

参赛队员姓名: 严子渔 彭彦迪 范思齐

中学: 华南师范大学附属中学

省份: 广东省

国家/地区: 中华人民共和国

指导教师姓名: 黄秉刚

论文题目: 中国民用航线网络的测度以及相关实证研究

本参赛团队声明所提交的论文是在指导老师指导下进行的研究工作和取得的研究成果。尽本团队所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果。若有不实之处，本人愿意承担一切相关责任。

参赛队员： 严子渔 彭彦迪 范思齐 指导老师： 黄秉刚

2019 年 9 月 15 日

# 中国民用航线网路的测度以及相关实证研究

严子渔 彭彦迪 范思齐

摘要：

民用航线网络是交通网络至关重要的组成部分。飞机速度快，行程远，受地理束缚较小的特点使得航空运输具有不可替代性，其网络结构与传统的交通网络有着较大的不同。近十年来，随着生活水平的提高，我国对于民用航空的需求持续增加，而政府管制逐渐放松，使得中国航线网络的结构和体量都发生了较大的变化。本研究从宏观和微观角度，结合近五年的最新数据，对中国国内航空网络的整体结构，与节点个体在网络中相对重要性以及两者变化趋势进行了测度，并通过实证分析，结合网络外部性与经济的规模效应理论对其相关因素进行分析。通过测度和实证分析，本文得出了以下结论：中国国内航线网络是一个具有无标度性的“小世界”网络，但受到网络外部性的影响，其无标度性在逐渐增强，有向无标度网络演化的趋势；受到规模经济效益以及需求存量的影响，新开航线大多都在非枢纽机场，因此削弱了枢纽机场对于网络总体通达性的贡献与相对重要性，但枢纽机场的航线数量与旅客数量依然在增长；当前我国航空运输需求的主要增长来源，是由于消费观念转变带来的旅游需求的增量，与经济和收入水平的增长关联性相对较弱。

关键词：

民用航线 复杂网络 需求量 网络外部性 规模经济性 实证分析

# 中国民用航线网络的测度以及相关实证研究

一、引言	1
二、文献综述	3
三、研究方法及数据	3
四、对中国航线网络总体性质的测度	5
五、对中国航线网络具体节点的测度	8
六、中国航线网络的影响因素的实证分析	12
七、总结与展望	16
附录1：获取数据时采用的爬虫程序说明	21
附录2：致谢	27
参考文献	28

## 一、引言

自改革开放以来，国民经济大幅提升，人民生活水平稳固上升。经过四十年的高速发展，人们对美好生活需要日益上涨，不仅仅是对物质的要求，也是对政治经济文化等诸多领域的更高期望。随着人们收入水准的提高，对生活品质和更高需求层次的要求也越来越高。人们温饱问题得以解决，可支配收入分配在出行上的份额逐渐升高。因此，全国出行的需求与日俱增。

而自飞机问世以来，其在交通运输中的重要性一直在稳步上升：人们利用飞机进行远距离飞行或者运输设备物资，而与其他交通方式相比，飞机具有速度快，行程远，不受特殊地理原因束缚等优点。这些特质使得航空运输具有不可替代性。

近十年来，随着对于民用航空的需求持续增加，我国对民用航空的管制逐渐放松，国有企业对航线的垄断逐渐被打破，市场环境变得更为开放。开放的市场使得各大私人航空公司进入市场壁垒减小，直接导致了航线以及机场数量的激增。相较2005年全国境内仅有141座民航机场与1378条国内航线，截至2019年7月，我国境内共有227个有定期航班的民用机场。

随着需求量增加，民航网络的结构性质势必发生了持续而显著的变化。全国航线数量持续增加，对原有的航空枢纽，以及民航机型的研发势必会带来影响。在近年来对航空运输的总需求增加的背景下，对能够承载更多旅客，并提供更舒适的飞行体验的超大型客机的研发与引入的关注度近年来逐渐升高，为此。然而从超大型客机在航空市场上的表现来看，随着A380与波音747客运型的停产，超大型客机的表现远远不如市场的预期。由此可见，伴随着对航空运输需求量增加的，还有航空网络日渐复杂的内部演化规律带来的对客机的新要求，而非仅仅是将网络的每一部分都简单地放大。。因此准确地把握航线网络演化的趋势，对于航空业所涉及的各个部分，无论是飞机研发，还是航空公司、租赁公司与相关的商业、旅游产业链都有相当重要的意义。

随着中国经济的快速发展，中国的航线网络也在以极高的速度演化，因此结合最新的数据，对中国航线网络从宏观和微观角度的演化趋势，及其可能的内在原因进行分析是十分有必要的。通过最新的数据，本文对中国航线网络的整体，以及网络中的节点的一些性质进行了测度，并与相关的影响因素进行了实证分析。其中，航线作为网络所具有的外部性，以及随着需求量增长而带来的规模经济效益是最有可能的潜在原因：非枢纽机场的乘客数量增

加，得以开辟更多的直飞航线，使得传统的枢纽机场对于航线网络通达性的贡献以及对周边地区的虹吸效应不可避免的减弱。

通过从宏观到微观的测度，本文主要得到了以下的初步结论：

1. 中国的民用航线网络是一个小世界网络，但具有无标度性且无标度性在逐渐增强，而小世界特性逐渐减弱，有成为无标度网络的趋势。这与美国航空网络在过去演化的趋势大致相同。
2. 由于航线网络的规模经济性，随着非枢纽节点的需求量增长，其连接航线的数量会以较原有枢纽更快的速度增长，从而降低了原有枢纽对于网络整体的通达性的贡献以及相对重要性。
3. 尽管对于网络的相对重要性减弱，但传统枢纽的流量与航线量仍然在持续增长，其地位仍然难以撼动。
4. 当前我国航空运输需求的主要增长来源，是由于消费观念转变而带来的旅游需求的增量，与经济和收入水平的增长关联性不强。

## 二、文献综述

民用航线及其效用属于复杂网络理论在经济学中的应用，是产业经济学的一部分，兴起于上世界90年代。以色列经济学家Oz Shy从博弈论的角度得出在一定条件下，航线具有网络经济性，并得到了一些与放松管制、代码共享等相关的结论。2000年前后，Harris以及等人Guimera等人对北美和全球航线网络进行了一系列的测度。两组学者分别得出了北美航空网络是一个无标度网络，而世界航空网络则是一个具有小世界网络基本特征的网络系统的结论。

对比之下，中国的航空网络则展现出了一种相异的结构。2007年，刘宏鲲等人对我国航空网络进行了一次系统的测度，表明我国的民用航线网络是一个小世界网络。该研究也将节点城市本身的大小（人口数量）纳入了考虑因素，非常有效的把民用航线网络的实际情况与经济学理论的相联系，使得读者更容易理解，也更具有说服性，为我们的研究提供了一定的基础。但美中不足的是，该研究仅仅考虑了人口因素，仍不够全面。因此，在此基础上，我们将还会着眼于城市本身的更多因素，不仅有人口规模带来的更大需求，还有如经济发展带来的购买力提升收入增高等情况，在更广的时间跨度下对网络结构变化的原因进行实证分析。

而王姣娥等人在2009年发布的研究则更加深入的研究了我国的航空网络的结构。除了与刘宏鲲等人已经得出的结论以外，还得出了我国结构层次不明显的结论并预测将来我国航空结构将日趋复杂。该研究相较之前的相关研究而言更加详细，也更加聚焦于我国的航空网络，但创新成分不足，并无研究方法和研究思路上的突破，无法给予后人更多研究新思路与新方向。

闫玲玲等人在2016年发表的论文则是从五个不同的指数衡量我国航空网络的脆弱性，许多参量对于我们而言也具有极高的参考价值。如介数中心性这一衡量标准对于该研究而言可以衡量节点影响力，且这意味着容易发生拥堵的程度。其对于测度指标的选取对本文从微观角度的测度具有借鉴价值。但是因为其研究目的在于脆弱性，而没有对于网络总体性质的测度、结构随时间的变化以及内在原因作出进一步的研究。而在我们的研究中，我们将从时间的角度剖析我国航空网络情况及变化趋势，并结合不同因素分析其趋势产生原因。

苏延芳与肖兴志二人的反事实研究也为我们的研究提供了良好的理论基础。他们通过对于美国放松航空管制之后的演变入手，对Hub-and-Spoke（HS）网络和Fully Connected

(FC) 网络两种网络体系的运营成本和运量等因素综合考虑，基于已有数据的基础上，对放松管制后的我国网络进行了预测。他们的反事实研究指出随着管制放松与需求量的增大，HS网络会逐渐取代FC网络，而这一特点在下文中也将会被运用。但是苏延芳与肖兴志二人的研究数据过时，最新的数据也仅仅是在2004年。航空科技发展日新月异，而航空网络结构也早已与15年前截然不同，这也使得我们基于最新数据的测度与实证分析变得十分必要。

航空网络作为一个网络系统，具有外部性。陈蓉和郭晓武二人对网络外部性的解析介绍了网络外部性的含义，来源与构成，分类等，很好的为读者提供了一个网络外部性的基本认识，但是也正因如此，针对性不强，无法对某一特定网络所具有的外部性进行具体的分析与阐述。而本文将结合航空网络的规模效应以及外部性，对航空网络结构改变的因素进行分析。李煜伟以及倪鹏飞二人也基于网络本身具有外部性这一特征，利用了网络的分析工具和新经济地理理论，分析了城市间的运输网络与城市本身的经济增长，改进了新古典区域经济增长模型，得出了应该通过增加运输网络间的集聚系数并降低平均路径长度，加快其要素的流动以加大经济增长。李倪二人的研究对城市的经济发展与城市之间的交通运输进行了深入分析，对我们的研究也大有启发。但是该研究更倾向于交通网络生产要素的运输能力，而民用航空网络则注重于载人能力，这两者在分析上有些许区别，前者注重于运输成本与效率等因素，而后者则更注重于经济与人口增长等带来的需求激增。同时，因为计算量等诸多原因，该研究的样本较少，文中仅仅选用了18个城市进行分析，一般性相对不足，而本文将对全国两百余机场进行测度，保证了结论的一般性。

