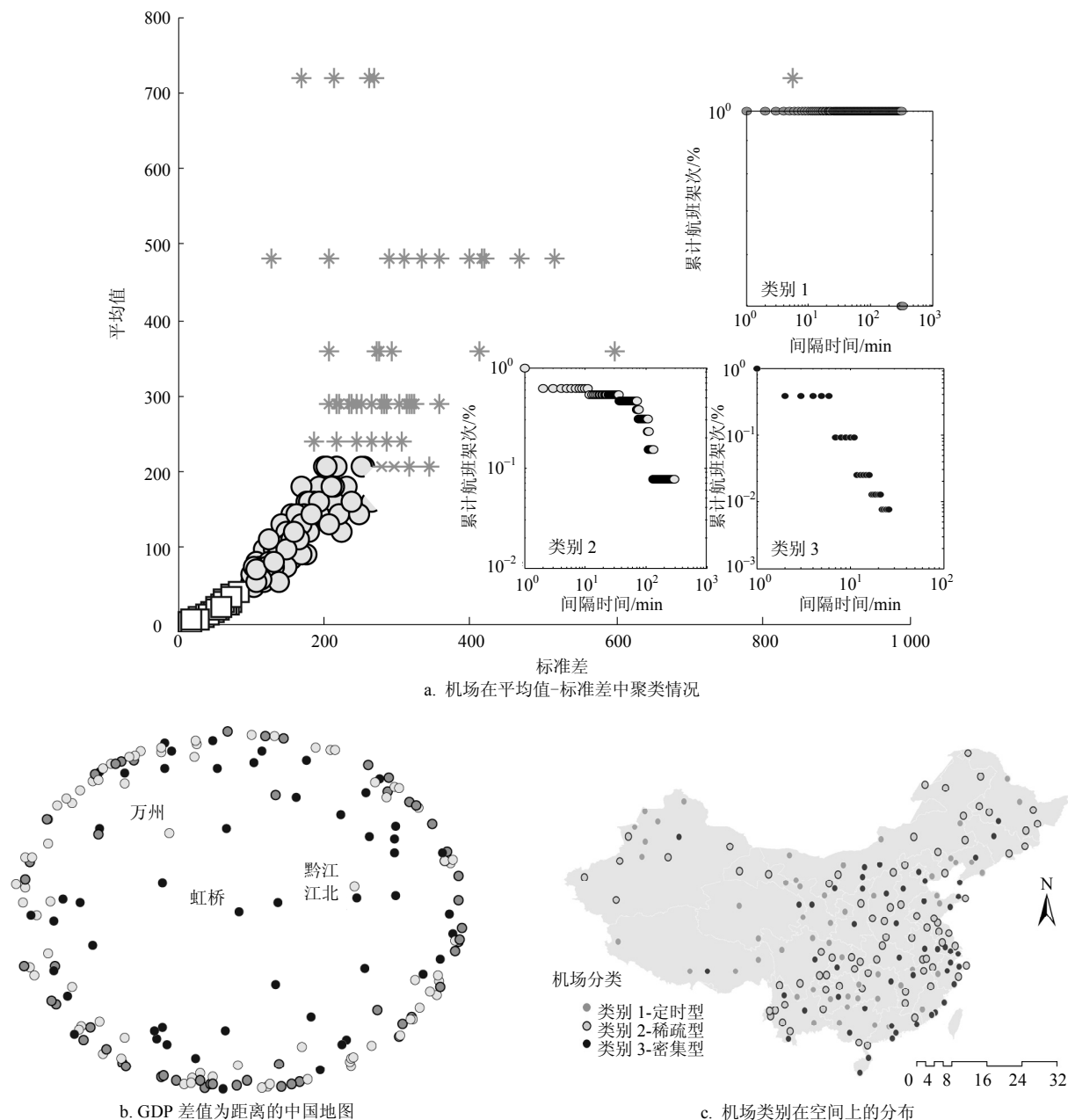


图5 基于机场爆发行为的分类

6 结束语

本文基于航班时刻表,首次将时间信息引入中国航空网络拓扑结构中,分别应用传统网络分析方法和时序网络分析方法对网络的静态和动态拓扑属性进行研究。在网络静态拓扑方面,与已有研究相比,本文数据集中新增机场和航线使网络呈现出更加显著的无标度特性和小世界现象。通过使用时间窗口方法将航空网络分割为24个子网络序列,研究其组成元素和拓扑统计量的时序特征,包括航班数量、统计量的分布函数和两两之间的相关性。结果表明,航空网络从6:00开始运作且和其他交通工具

一样在13:00和17:00出现高峰;网络度分布和介数分布激增后缓慢趋于稳定,而集聚系数分布则会先下降再回升而后趋于稳定;度-介数为逐渐增强的正相关关系,度-集聚系数和介数-集聚系数在8:00时出现正相关和负相关的转折点。此外,本文还研究了网络中连续事件发生的时间间隔呈现出来的爆发性对传播的影响,以及基于爆发性行为的节点分类方法。结果发现,网络爆发性对传播起着延缓作用(稳定时间长,传播范围小);网络中存在时间间隔分布差异明显的三类节点,且与其空间位置有关。

