

# 利用链路预测推断网络演化机制

刘宏鲲<sup>①</sup>, 吕琳媛<sup>②</sup>, 周涛<sup>③④\*</sup>

① 西南财经大学统计学院, 成都 610074;

② 弗里堡大学物理系, 瑞士弗里堡, CH-1700;

③ 电子科技大学互联网科学中心, 成都 610054;

④ 中国科学技术大学近代物理系, 合肥 230026

**摘要:** 直接建立演化模型推测影响网络演化的因素是目前研究网络演化机制的常用方法, 但由于可供比较的结构特征量太多, 不同的模型之间难以进行定量化的比较。链路预测方法能清晰直观地对各种因素进行辨别, 因此提供了一种更好的定量化的方法。本文首先介绍基于节点相似性的链路预测方法, 然后讨论利用链路预测推测网络演化机制的基本框架。在以中国城市航空网络为例的实证分析中发现, 当单独利用结构(共同邻点)和节点属性(距离、人口、GDP 和第三产业产值)作为相似性算法的因素时, 基于共同邻居的算法预测准确度最高, 暗示网络演化主要受结构因素影响, 其次才是外在因素。而将四种基于节点属性的算法与基于结构的算法耦合进行计算时, 共同邻点配合第三产业产值效果最好, 并且在预测效果最佳时对应的最优参数处, 第三产业产值在其中承担了比共同邻点更为重要的作用。本文为研究网络演化模型提供了全新的视角和分析工具。

**关键词:** 链路预测 复杂网络 演化机制 航空网络

自从对复杂网络的研究热潮出现后, 对网络结构性质的研究吸引了众多学者的目光, 包括度分布、集聚性质、社团结构、节点中心性、节点间的路径长度等等。不但研究各种网络静态结构的成果众多, 而且从动态的角度揭示网络演化的机制、探寻各种微观因素对网络结构形成的影响, 成果也丰厚<sup>[1-7]</sup>。

直接建立演化模型推测影响网络演化的因素是目前研究网络演化机制的常用方法。基于节点度的优先连接机制可以用来产生演化的无标度网络(网络具有幂律度分布), 例如, Barabasi-Albert (BA) 模型<sup>[7]</sup>, 该模型以节点度为指标, 通过优先连接机制, 生成了幂指数为3的度分布。当然, 优先连接仅仅是生成无标度网络的可能机制之一, 尔后许多研究者提出了多种不同于优先连接的其他可以导致幂律度分布的演化机制<sup>[8-11]</sup>。这类主流建模方法的基本思路是, 对基于某些因素构建出的生成网络分析其统计特征, 如果具有和真实网络接近的统计性质, 那么就认为这些因素对网络的结构影响显著, 也即这些因素是网络演化的重要机制, 否则认为这些因素对网络结构的影响不显著。但是, 考虑分别由不同因素驱动的演化

---

国家自然科学基金重点项目 10635040, 国家自然科学基金面上项目 11075031, 瑞士国家自然科学基金项目 200020-121848, 西南财经大学“211 工程三期”统计学国家重点学科建设项目, 西南财经大学“211 工程三期”青年教师成长项目资助。

\*E-mail: zhutou@ustc.edu

模型，那些衡量模拟网络与真实网络相似度的众多结构量化指标，很有可能表现并不一致，以致难以辨析哪些才是影响网络演化的主导因素，以及这些主导因素在网络演化过程中分别起到了多大的作用。举个例子来说，计算机互联网的理论模型非常多，其中有些能够更精确地重现网络度分布和异配特性<sup>[12]</sup>，另一些则能够更好刻画网络 $k$ -core分解后的结构<sup>[13]</sup>，那么到底哪类模型更好呢，能不能在一个统一的框架下进行比较呢？传统建模研究无法回答这些问题，事实上也无法建造一个统一的平台比较这些模型——这些局限性一定程度上制约了网络演化模型研究的发展。