总的而言，尽管对于世界以及我国航线网络的结构，有不少论文进行了测度，但是中文文献的数据都相对较旧，且仅仅聚焦于单一时间点的航线网络结构，而没有涉及航线网络发展演化的内在规律，或仅仅从网络外部性理论进行推导，而没有对实际的变化趋势进行测度与实证分析。而且大部分论文仅对本文涉及的宏观测度部分，而没有从微观角度从具体节点的中间中心度等性质的变化及其原因进行分析。因此，本文力求从实证分析预测度的方法出发，结合宏观和微观两个角度探究航线网络变化的趋势及其内在原因，来填补此前中文文献对此方面内容的空缺。

### 三、研究方法及数据

任何网络都可以看作是按照某种方法将节点连接而成的系统。18世纪，欧拉提出了七桥问题，将横跨城市内河流的七座抽象为七条边，并证明了无法从任意节点出发经过所有边再回到原点，开创了图论的先河。

1960年代，东欧数学家Erdos和Renyi创造了ER随机图模型，开创了对抽象网络的系统研究。在ER随机图模型中，任意一对节点对相连的概率都为 $p$ ，因此一个含有 $N$ 个节点的随机图中边的总数可以用一个服从Poisson分布，期望值为 $p[N(N - 1)/2]$ 的随机变量表示。因此对于一个有 $N$ 个节点的ER随机图，其总共有 $M$ 条边的概率为

$$p^M(1 - p)^{\frac{N(N - 1)}{2 - M}}$$

将航空网络抽象化，用点集 $k = [k_1, k_2 \dots k_n]$ 代表所有机场，用边集 $E = [e_1, e_2 \dots e_m]$ 代表所有航路，则网络总体可以记为无向图 $G = [k \quad E]$ ，总节点数表示为 $N = |k|$ ，而总连接数可以记为 $M = |E|$ 。网络的邻接矩阵可以记为 $A_{n \times n} = (a_{ij})$ ，当节点 $v_i$ 和 $v_j$ 之间有航线连接的时候 $a_{ij} = 1$ ，反之则 $a_{ij} = 0$ 。

节点 $v_i$ 的度(degree) $k_i$ 指的是与节点直接相邻的航线的数目。一般而言 $k_i$ 越大，节点在网络中越重要。度的表达形式为

$$k_i = \sum_{j=1}^N a_{ij}$$

航线网络的密度(density)为网络中实际存在的航线数与理论可容纳的边数上限的比值。一个具有 $N$ 个节点与 $L$ 条航线的航空网络密度为

$$d(G) = \frac{2L}{N(N - 1)}$$

网络中节点的距离 $d_{ij}$ 节点之间最短路径所需连接的次数，在航空网络中可以理解为连接两个城市所需最短的飞行次数，与城市对的物理距离无关。

网络的平均路径长度为网络内部任意两个节点间所需连接次数的平均值，记为

$$L = \frac{2}{N(N - 1)} \sum_{i \geq j} d_{ij}$$

网络中节点的度中心性指的是按照网络的度的大小对网络节点进行排列所得到的排名。因为不同的网络中所存在的节点的数量不同，为了便于比较，一般节点的中心性在比较前会进行归一化

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

对现实中复杂网络的实践和测度在Erdos提出随机图理论之前就已经出现。美国社会学家Milgram提出了六度分隔理论，即只需要6段关系就能连接任何两个不认识的人。这是复杂网络的小世界现象第一次被提出。1998年，J.Watts等人第一次提出了以高集聚系数（或高簇系数）与低平均路径长度为特点的小世界网络（即WS图）。

J. Watts提出了用于衡量网络节点聚类情况的集聚系数。一个节点*i*的局部集聚系数指其所有相邻节点之间连接的数量与最大的可能连接数量之比，即：

$$C_i = \frac{1}{k_i(k_i - 1)} \sum_{j \neq k} a_{ij} a_{ik} a_{jk}$$

其中*k<sub>i</sub>*表示节点*i*的度，*a<sub>ij</sub>*表示邻接矩阵元，当节点*i, j*之间有直接连接时值为1，否则为0。任意节点的局部集聚系数介于0与1之间。当一网络的集聚系数大于同样的顶点集合上构造的，具有相近平均路径长度的随机网络的集聚系数时，此网络可以判断为小世界网络。

1999年，Barabási等人的研究揭示了网络的无标度特性，即存在一种复杂网络中大部分节点只和相邻的几个很少的节点连接，而一小部分节点与大量节点连接。一般而言，网络中节点的度的分布遵守幂律分布（Zipf齐夫分布），节点分布于幂律分布的拟合程度可以作为判断网络是否为无标度网络重要方法。任意抽取无标度网络内的节点，其度数*d*是自然数*k*的可能性为

$$P(d = k) \propto \frac{1}{k^\gamma}$$

即*d = k*的概率与*k*的某负次方相关，概率随*k*值降低。任何节点的度的分布与幂律分布拟合的网络都可以认为具有一定的无标度性。若节点的度的分布与幂律分布拟合程度越高，则可以认为网络的无标度性越强。一般而言，大型无标度网络的*γ*值介于2-3之间，这也可以作为判断网络无标度性的辅助条件。

本研究的测度基于通过爬虫程序（代码见附录）从机票代理以及航班追踪网站variflight以及中航信收集的2015-2019年中国境内定期航线班次表，包含一周执行的班次，机型以及起降地。本研究对于数据有以下设定：

- (1) 假设航线网络是一个对称网络，即每周从A地飞往B地的航班数以及提供的座位数，与每周从B地飞回A地的航班数以及提供的座位数相同。
- (2) 爬虫程序收集到的所有信息只包含公开售票的定期航班，不包括包机航班、空载调机航班、通用航空、私人飞机等未公开售票的航班。
- (3) 统计航班数据时以周为单位，即所收集的航班每周至少往返飞行一次。
- (4) 统计数据中只包含中国境内（不含港、澳、台）地级市及以上城市。由于本研究以城市为研究节点，位于同一城市且其间没有航班的机场的数据会被合并，如上海浦东机场与虹桥机场，北京首都机场与南苑机场；重庆江北机场与万州机场由于情况特殊，且两者之间有定期航班，故不合并进行统计。
- (5) 本研究收集的统计数据为售票时的飞行计划，与实际执行可能有一定出入。在计算两地间每周提供的座位数量时，本研究按照飞机设计的座位布置计算，不考虑超售、预留座位等情况。
- (6) 所有数据以邻接矩阵 $A_{n \times n}$ 记录， $n$ 为网络内节点数量，矩阵中 $a_{ij}$ 代表节点对 $(i, j)$ 之间的航班量，若 $(i, j)$ 间无航班，则 $a_{ij} = 0$ 。

通过以上方法，本文共收集2015-2019年每年1月、6月第一周全国各节点间每周航班量。截至2019年7月，全国共有民航机场238座，其中获得可测度数据的有227座；总共有5231个城市对之间有航班飞行；网络中所有节点的平均度为23.1，其中度最大的为北京（143）。

#### 四、对中国航线网络总体性质的测度

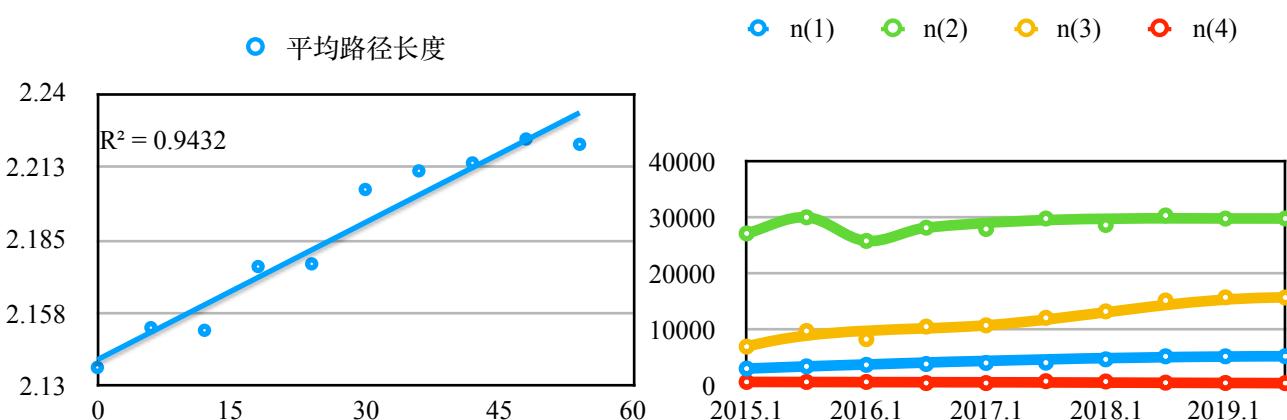
对航线网络宏观总体特性的测度，有助于更好地理解网络的基本性质及验证其分布的变化。此处最重要的两种特性为网络的小世界特性与无标度特性。小世界网络大部分节点相互不连接，但平均路径长度较短；无标度网络指大部分节点的度较低，而少部分节点的度较高。部分观点认为无标度网络属于小世界网络的一种，但无标度与小世界网络在度的分布等方面有不同的特性。无标度特性意味着一部分节点较高的度以及较低的集聚系数；小世界特性则分布形态更为多样，集聚系数较高。