本文研究是对已有静态航空网络分析的一个重要补充。充分考虑航班时间信息对网络结构和功能

的影响,对进一步科学合理地开展航班延误级联失效^[22]、人群移动与疾病传播等网络动力学问题的研究^[23]具有重要的科学意义。

参 考 文 献

- [1] ZHANG Jun, CAO Xian-bin, DU Wen-bo, et al. Evolution of Chinese airport network[J]. *Physica A: Statistical Mechanics and its Applications*, 2010, 389(18): 3922-3931.
- [2] ALBERT R, BARABASI A L. Statistical mechanics of complex networks[J]. *Reviews of Modern Physics*, 2002, 74: 47-97.
- [3] NEWMAN M. The structure and function of complex networks[J]. *SIAM Review*, 2003, 45: 167-256.
- [4] GUIMERA R, AMARAL L. Modeling the world-wide airport network[J]. *European Physical Journal B - Condensed Matter*, 2004, 38(2): 381-385.
- [5] BAGLER G. Analysis of the airport network of India as a complex weighted network[J]. *Physica A Statistical Mechanics & Its Applications*, 2008, 387(12): 2972-2980.
- [6] BROCKMANN D, HELBIG D. The hidden geometry of complex, network-driven contagion phenomena[J]. *Science*, 2013, 342(6164): 1337-1342.
- [7] GOSWAMI B, SHEKATKAR S M, RHEINWALT A, et al. A random interacting network model for complex networks[J]. *Scientific Reports*, 2015, 5: 18183.
- [8] BARRAT A, BARTHELEMY M, VESPIGNANI A. Modeling the evolution of weighted networks[J]. *American Physical Society*, 2004, 70(2): 066149.
- [9] LI W, CAI X. Statistical analysis of airport network of China[J]. *Physical Review E*, 2004, 69(4): 266-289.
- [10] WANG Jiao-e, MO Hui-hui, WANG Fa-hui. Evolution of air transport network of China 1930-2012[J]. *Journal of Transport Geography*, 2014, 40: 145-158.
- [11] WANG Jiao-e, MO Hui-hui. Complex evolution process of China's air transport network[J]. *Journal of Transportation Systems Engineering*, 2014, 14(1): 71-80.
- [12] LIU Xing-jian, DERUDDER B, GARCIA C G. Exploring the co-evolution of the geographies of air transport aviation and corporate networks[J]. *Journal of Transport Geography*, 2013, 30: 26-36.
- [13] XIAO Yu, WANG Fa-hui, LIU Yu, et al. Reconstructing gravitational attractions of major cities in China from air passenger flow data, 2001-2008: a particle swarm optimization approach[J]. *The Professional Geographer*, 2013, 65(2): 265-282.
- [14] PAN R K, SARANKI J. Path lengths, correlations, and centrality in temporal networks[J]. *Physical Review E*, 2011, 84(1Pt2): 1577-1589.
- [15] 刘宏鲲. 中国航空网络的结构及其影响因素分析[D]. 成都: 西南交通大学, 2007.
- LIU Hong-kun. Analyzing the structure of Chinese aviation network and impact factors[D]. Chengdu: Southwest Jiaotong University, 2007.
- [16] 刘宏鲲, 周涛. 中国城市航空网络的实证研究与分析[J]. *物理学报*, 2007, 56: 106-112.
- LIU Hong-kun, ZHOU Tao. Empirical study of Chinese city airline network[J]. *Physical Review*, 2007, 56: 106-112.
- [17] 朱冰心. 中国航空网络结构及其演化分析[D]. 上海: 复旦大学, 2009.
- ZHU Bing-xin. Analysis of Chinese aviation network structure and its evolution[D]. Shanghai: Fudan University, 2009.
- [18] BASU P, BARNOY A, RAMANATHAN R, et al. Modeling and analysis of time-varying graphs[EB/OL]. 2015-10-20. <https://arxiv.org/abs/1012.0260>.
- [19] 曾小舟, 唐笑笑, 江可申. 基于复杂网络理论的中国航空网络结构实证研究[J]. *交通运输系统工程与信息*, 2011, 11(6): 175-181.
- ZENG Xiao-zhou, TANG Xiao-xiao, JIANG Ke-sheng. Empirical study of Chinese airline network structure based on complex network theory[J]. *Journal of Transportation Systems Engineering and Information Technology*, 2011, 11(6): 175-181.
- [20] SUN Li-jun, AXHAUSEN K W, LEE D H, et al. Understanding metropolitan patterns of daily encounters[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(34): 13774-13779.
- [21] 周涛, 韩筱璞, 闫小勇, 等. 人类行为时空特性的统计力学[J]. *电子科技大学学报*, 2013, 42(4): 481-540.
- ZHOU Tao, HAN Xiao-pu, YAN Xiao-yong, et al. Statistical mechanics on temporal and spatial activities of human[J]. *Journal of University of Electronic Science and Technology of China*, 2013, 42(4): 481-540.
- [22] 邵荃, 朱燕, 贾萌, 等. 基于复杂网络理论的航班延误波及分析[J]. *航空计算技术*, 2015(4): 24-28.
- SHAO Quan, ZHU Yan, JIA Meng, et al. Analysis of flight delay propagation based on complex network theory[J]. *Aeronautical Computing Technique*, 2015(4): 24-28.
- [23] BENGTSSON L, GAUDART J, LU X, et al. Using mobile phone data to predict the spatial spread of cholera[J]. *Scientific Reports*, 2015, 5: 8923.
- [24] OK旅行网. 2014年中国航班时刻表[EB/OL]. [2016-05-10]. http://jipiao.oklx.com/cn_airfield_schedule.aspx.
- OK Travel Network. China flight schedule in 2014[EB/OL]. [2016-05-10]. http://jipiao.oklx.com/cn_airfield_schedule.aspx.

编辑 叶 芳