近几年，网络链路预测受到了广泛的关注。网络中的链路预测是指根据网络中节点的特征或已经存在的边（结构特征），预测两个节点间边的存在性<sup>[14,15,16]</sup>。这种预测既包含了对未知链接（existent yet unknown links, missing links）的预测，也包含了对未来链接（future links）的预测。其中，基于节点相似性的链路预测是链路预测研究的主流方法之一，该类方法的一个重要假设就是，两个节点之间相似性越大，它们之间存在链接的可能性就越大。相似性的定义可以有很多种，最简单的是基于节点属性的定义<sup>[17]</sup>。如果两个节点拥有许多共同的特征，就认为这两个节点是相似的。但是，由于很多情况下获取节点属性信息非常困难，因此在一些系统中基于属性的相似性算法很难实现，例如在线社交网络中用户的个人信息是保密的或者是虚假的。另外一类更加可靠的方法是基于网络结构的相似性，称为结构相似性。Liben-Nowell 和 Kleinberg 提出了基于网络拓扑结构相似性的定义方法，并将这些指标分为基于节点和基于路径两类<sup>[18]</sup>，在对大型科学家合作网络进行的实证研究中，他们发现仅考虑节点共同邻居的方法和 Adamic-Adar Index<sup>[19]</sup>（AA 指数）是预测准确性最好的方法。周涛、吕琳媛和张翼成用9种基于局部信息的相似性指标对6种现实网络进行了准确性的对比，进一步验证了 Liben-Nowell 和 Kleinberg 的研究结果，并提出两种准确性更高的相似性指标：资源分配指数（resource allocation index）和局部路径指数（local path index）<sup>[20]</sup>。最近其他小组的研究结果显示，新提出来的相似性指标在进行群落划分<sup>[21]</sup>、含权网络权重设置<sup>[22]</sup>和处理含噪网络链路预测<sup>[23]</sup>的时候也比原有指标好。一些更复杂的物理过程，例如局部随机游走，最近也被应用于度量网络相似性，并籍此提高链路预测的准确性<sup>[24]</sup>。

总的来说，链路预测就是在网络中根据节点的属性或已经存在的边（结构特征），选取某一因素或混合因素作为基础，通过计算各种预测方法的准确性，找到适合某一网络的最佳预测方法，从而预测网络中未知的边和未来可能产生的边。事实上，这种方法为挖掘演化模型重要驱动因素和公平评价提供了可能性。与直接通过建立演化模型找到影响网络演化的因素相比，由于链路预测能够量化预测方法的准确度，可以更加清晰直观地对各种因素进行细致辨别，因此，链路预测在分析网络演化机制上提供了更好的可比较的量化方法。在本文中，我们将首先介绍链路预测的方法与评价指标，然后讨论如何利用此方法推测决定网络演化的主要因素，最后以中国城市航空网络为例验证此方法的有效性。

## 1. 链路预测的方法与评价指标

考虑无向的简单网络  $G(V, E)$ ,  $V$  是节点集合,  $E$  是边的集合, 不考虑多重边和自连边。网络总的节点数为  $N$  ( $N = |V|$ ), 边数为  $M$  ( $M = |E|$ )。由此得该网络共有  $N(N-1)/2$  个节点对, 即全集  $U$ 。给定一种链路预测的方法, 对每对没有连边的节点对  $x, y$  ( $\in U \setminus E$ ) 赋予一个分数值  $s_{xy}$ 。由于  $G$  是无向的, 因此, 分值是对称的, 即  $s_{xy} = s_{yx}$ , 然后将所有未连接的节点对按照该分数值从大到小排序, 排在最前面的节点对出现连边的概率最大。

为了检测算法的准确性, 需把已知边的集合  $E$  随机地分为两部分: 一部分是训练集  $E^T$ , 作为已知信息用来计算分数值; 另一部分是测试集合  $E^P$ , 用来进行测试, 这个集合中的信息不能用来进行预测。显然,  $E = E^T \cup E^P$ , 并且  $E^T \cap E^P = \emptyset$ 。例如, 将网络所有连边的 10% 作为测试边从网络中删除, 从而根据剩下的 90% 的边的信息预测那些被删除的链边。在此, 将属于  $U$  但不属于  $E^T$  的边定义为不存在的边。

衡量链路预测算法精确度的指标有三种  $AUC$ ,  $Precision$  和  $Ranking Score$ 。三种衡量指标的侧重点不同,  $AUC$  (area under the receiver operating characteristic curve) [25] 是从整体上衡量算法的精确度,  $Precision$  只考虑排在前  $L$  位的链边, 而  $Ranking Score$  更多考虑了所预测的边的排序。关于这三种指标的详细介绍可参见文献[18]。