一般认为全球的航空网络是一个小世界网络，具有一定的无标度性质；而一部分国家，如美国的航空网络则是无标度网络。因此航空网络的无标度性是否有随网络发展而增加的内在趋势，是一个十分重要的问题。

此处使用ucinet软件对2015-19年中国的航空网络进行测度，从平均路径长度、局部集聚系数以及节点的度的分布三个方面判断航空网络的无标度性质与小世界性质及其变化。

设 $n(p)$ 为网络内最短路径长度为 $p$ 的节点对个数， $\hat{n}$ 为航线网络的平均路径长度，则自2015-2019年，中国航空网络的节点对路径长度分布情况如下：

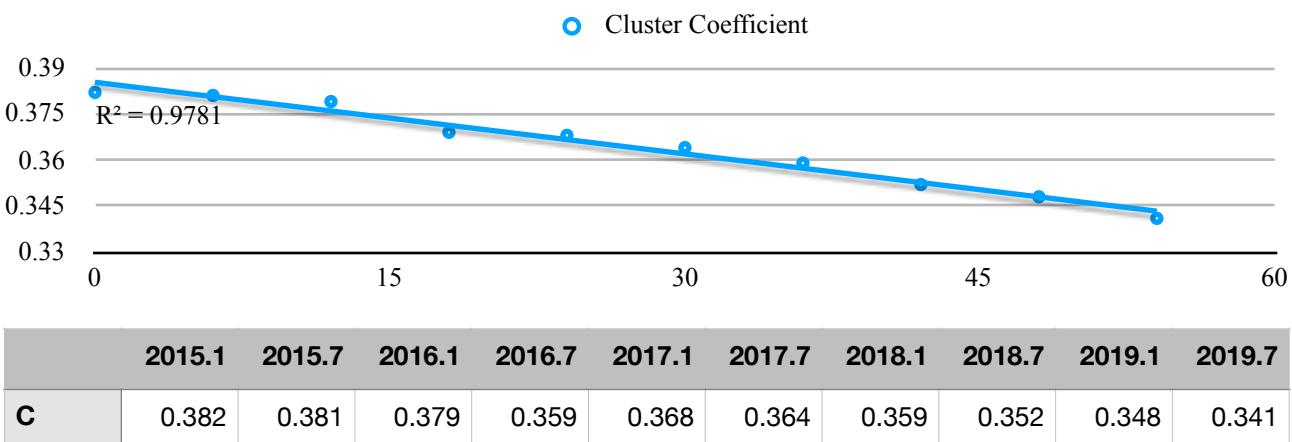
$n(p)$	2015.1	2015.7	2016.1	2016.7	2017.1	2017.7	2018.1	2018.7	2019.1	2019.7
$n(1)$	3018	3412	3683	3841	4031	4083	4673	5210	5211	5231
$n(2)$	27235	30132	25901	28272	28045	29915	28761	30458	29902	29918
$n(3)$	6952	9761	8287	10538	10765	12116	13269	15226	15783	15755
$n(4)$	624	620	624	420	420	754	703	490	438	408
$\hat{n}$	2.137	2.152	2.151	2.175	2.176	2.204	2.211	2.214	2.223	2.221



由此可见，相比美国网络等被认为是无标度网络的复杂网络，中国航空网络拥有较低的平均路径长度，这是十分明显的小世界网络性质。但可见的是，随着时间的推移，平均路径长度正在缓慢增加。这可能是因为中国航空网络逐渐向边远地区的节点扩张，显著地增加了整个网络的平均路径长度，因此其小世界性质随着扩张而逐渐减弱。

自2015年至2019年，在四年间，测度中有效的节点（即有定期航班的机场）数量自190个增加至227个，其中2015年及以后建成的机场在2019年的平均度为6.2，远低于2019年中国航线网络节点平均度22.3。因此，网络中节点的度的分布的“长尾效应”也更为明显，间接地说明了网络无标度性的逐渐增强。

另一判断网络小世界性质的指标为网络的总体集聚系数，同样地，经ucinet对数据的测度，我们可以得到中国航线网络近五年的集聚系数C。一般而言，网络的集聚系数越高，则网络整体越趋近于一个小世界网络。



此时我们可以发现，自2015年到2019年，中国航空网络的聚类系数持续下降，表明其基于空间的聚类性正在减弱，小世界性质在逐渐降低。这可能来源于高铁等替代交通工具对于特定城市对所产生的冲击，从而带来了部分节点局部集聚系数的减弱。关于高铁的开通对于航线网络的冲击，本文将会在后文中进行讨论。

网络的无标度性，可以定义为其节点的度的分布遵守幂律分布（齐夫分布），即随机抽取无标度网络内的节点，其度数d是自然数k的概率为

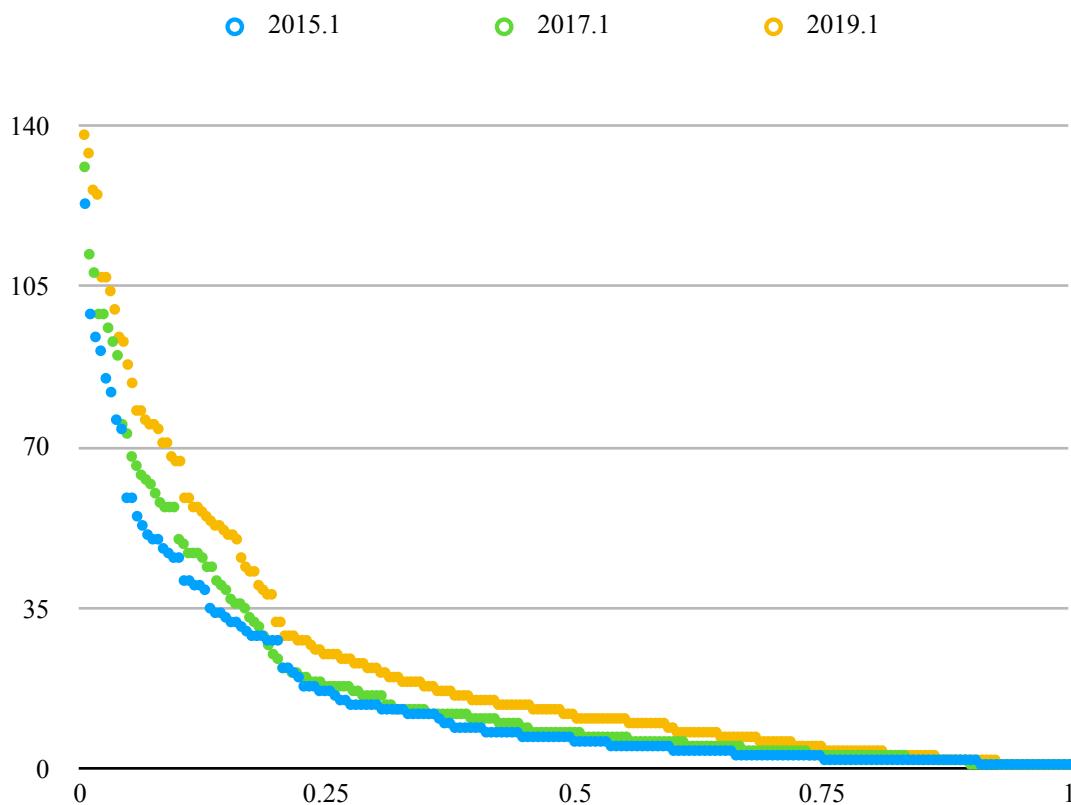
$$P(d = k) \propto \frac{1}{k^\gamma}$$

当网络节点的度的分布与幂律分布的拟合度越高时，网络的无标度性越强。当网络遵循幂律分布时，将会有极少量度数极高的节点，与大量度数较低的节点，节点度的平均数大于节点

度的中位数。下图中横坐标为特定度数的累积概率，而纵坐标为对应的节点度数。则此处对于度的分布的回归方程应为

$$y = Cx^{-\gamma}$$

其中C为常数， $\gamma$ 则是函数的幂次，一般现实中的无标度网络 $\gamma$ 值介于2-3之间。



上图为2015、2017与2019年1月的国内航线节点的度的分布，因为表格空间没有显示所有的数据。通过拟合得到：

2015年1月，网络度分布的回归方程为 $y = 3.5239x^{-1.167}$ ,  $\gamma = 1.167$ ,  $R^2 = 0.7929$ ;

2017年1月，网络度分布的回归方程为 $y = 2.4996x^{-1.204}$ ,  $\gamma = 1.204$ ,  $R^2 = 0.8456$ ;

2019年1月，网络度分布的回归方程为 $y = 2.1070x^{-1.219}$ ,  $\gamma = 1.219$ ,  $R^2 = 0.8687$ 。

由此可以发现，我国航空网络节点的度的分布具有一定的无标度性，但是并不是严格意义上的无标度网络。与此同时，我国网络的无标度性正在随时间而加强，而小世界性正在随时间而减弱。无论是测度得到的 $\gamma$ 值还是与幂律分布拟合水平的 $R^2$ 都在显著地向着无标度网络的水平靠近。因此尽管当前我国网络仍是小世界网络，随着我国对航空需求的进一步发展，以及对航空业准入管制的放松，我国的航线网络有类似美国航空网络自1979年《开放天空法案》后一样，逐渐演化成为一个无标度网络的趋势。

这一测度结果可以通过苏延芳等人对于航空网络测度的反事实研究解释，即随着管制放松与需求量的增大，Hub-and-Spoke（HS）网络会逐渐取代Fully Connected（FC）网络。即航空公司更倾向于在依靠特定枢纽运行，优先在度数较高的节点建立连接，并与其他航空公司建立代码共享，而不是独自建立整个国内航线网络；全国的航空网络整体将会逐渐从小世界网络过渡至无标度网络。

这一航线网络发展的内在规律，可以通过网络外部性(network externality)解释。即使不考虑需求量大小，所有节点对之间的互联仍不是效用最大的网络。每条航线的效用 $U(x)$ 可以表示为目标函数

$$U(x) = n(1 - x)p$$

其中 $n$ 为网络中已有的航线数量； $x$ 为航线开通的先后顺序； $p$ 为航线的成本。其中 $n, x$ 先进行归一化， $x$ 越小，航线开通越早。若 $U(x) > 0$ ，则航线开通，反之亦然。对于航空业而言，其利润最大化的目标方程为

$$\begin{cases} \max \pi \hat{x} = p \hat{x} \\ s.t. p = \hat{x}(1 - \hat{x}) \end{cases}$$

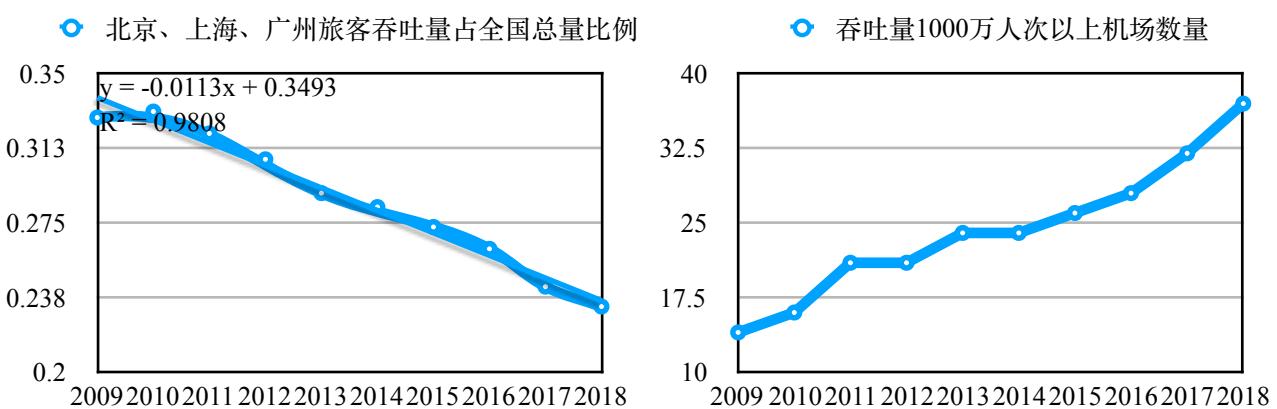
可得 $x = 0$ 或 $\frac{2}{3}$ 时， $\frac{d\pi}{dx} = 0$ ，则总利润 $\pi$ 在 $x = \frac{2}{3}$ 时达到最大值。

在苏延芳等人的论文中，这一理论模型得出了美国航空网络向无标度网络发展的内在动机，即完全点对点的连接与只连接附近节点的小世界网络（这在航空业早期十分常见，因为飞机航程和客流不足，只能采取经停的方式执飞远程航线，各节点实际上只与附近一小部分节点直接连接）并不能满足航空公司盈利的目的。美国航空网络数十年的演化已经证明了这一点，而当前中国航线网络的变化趋势也符合这一规律。

## 五、对中国航线网络具体节点的测度

结合上一部分对于中国国内航线网络结构的总体测度，本部分将会以节点城市为单位对航线网络进行测度。

从总体的吞吐量数据来看2018年，北京、上海、广州三地城市机场占全国境内机场吞吐量23.3%，较上年下降1%；吞吐量200万人次以下机场占全国总吞吐量6.9%，较上年下降0.4%。从吞吐量占全国总吞吐量的比例上看，北京、上海、广州占全国总量的比例一直在持续下降，而相应地，除这三座城市外，吞吐量1000万人次以上的城市航空吞吐量占境内总吞吐量的比例一直在持续上升，其中主要包括成都、西安、昆明、乌鲁木齐等“地方”支线枢纽，以及深圳、杭州、南京等崛起的“新一线城市”。



从上图看，自2009年以来，北京、上海、广州三地吞吐量占全国之比呈几乎线性的下降趋势 ( $R^2 = 0.9808$ )，与之相对应的则是全国吞吐量1000万人次以上机场数量的显著增加，2018年时几乎是2009年的三倍多。

但是节点的“重要性”可以一分为二看待：除了节点流量对总流量的比例外，节点对于网络的通达性以及网络整体的结构变化也是较为重要的因素。此部分采用网络的度和度的中心性 (Degree&Degree Centrality)、中间中心度(Betweenness Centrality)以及网络的接近中心度(closeness centrality)对节点在网络中重要性的变化进行度量。

### (一) 测度指标

#### 1. 度和度的中心性 (Degree&Degree Centrality)

航线网络中的节点 $v_i$ 的度(degree) $k_i$ 指的是与节点直接相邻的航线的数目。一般而言 $k_i$ 越大，节点在网络中越重要。度的表达形式为

$$k_i = \sum_{j=1}^N a_{ij}$$

网络中节点的度中心性(degree centrality)指的是按照度的大小对网络节点进行排列所得到的排名，与节点对于其邻居节点的直接影响力相关。因为不同年份的航空网络网络节点的数量不同，为了便于比较，一般节点的中心性在比较前会进行归一化（即将排名除以矩阵内节点数，最后一名为1）

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

## 2. 中间中心度(Betweenness Centrality)

节点*i*相对于节点对(*p, q*)的中间中心度，指的是节点*i*处于连接*p, q*的最短路径上的可能，或经过点Y且连接*p, q*的最短路径占两点之间的最短路径总数之比。节点*p, q*之间的最短路径的定义为此路径所经过的边数最短，而非物理网络中所经过的距离最短——在本文中即为飞行次数最短，而非飞行距离最短。此时可以定义节点*i*的中间中心度：

$$g(i) = \sum_{i \neq p \neq q} \frac{\sigma_{st}(i)}{\sigma_{st}}$$

若节点的中间中心度上升，则说明其节点对于网络系统的通达性的相对作用上升；若节点的中间中心度下降，则说明节点对网络系统的相对作用下降。通过对2009-2018年全国航空网络各节点的中间中心度进行度量，并随时间进行比较，我们可以得到特定节点在网络通达性上的作用呈上升或是下降的趋势。

## 3. 接近中心度(closeness centrality)

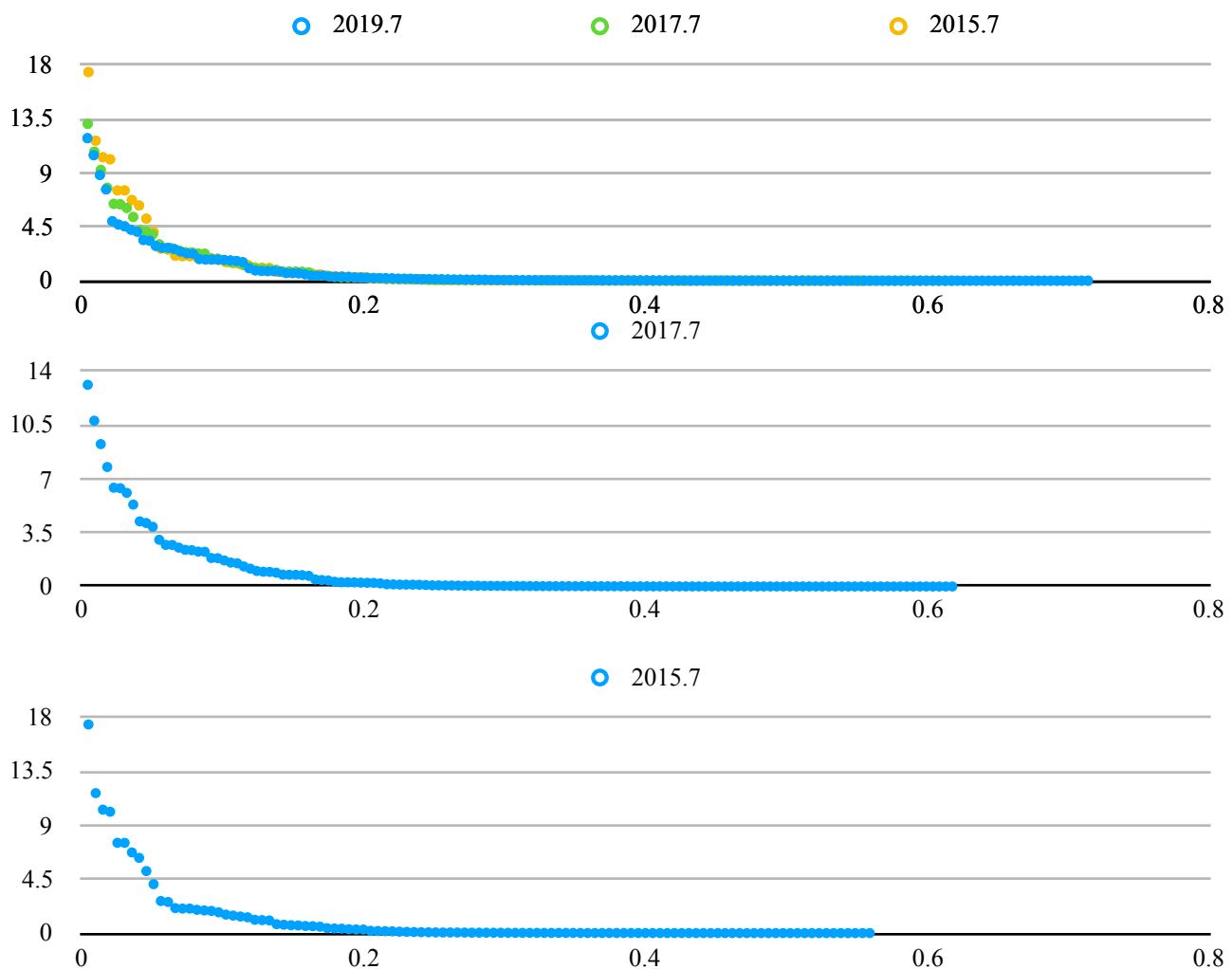
与中间中心度衡量节点在网络中的相对重要性不同，接近中心度衡量的是节点与其相邻节点的通达性，即绝对意义上与其他节点互动的程度。因为中国航空网络中不存在完全不互相连通的节点对，因此可以记节点对(*p, q*)间最短路径长度为*d<sub>pq</sub>*，则对于任意节点*p*，到网络其他任意节点的平均最短路径长度为