本文采用  $Ranking Score$  的评价方法。  $Ranking Score$  主要考虑测试集中的边在最终排序中的位置。令  $H = U - E^T$  为未知边的集合 (包括测试集中的边),  $r_i$  表示未知边  $i \in E^P$  在排序中的排名。那么这条未知边的  $Ranking Score$  值为  $RankS_i = r_i / |H|$ , 遍历所有在测试集中的边得到系统的  $Ranking Score$  值为

$$RankS = \frac{1}{|E^P|} \sum_{i \in E^P} RankS_i = \frac{1}{|E^P|} \sum_{i \in E^P} \frac{r_i}{|H|} \quad (3)$$

显然,  $RankS$  值越小表示排在越前面, 也就意味着被成功预测的概率越大, 因此算法精确度越高。

在有些规模较小的网络中, 一种更加精确的数据集划分方法是  $leave-one-out$ , 即每次从网络中选取一条边作为测试边, 然后预测这条边出现的可能性, 然后应用  $Ranking Score$  对这条边的预测效果进行评价。对网络的所有边都进行一次这样的预测 (一共进行  $M$  次), 得到  $RankS$  的平均值, 即为整个网络的预测精度。值得注意的是, 这种方法由于每次都要重新计算相似性, 因此不适用于规模很大的网络。

链路预测的本质是挖掘导致连边产生的原因, 这同时也是网络演化模型所关心的问题。一个演化模型原则上都可以对应于一种链路预测的算法。因此我们就可以借助链路预测的框架和评价方法量化的对不同演化模型所对应的链路预测算法进行评价, 从而间接的对演化模型进行比较和评价。我们最终希望能够在这种方法论的指导下建立一个公平的比较不同演化模型的平台。下面, 我们将以中国城市航空网络为例, 该网络包含 121 个节点 1466 条边, 属于规模较小的网络, 因此使用  $leave-one-out$  的方法进行预测。

## 2. 中国航空网络的链路预测

本文研究的中国城市航空网络以通航城市为节点，两个城市间的直飞航线为边(不包括有经停的航线)，共包含 121 个节点和 1466 条边。统计数据涵盖了国内主要航空公司，包括中国国际航空公司、东方航空公司、南方航空公司、上海航空公司、山东航空公司、四川航空公司、海南航空公司和厦门航空公司，在 2006 年提供的所有航班。对于城市有一个以上机场的将其数据合并，例如，上海有虹桥机场和浦东机场，重庆有江北机场和万州梁平机场。统计数据中只包括行政级别为县级市及以上的城市，但不包括香港和澳门。所有数据来源于网络。中国城市航空网络的平均度为 11.388，簇系数 0.748，平均最短路径长度 2.263，是一个典型的小世界网络，且其度分布符合双段幂律分布<sup>[26]</sup>。

一类链路预测方法是基于网络结构相似性的，即根据网络结构信息定义两节点的相似性，并假设两个更加相似的节点间产生连边的概率更大。这种相似性的定义，可以是最简单的基于共同邻居数量 (Common Neighbors, 简称 CN)<sup>[27]</sup>的方法，也可以是较复杂的基于随机游走过程或者矩阵森林等较复杂的方法。虽然复杂的算法考虑更多的信息，有时候表现不一定有基于局部信息的算法好<sup>[23]</sup>。文献<sup>[20]</sup>中比较了 9 种基于共同邻居的局部相似性算法，结果显示最简单的 CN 表现较好，而且对航空网络的预测比较准确，AUC 可达到 0.9 以上。另外，崔爱香等人提出了一种由共同邻居驱动的网络演化模型，结果显示这个模型不仅能够重现网络的全局性质，如服从幂律的度分布，还可以重现局部的拓部结构性性质，如完全子图的度的幂律分布<sup>[28]</sup>。因此本文采用共同邻居的相似性定义作为基准方法，即两个节点如果有更多的共同邻居，那么它们之间就更有可能会产生连边。对于节点  $x$ ， $\Gamma(x)$  表示  $x$  的邻居集合。因此，共同邻居指标  $s_{xy}$  就是共同邻居集合中元素的个数，即：