$$d_p = \frac{1}{N - 1} \sum_{p \neq q} d_{pq}$$

节点*p*的接近中心度则为 $\frac{1}{d_p}$ 。节点的接近中心度越大，则节点越处于网络中的“中心”地位。

## (二) 测度过程

对2015至2019年中国航线网络的以上三个指标进行测度，以测度得到的值为纵坐标，各节点测度值在网络内归一化的排名为横坐标（对度和度的中心性而言，即度为纵坐标，度的中心性为横坐标），可以得到以下的结果



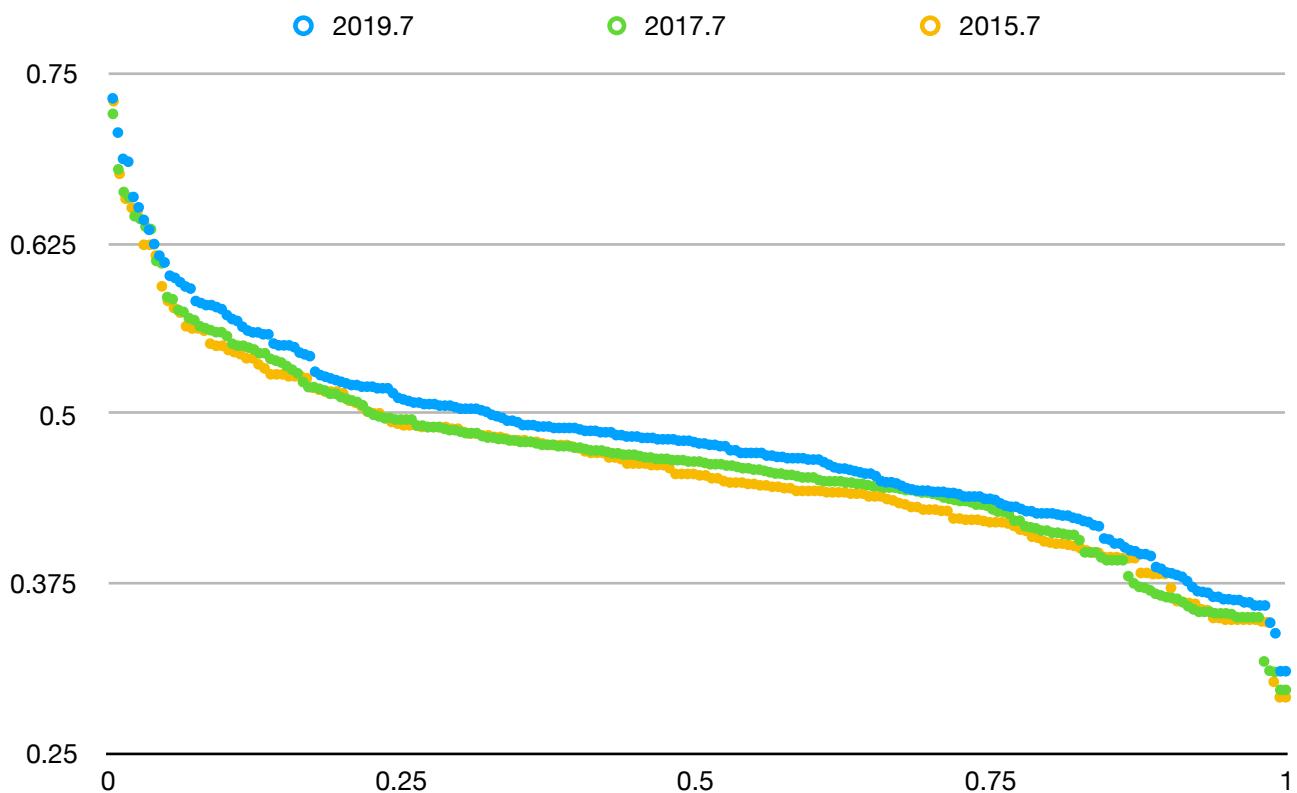
首先是节点的中间中心度( $\times 10^3$ )的分布，上图在绘制时删去了中间中心度过小的点。

第一部分2015年（黄）与2019（蓝）的直接对比显示，中间中心度较高的节点的中间中心度，自2015年至2019年，发生了显著的下降。其次则可以过来自2015, 2017, 2019年的三份数据的对比发现，数据的“尾部”长度在增加，即大部分的中小机场的中间中心度呈现了持续上升的趋势。

造成这种趋势的潜在原因，可能来自于随着需求增加，城市对间的航线增多，原有的大型枢纽对于网络系统的通达性贡献降低，而大部分中小城市因为需求增加，航线数量也增加，因此出现在了中间中心度的分布图上。且因为经济的规模效应，随着周边地区的需求增加，新的机场被修建，原有的大型枢纽的对周边的虹吸效应在下降。对于造成大型枢纽中间中心度下降的潜在原因，本研究将会在后文中进行实证分析。

与节点的中间中心度衡量节点对网络整体通达性的相对贡献不同，节点的接近中心度衡量的是节点本身与周围节点的连接通达情况，受其他节点之间连接状态的影响较小。

2015-19年我国航空网络节点的接近中心度分布情况如下：



通过对2015年、2017年和2019年7月的比较，可以发现网络中大部分节点的接近中心度依然在随着时间而上升，无论其中间接近度上升还是下降。但是我们也可以清晰地发现原本接近中心度较高的节点，随时间而上升的幅度远没有其他节点明显。尽管结合此前判断网络无标度性所做的节点的度与度中心性的测度可以得出，这些节点作为枢纽的地位不可动摇，但是其新增路径的速度没有重要性稍低的节点新增的速度快，因此可能造成了这些中间中心度的下降。对于节点接近中心度的相关潜在印象因素，也将会在本研究的后文涉及。

与此同时另一个可以观察到的现象是，大部分节点的接近中心性都维持在相对客观的水平，与接近中心性最高的节点相差没有节点的度的分布那样悬殊，这也说明了网络的大部分节点都具有尚可的通达性，印证了中国航空网络目前仍然是一个小世界网络的特点。

对于枢纽节点而言，接近中心度近年来仍然在上升，说明依然不断的有新的航线在枢纽节点开航。然而大部分枢纽节点的中间中心度在持续下降，表明其对于网络总体的通达性贡献逐渐减小，因为更多其他节点之间的航线开辟，减少了节点对之间的最短路径经过大型枢纽的概率。而更多其他节点之间的航线开辟，则可以用经济的规模效应解释——当一地区对航线服务的需求增加，则会开辟新的直航航班，从而减弱中心机场的虹吸效应，乘客无需至中心机场中转。

## 六、中国航线网络的影响因素实证分析

前两个部分对我国航空网络的结构以及相关的变化趋势进行了一定的分析，本部分将会联系这些结构背后的一些潜在因素进行实证分析。在前文的测度中，我们发现，尽管网络整体的无标度性上升，但是仍然有一部分节点的中间中心度随时间下降，其中有不少大型枢纽节点。为了对这些节点的中间中心度随时间下降的相关原因以及潜在因素进行测度，我们试使用多元线性回归模型对相关的变量进行测度。

为了将分析聚焦在中间中心度随时间下降的节点上，选取了2015-2019年中间中心度下降的节点作为测度样本。此处实证分析的待解释变量选用了节点2015年与2019年中间中心度变化量。由于航线网络所涉及的要素较为复杂，在此分为三大类7个解释变量，涉及需求与购买力，市场规模与潜在的替代交通方式：

与市场规模带来的需求相关：

当地户籍人口变化率(Citizen\_Population)

当地常住人口变化率(Resident Population)

与需求总量相关：

上年当地机场旅客吞吐量增长率(Airport\_Capacity)

与购买力带来的需求相关：

当地人均可支配收入变化率(Income\_per\_capita)

当地国内生产总值变化率(Gross\_Domestic\_Product)

与需求类型相关：

当地年游客人数变化率(Tourist\_Population)

通过Eviews软件对以上变量进行多元线性回归分析，可以得到结果如下(见下页)

Variable	Coefficient	Std. Error	t-Statistic	Prob
C	-3.347418	2.066556	-1.619805	0.1195
Income_per_capita	-1.493525	4.156342	-0.359336	0.7228
Gross_Domestic_Product	-2.1070784	1.655817	-1.311005	0.2034
Civil_Automobile	2.156336	1.226769	1.757736	0.1327
Citizen_Population	2.119579	3.289123	0.644421	0.5260
Resident_Population	3.772241	4.074474	0.925823	0.3646
Tourist_Population	1.476518	0.945646	1.561386	<b>0.0927*</b>
Airport_Capacity	0.889559	0.392333	2.267360	<b>0.0335*</b>
<i>R</i> <sup>2</sup>				0.69
Adjusted <i>R</i> <sup>2</sup>				0.38

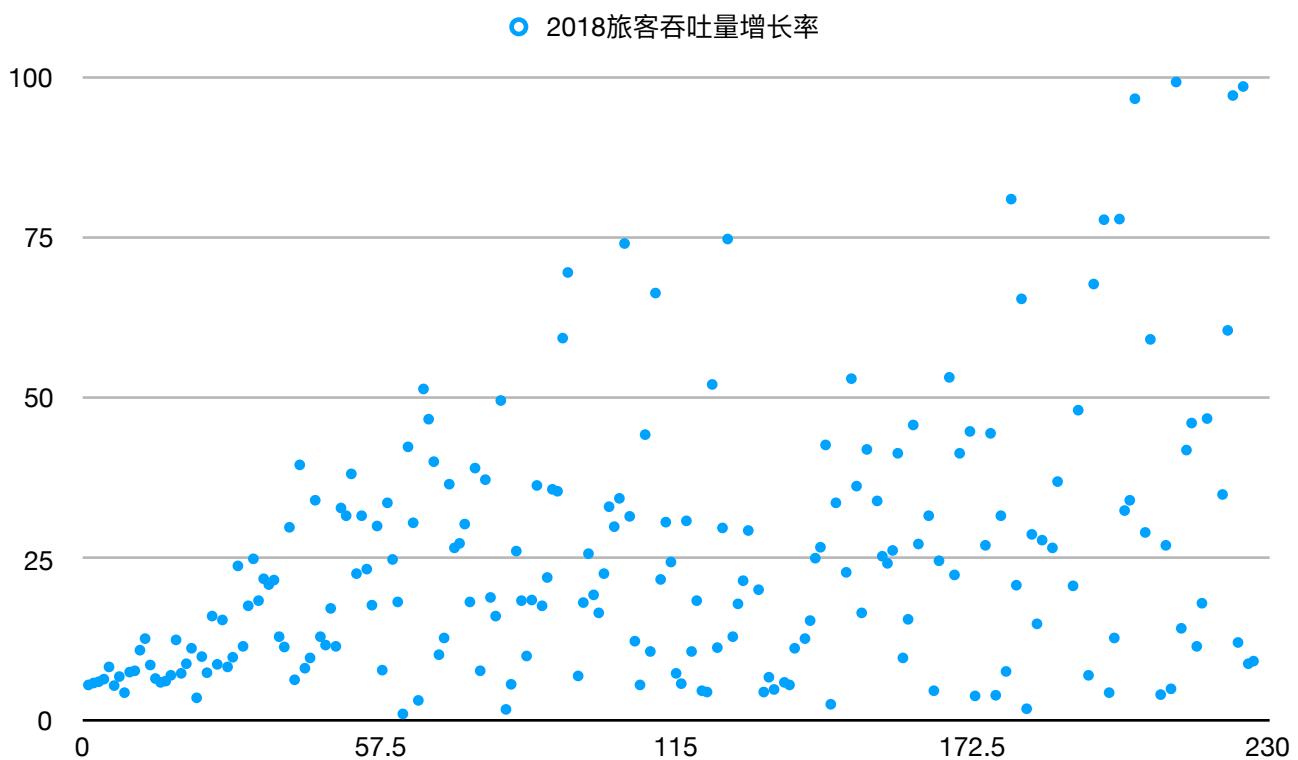
多元分析得到的结果证明了年游客人数变化率以及机场的旅客吞吐量变化率（增长率）与节点中间中心度的变化率有着较为显著的联系。相较之下，人均可支配收入和地区生产总值的增长，以及户籍人口与常住人口的增长速度，和当地机场在网络中的相对重要性的增长并无明显的相关性。

与预期的替代关系不同，当地民用汽车总数与机场节点的的相对重要性反而呈现了微弱的正相关性，说明汽车和飞机作为交通工具的相互替代性非常微弱，而此时微弱的正相关性来源于汽车拥有量和当地生活水平的联系，而航空运输的需求同样与当地生活水平有关，因此并不是一个有效的测度标准。

机场在网络中的重要性的增速和人口增速关联度较低，不仅说明了当前人口增长较快，人口基数较大的地区基本已经被航空网络覆盖，更说明当前与人口相关的潜在需求规模，对航空运输的需求中体现的并不明显。当前航空网络规模的增量，不再是来自人口增长的市场规模的增加，而是来自人民生活水平以及出行动机的提高。

机场在网络中重要的变化与代表总体经济活动的GDP的变化和人均可支配收入的变化相对较低的相关度，以及和年游客人数变化相对较高的相关度，说明当前中国航空网络规

模的增量以观光旅游为主，而非总体的经济和商业活动，因此新增的节点也多以具有旅游价值，而非具有商业价值的目的地为主，机场规模和在网络中的重要性的增长与目的地当地的经济发展水平关联性也在降低（因为前往这些重要性增加的新兴节点的旅客多为旅游者，如图木舒克及额济纳/塞汗桃来等）。



上图为横坐标为2018年中国航空网络节点的旅客吞吐量排名，而纵坐标为当年节点旅客吞吐量的同比增长百分点。由此可以发现，增长率较高的节点大多分布在吞吐量排名靠后的部分，即一般而言，吞吐量增长率较大的节点原本规模较小。再通过前文实证分析得到的中国航空网络中节点的相对重要性的增长速度与客流量增长速度的显著关联，可以得到节点相对重要性和原本节点规模的潜在联系，也可以解释此前测度时发现大型节点普遍相对重要性在下降，而小型节点普遍相对重要性在上升的问题——各节点的客流量存量不同，而增长速度不同；规模越小的节点更有可能增长率较大，从而中间中心度（相对重要性）增长。这验证了经济的规模效应——当节点的需求量跨过某个“门槛”时，从该节点而非附近的枢纽始发的航班将会具有正规模效应，因此增长率将会比网络中节点增长的普遍值大。

## 七、总结与展望

本文的研究思路从网络整体到具体节点，再到节点个体分布的其中一个现象的潜在原因，从宏观到微观，从抽象到具体。

首先通过对网络平均路径长度，局部集聚系数与节点度的分布的测度，本文得到，相较于美国航空网络等无标度航线网络，中国的国内航线网络仍是一个小世界网络，具有高集聚系数与低平均路径长度；但是随着网络逐渐的发展以及管制的放松，因为网络的外部性，中国的航线网络的小世界性正在逐渐减弱，而无标度性在逐渐增强，节点的度的分布逐渐接近大尺度无标度网络的定义。结合参考文献得出的结论，中国的航空网络演变特点与美国的航空网络类似，都经历了或正在经历在放开管制后从小世界网络过渡至无标度网络的过程，与枢纽节点直接相连的航线量将会长期持续增加。

但是回到具体节点的测度上，我国航空网络的一部分枢纽节点在网络中的相对重要性，即中间中心度不升反降，通过进一步的测度与实证分析，本文发现节点在网络中的中间中心度的变化，与节点对于航空运输的需求量增速显著相关。因此，各地在航空网络发展时增速的不均，势必造成航空网络中枢节点通达性的相对重要性的下降，但是与其直接相连的节点数与旅客流量不会下降。我们认为各城市间对航空运输的需求量的增速表现出的不同，与经济的规模效应有关——当越过某一“门槛”时，从节点出发的航线才会有经济效益，因此航线才会开设，因此在地区之间经济增长速度相同的情况下，对航空运输的需求的增速是不连续的，从而造成航空运输存量不同的节点，即使经济增速相同，体现出的航空运输需求也是不同的。通常重要性较低的城市会越过较多这样的“门槛”，因此一般会有更高的需求增速，从而带来节点重要性的快速上升。

此外，在实证分析的过程中，我们发现当前航线网络节点的重要性，与国内生产总值以及可支配收入等总体购买力相关的指数，以及户籍人口和常住人口等与市场规模上相关的指数关联性不高。相反地，当前我国航线网络节点的相对重要性与旅客人数的增量息息相关，反映了当前我国航空运输的需求增量更多是由旅游需求的增加带来的，而旅游需求的增加更多来自消费者观念意识的转变（通常消费观念会相比消费力的变化滞后），而非购买力绝对值的增加。

相较于前人的研究，考虑到目前中国航线网络快速发展的现实，本文采用最新的数据，对中国航空网络从宏观到微观的变化趋势及其内在逻辑进行了分析，而非仅仅简单地进

行测度，从而得出了中国航空网络未来将会从小世界网络逐渐变为无标度网络的趋势。由于中国航空网络尚未完全成熟，新开辟的节点（或原本较小的节点）在经济的规模效应（或“门槛效应”）影响下，增速超过了原有的大型及枢纽节点，造成原有枢纽节点总体重要性的下降；替代路线随着节点间连接增加而增加，原有的枢纽对于网络的通达性的贡献大不如前，而对网络的相对重要性在下降。

此外，论文的实证分析也为中国航空运输需求开辟了新的视角。从节点的年游客数同比增长率与节点相对重要性的增长率的密切联系可以得到，在当下以及不远的未来，随着人们消费的意识的转变而增加的旅游需求，将会成为航空运输需求增量的主要部分。

从当前中国航线网络演化的趋势看，短期内新增的航线将大部分以非枢纽之间的连接为主，而从长期的趋势看，枢纽的连接度也将会进一步增长；旅游客流将会取代商务客流，成为未来一段时间内需求增量的主流。因此，在机型的选择上，高度的灵活性将会成为第一标准——即能服务短程直线，也能服务主要干线；既能承载商务客流，也能承载旅游客流。旅游客流的特点是对价格敏感，对路线高度定制化，但对便捷性与舒适性的要求没有商务客流高。这对航空公司的管理提出了全新的，更高的要求。因此用途较为单一的超大型客机势必无法成为主流，而在使用范围和操作的灵活性将会史无前例地重要。

因此未来的我国民航市场，在机型的选择上应当充分考虑到使用的灵活性，而在运营上将会高度精细化，充分考虑以旅游客流为代表的各细分乘客群体的需求（相比之下，十年前随着超大型客机引入的是对商务客流的充分考虑）。同样地，航空公司的运营架构，将会在保证枢纽的地位的前提下，进一步向非枢纽城市下沉。以南方航空为例，其以其枢纽的广州飞行部为班底成立了南航飞行总队，巩固其在广州枢纽的地位；同时成立了子公司重庆航空，进一步挖掘较小节点的规模效应带来的新增需求红利。在可见的未来，随着超大型客机消失带来的，是航空网络演化的进一步复杂化，从而给航空公司带来新的巨大挑战。

## 附录1：获取数据时采用的爬虫程序说明

分析数据来自于飞常准官网的航线数据部分（URL: <http://map.variflight.com/stat>），使用基于Python3.7编写的网络爬虫进行爬取并归档入库。爬虫程序的主题部分代码如下：

```
from symbol import parameters

import random

import pymysql

import requests

doc = open('airline.txt','w')

import json

url = 'http://map.variflight.com/mst?_c=TTT'

user_agent_list = [
    "Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.1 "
    "(KHTML, like Gecko) Chrome/22.0.1207.1 Safari/537.1",
    "Mozilla/5.0 (X11; CrOS i686 2268.111.0) AppleWebKit/ "
    "536.11 "
    "(KHTML, like Gecko) Chrome/20.0.1132.57 Safari/536.11",
    "Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/536.6 "
    "(KHTML, like Gecko) Chrome/20.0.1092.0 Safari/536.6",
    "Mozilla/5.0 (Windows NT 6.2) AppleWebKit/536.6 "
    "(KHTML, like Gecko) Chrome/20.0.1090.0 Safari/536.6",
    "Mozilla/5.0 (Windows NT 6.2; WOW64) AppleWebKit/537.1 "
    "(KHTML, like Gecko) Chrome/19.77.34.5 Safari/537.1",
    "Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/536.5 "
    "(KHTML, like Gecko) Chrome/19.0.1084.9 Safari/536.5",
    "Mozilla/5.0 (Windows NT 6.0) AppleWebKit/536.5 "
    "(KHTML, like Gecko) Chrome/19.0.1084.36 Safari/536.5",
```

```
"Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/536.3 "
"(KHTML, like Gecko) Chrome/19.0.1063.0 Safari/536.3",
"Mozilla/5.0 (Windows NT 5.1) AppleWebKit/536.3 "
"(KHTML, like Gecko) Chrome/19.0.1063.0 Safari/536.3",
"Mozilla/5.0 (Macintosh; Intel Mac OS X 10_8_0)

AppleWebKit/536.3 "

"(KHTML, like Gecko) Chrome/19.0.1063.0 Safari/536.3",
"Mozilla/5.0 (Windows NT 6.2) AppleWebKit/536.3 "
"(KHTML, like Gecko) Chrome/19.0.1062.0 Safari/536.3",
"Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/536.3 "
"(KHTML, like Gecko) Chrome/19.0.1062.0 Safari/536.3",
"Mozilla/5.0 (Windows NT 6.2) AppleWebKit/536.3 "
"(KHTML, like Gecko) Chrome/19.0.1061.1 Safari/536.3",
"Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/536.3 "
"(KHTML, like Gecko) Chrome/19.0.1061.1 Safari/536.3",
"Mozilla/5.0 (Windows NT 6.1) AppleWebKit/536.3 "
"(KHTML, like Gecko) Chrome/19.0.1061.1 Safari/536.3",
"Mozilla/5.0 (Windows NT 6.2) AppleWebKit/536.3 "
"(KHTML, like Gecko) Chrome/19.0.1061.0 Safari/536.3",
"Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/535.24 "
"(KHTML, like Gecko) Chrome/19.0.1055.1 Safari/535.24",
"Mozilla/5.0 (Windows NT 6.2; WOW64) AppleWebKit/535.24 "
"(KHTML, like Gecko) Chrome/19.0.1055.1 Safari/535.24"

]
```

```
headers = {

'Accept': 'application/json, text/javascript, */*; q=0.01',
'Accept-Encoding': 'gzip, deflate',
'Accept-Language': 'zh,zh-CN;q=0.9,en;q=0.8',
'Connection': 'keep-alive',
'Content-Type': 'application/x-www-form-urlencoded; charset=UTF-8',
'Origin': 'http://map.variflight.com',
```

```

'User-Agent': random.choice(user_agent_list),
'Referer': 'http://map.variflight.com/stat',
'X-Requested-With': 'XMLHttpRequest',
'Cookie': '_ga=GA1.2.1619466040.1567428928;
Hm_lvt_d1f759cd744b691c20c25f874cadc061=1567428861,1567493603,1567
510128; PHPSESSID=oogonbhjbtqjusprg19oa31sr2;
Hm_lvt_2ad30bb2583cfacb7d56e2a6e1444d4b=1567772903,1567954693,1567
992450,1567997232;
Hm_lpvt_2ad30bb2583cfacb7d56e2a6e1444d4b=1567997798',
'Host': 'map.variflight.com'
}

```

```

parameters =
"queryDate1=D1&queryDate2=D2&isDirect=1&isStop=0&isOwn=1&isShare=0
&obj=AF&type=APTx&isCX=1&isC3=0&isNC=0"

date1 = ['2015-01-01', '2015-07-01', '2016-01-01', '2016-07-01',
'2017-01-01', '2017-07-01', '2018-01-01', '2018-07-01', '2019-01-
01',
'2019-07-01']
date2 = ['2015-01-07', '2015-07-07', '2016-01-07', '2016-07-07',
'2017-01-07', '2017-07-07', '2018-01-07', '2018-07-07', '2019-01-
07',
'2019-07-07']

obj =
['AXF', 'EJN', 'RHT', 'THQ', 'AAT', 'ACX', 'AEB', 'AHJ', 'AKU', 'AOG', 'AQG', 'AVA',
'BAR', 'BFJ', 'BPL', 'BSD', 'CDE',
'CHG', 'CIF', 'CIH', 'CWJ', 'DAT', 'DAX', 'DBC', 'DCY', 'DDG', 'DLU', 'DQA', 'DTU',
'ENH', 'ENY', 'ERL', 'FUG', 'FUO', 'FYJ', 'FYN', 'GMQ',
'GYS', 'GYU', 'GZG', 'HCJ', 'HDG', 'HEK', 'HIA', 'HJJ', 'HLD', 'HLH', 'HMI', 'HNY',
'HPG', 'HTT', 'HUO', 'HUZ', 'HXD', 'HYN', 'HZG', 'HZH',
'IQM', 'IQN', 'JDZ', 'JGD', 'JGN', 'JGS', 'JIC', 'JIQ', 'JMJ', 'JMU', 'JNG', 'JNZ',
'JSJ', 'JUH', 'JUZ', 'JXA', 'JXH', 'KCA', 'KGT', 'KJH',
'KJI', 'KOW', 'LCX', 'LDS', 'LFQ', 'LLB', 'LLF', 'LLV', 'LNJ', 'LNL', 'LPF', 'LUM', 'L

```

```
LZO','LZY','MDG','MXZ','NAO','NAY','NBS','NDG',
'NLH','NLT','NZL','OHE','PZI','QSZ','RIZ','RKZ','RLK','RQA','SHF','SQD',
'SQJ','SYM','TCG','TCZ','TEN','TGO','TNH','TVS',
'TWC','UCB','UYN','WDS','WGN','WMT','WNH','WSK','WUA','WUS','WUT','WUZ',
'WXN','XAI','XIL','XNT','YBP','YIC','YIE','YIN',
'YIW','YKH','YNJ','YNZ','YSQ','YUS','YXG','YYA','YZY','ZAT','ZHY','ZQZ',
'ZI','BAV','BHY','BPE','BPX','CGD','DIG','DNH',
'DOY','DYG','GOQ','HSN','HTN','JHG','JJN','JZH','KRL','KRY','LJG','LYA',
'LYG','LYI','LZH','MIG','NGQ','NNY','NTG','NZH',
'SWA','TXN','WEF','WEH','XFN','XIC','XUZ','YCU','YIH','ZHA','CGQ','CZX',
'DLC','DSN','FOC','HAK','HET','HFE','HRB','INC',
'KHG','KHN','KWE','KWL','LHW','LXA','NGB','NNG','SHA','SHE','SJW','SYX',
'TAO','TLQ','TNA','TYN','URC','WNZ','WUX','XMN',
'XNN','YNT','YTY','ZUH','CAN','CGO','CKG','CSX','CTU','HGH','KMG','NKG',
'PEK','PKX','PVG','SZX','TSN','WUH','XIY']
```

```
for airports in obj:
    temp = parameters.replace('AF', airports)
    #print(temp)

    for i in range(10):
        temp1 = temp.replace('D1',date1[i])
        temp2 = temp1.replace('D2',date2[i])
        print(temp2)
        json_data = requests.post(url, data=temp2, headers=headers)
        #json_data = json_data.encoding("utf-8")
        #json_data = json.loads(json_data)
        #json_data = json_data['data']['main']
        json_data.encoding = 'utf-8'
        html = json_data.text
        j = json.loads(html)[ "data" ][ "main" ]
        print(j)
        print(temp2,file=doc)
        print(j,file=doc)
```

```
doc.close()
```

飞常准的航线数据网页使用的是基于Ajax的网页动态加载技术，表现为修改查询数据进行查询时网页的URL没有发生变化。因此无法使用基本的get指令直接爬取URL。在浏览器中按下F12进入开发者模式，选中Network分类的XHR数据查看网页加载时的请求数据。网页的请求数据如图所示：

从请求数据可以知道该网页的数据请求方法为Post，请求数据的源网页为[http://map.variflight.com/mst?\\_c=TTT](http://map.variflight.com/mst?_c=TTT)。在Request Headers中可以得到请求头，Form Data则是动态网页加载不同数据时向服务器发送的数据请求项。请求头是向服务器请求数据时指明请求类型的数据。因为本程序爬取都是同一类数据，无需更换请求头，直接按照开发者模式中查询到的请求头向服务器模拟请求即可。

如图中我所查询的是北京首都机场的航班数据，查询选项包括直飞、经停和执飞，地理选项则是境内和跨境。因此请求数据中isDirect, isDomestic, isStop, isCross和isOwn的请求数据都是1，代表我选择了此类数据，isShare则是0，代表我没有选择共享航线。“obj”则是我查询的机场代码，type则是飞常准服务器的一个自定请求数据。经过研究和试验发现，飞常准的数据查询结果可以展开成详细数据，当点击“查看”按钮查看详细数据时请求数据为APTx，请求概览数据时则是APT。在了解飞常准服务器的请求原理后，可以通过爬虫程序向服务器模拟请求数据，通过更换请求数据类型来获得不同数据。

程序中的“parameters”参数部分就是要向服务器模拟请求的数据，“queryDate1”和“queryDate2”是查询的起始时间，“obj”则是查询的机场。下表所示便是用于时间和机场循环的列表。

程序的第一个for循环用于从列表顺次中选取机场来替换参数中的AF。

第一个for循环中嵌套的第二个for循环用于替换参数中的D1和D2。

大部分网站都有针对爬虫的反爬措施，短时间内的同ip大量访问会导致服务器封锁制定ip。因此爬虫程序需要基于ip池准备反爬措施，在爬取中定时更换ip和user-agent，混淆网页反爬。如图所示是向网页模拟的请求头：

其中User-Agent用于模拟不同浏览器发出的申请进行反爬混淆。使用random函数从user\_agent\_list中随机选取使用。

“requests”函数用于按照上面已经分析出的请求头和请求数据向网页发送模拟请求并获取网页回复，是整个程序的核心函数。将返回的数据赋值给json\_data。

飞常准的网页使用基于Unicode的UTF-8进行编码，因此获取服务器返回数据后也要对数据进行编码，将编码后的json文件的字符串赋值给html，使用json.load函数抓取字符串中data中main的数据部分，赋值给j，将j的值写入txt文档并保存。

## 附录2: 致谢

首先，感谢我的另两位搭档过硬的水平与投入的态度，为本次论文的完成作出了不可替代的贡献。其中彭彦迪同学主要负责文献综述，以及对航班量数据矩阵与实证分析的数据处理；范思齐同学主要负责获得航班量数据的网络爬虫程序的编写；我负责其余部分的写作与统稿。

在此，也向我的学校华南师范大学附属中学以及指导老师黄秉刚老师表示诚挚的感谢。正是学校的氛围以及老师的鼓励促使我参加了本次丘成桐科学奖，论文的完成也离不开黄秉刚老师的悉心指导。

感谢中山大学的杨子晖教授，对本论文的测度方法和实证分析提出了宝贵的建议。除了他和黄秉刚老师，还有无数的人曾经对本论文提出大量具有建设性的建议，在此一并表示感谢。

最后，还需要感谢我和两位搭档的家人和朋友们，在本文的研究与写作期间一直陪伴并支持着我们的工作。没有他们的支持，论文同样难以完成。

严子渔

2019.9.15

## 参考文献

李健. “我国航空运输业放松管制绩效研究.”工业技术经济, no. 11, Nov. 2011, pp. 36-40.

荣朝和. 西方运输经济学. 经济科学出版社, 2002.

苏延芳, 肖兴志. “反事实研究：放松规制与中国航空网络演变.”中国民用航空, Jul. 2003, pp. 72-83.

闻中, 陈剑. “网络效应与网络外部性：概念的探讨与分析.”当代经济科学, vol 22, no. 6, 2000, pp. 13-20.

吴泗宗, 蒋海华. “对网络外部性的经济学分析.”同济大学学报(社会科学版), vol 13, no. 6, Dec. 2002, pp. 70-77.

张永莉, 张晓全. “我国城市间航空客运量影响因素的实证分析.”经济地理, vol 27, no. 4, Jul. 2007, pp. 658-661.

刘宏鲲, 张效莉, 曹嵬, 汪秉宏, 周涛. “中国城市航空网络航线连接机制分析.”中国科学, vol 39, no. 7, 2009, pp. 935-942.

闫玲玲, 陈增强, 张青. “基于度和聚类系数的中国航空网络重要性节点分析.”智能系统学报, vol 11, no. 5, Oct. 2016, pp. 586-593.

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.” 交通运输系统工程与信息, vol 11, no. 6, Dec. 2011, pp. 176-181.

薛俊菲. “基于航空网络的中国城市体系等级结构与分布格局.”地理研究, vol 27, no. 1, Jan. 2008, pp. 23-33.

杨涛, 曹文娟. “网络影响因子及其测度.”图书情报工作, no. 9, 2002, pp. 24-27.

李煜伟, 倪鹏飞. “外部性、运输网络与城市群经济增长.”

吴思竹, 张智雄. “网络中心度计算方法研究综述.”

杜海峰, 李树苗, 悅中山等. “小世界网络与无标度网络的社区结构研究.”物理学报, vol 56, no. 12, Dec. 2007, pp. 6886-6893.

王姣娥, 景悦. “中国城市网络等级结构特征及组织模式——基于铁路和航空流的比较.”地理学报, vol 72, no. 8, Aug. 2017, pp. 1508-1519.

金凤君, 王成军.“轴 - 辐侍服理念下的中国航空网络模式构筑.”*地理研究*, vol 24, no. 5, Sep. 2005, pp. 774-784.

刘世铎, 吴群琪.“基于运输需求的公路网可达性研究.”*长安大学学报*, vol 12, no. 1, Mar. 2010, pp. 40-45.

王兴隆, 张淑婷.“航空网络影响力节点识别与鲁棒性分析.”*科学技术与工程*, vol 19, no. 8, Mar. 2019, pp. 297-302.

皮天雷.“计量经济学中函数设定形式的检验.”*统计与决策*, no. 11, 2009, pp. 19-20.

中国民用航空局. *2018年民航行业发展统计公报*. May 2019.

Choi, Jay Pil. “Network Externality, Compatibility Choice, and Planned Obsolescence.” *The Journal of Industrial Economics*, vol. 42, no. 2, 1994, pp. 167–182. *JSTOR*, [www.jstor.org/stable/2950488](http://www.jstor.org/stable/2950488).

Economides, Nicholas. “Desirability of Compatibility in the Absence of Network Externalities.” *The American Economic Review*, vol. 79, no. 5, 1989, pp. 1165–1181. *JSTOR*, [www.jstor.org/stable/1831443](http://www.jstor.org/stable/1831443).

Guimerà, R., et al. “The Worldwide Air Transportation Network: Anomalous Centrality, Community Structure, and Cities' Global Roles.” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 22, 2005, pp. 7794–7799. *JSTOR*, [www.jstor.org/stable/3375673](http://www.jstor.org/stable/3375673).

Liebowitz, S. J., and Stephen E. Margolis. “Network Externality: An Uncommon Tragedy.” *The Journal of Economic Perspectives*, vol. 8, no. 2, 1994, pp. 133–150. *JSTOR*, [www.jstor.org/stable/2138540](http://www.jstor.org/stable/2138540).

Nagler, Matthew G. “NEGATIVE EXTERNALITIES, COMPETITION AND CONSUMER CHOICE.” *The Journal of Industrial Economics*, vol. 59, no. 3, 2011, pp. 396–421. *JSTOR*, [www.jstor.org/stable/41289460](http://www.jstor.org/stable/41289460).

Parker, Geoffrey G., and Marshall W. Van Alstyne. “Two-Sided Network Effects: A Theory of Information Product Design.” *Management Science*, vol. 51, no. 10, 2005, pp. 1494–1504. *JSTOR*, [www.jstor.org/stable/20110438](http://www.jstor.org/stable/20110438).

Shy, Oz. *the Economics of Network Industries*. Cambridge Univ. Press, 2001.

Xu, Zengwang, and Robert Harriss. “Exploring the Structure of the U.S. Intercity Passenger Air Transportation Network: a Weighted Complex Network Approach.” *GeoJournal*, vol. 73, no. 2, 2008, pp. 87–102. *JSTOR*, [www.jstor.org/stable/41148284](http://www.jstor.org/stable/41148284).