$$s_{xy} = |\Gamma(x) \cap \Gamma(y)| \quad (4)$$

在以往的航空网络演化研究中不仅考虑了网络的拓扑结构信息，有些还考虑了几何因素或节点的属性 (或称外部因素)。例如 Guimera 等人同时考虑优先连接和距离因素<sup>[29, 30]</sup>，文献<sup>[31]</sup>比较了人口、距离和经济与网络演化的关系，并基于此三因素分别建立了三个网络演化模型，结果发现基于第三产业产值建立模型所产生的仿真网络能够更好的符合实际的网络结构。本文根据城市的人口、经济 (GDP 和第三产业产值) 以及城市间的距离作为指标定义了四种相似性指标，即：

(i) 基于距离的相似性：城市距离被认为是影响航空网络构建的重要因素，航线数量随着距离的增加将减少，因此我们认为两节点间距离 (Dis) 越远这两个城市的相似性越小，即有航线的概率越小：

$$s_{xy}^{Dis} = \frac{1}{Dis(x, y)}, \quad (5)$$

其中，两个城市间的距离根据城市的经纬度，通过近似地把地球看作一个标准球体，并利用球面距离公式求得。

(ii) 基于人口数量的相似性：假设每个城市的人均飞行次数相同，定义两城市的相似性正比于它们各自的人口数 ( $P$ ) 的乘积

$$s_{xy}^{Popu} = P(x) \cdot P(y), \quad (6)$$

这里，人口指全市年末总人口数。

(iii) 基于城市经济指标 ( $GDP$ ) 的相似性：两城市的相似性正比于他们的  $GDP$  的乘积。

$$s_{xy}^{GDP} = GDP(x) \cdot GDP(y), \quad (7)$$

这里， $GDP$  是国内生产总值，指在一定时期内(一个季度或一年)，一个国家或地区的经济中所生产出的全部最终产品和劳务的价值；

(iv) 基于第三产业产值的相似性：考虑到航空业的主体是与服务业相关的，因此用第三产业产值 (Tertiary Industry, 简称  $TI$ ) 定义相似性：

$$s_{xy}^{TI} = TI(x) \cdot TI(y). \quad (8)$$

由于航空网络的数据为 2006 年的数据，因此，人口、 $GDP$  和第三产业产值将采用各空港城市 2005 年的相关数据 (来源于《中国城市统计年鉴 2005》)。上述五种基于节点相似性的算法：共同邻居、距离、人口、 $GDP$  和第三产业产值的预测准确度如表 1 所示。

表1 五种相似性算法的预测准确度

算法名称	RankS
共同邻居 ( $CN$ )	0.10185
距离 ( $Dis$ )	0.30081
人口 ( $Popu$ )	0.25475
国内生产总值 ( $GDP$ )	0.14588
第三产业产值 ( $TI$ )	0.11954

可以发现，当单独以这五种因素作为相似性算法的基础时， $CN$  指标预测准确度最高。这说明共同邻居的相似性定义能够符合网络的结构特征。在四个基于节点属性的相似性中，基于距离的相似性预测效果最差，而第三产业产值表现最好，这与文献<sup>[31]</sup>中的结论十分吻合，即在中国城市航空网络中，节点的第三产业产值是影响网络结构及演化的重要因素。

事物呈现的状态往往是多种因素共同作用的结果，我们将基于拓扑结构的  $CN$  指标与其他四个基于几何距离或节点属性的指标分别耦合起来进行预测。本文采用最简单的线性方式，即

$$s = \lambda \cdot s^{CN} + (1 - \lambda) \cdot s^{Attr} \quad (6)$$

其中， $s^{CN}$  表示基于结构的方法， $s^{Attr}$  表示基于几何因素和节点属性的方法，参数  $\lambda \in [0, 1]$ 。

当  $\lambda = 1$  时算法回到  $s^{CN}$ ，当  $\lambda = 0$  时算法回到  $s^{Attr}$ 。为了便于计算，将每一种相似性的值归一化到区间[0,1]，即除以其最大值。

耦合算法的精确度如图 1 所示，其中子图 (a)、(b)、(c) 和 (d) 分别展现了四种耦合算法随着参数  $\lambda$  变化预测准确度的变化。由图可见四种算法都有一个最优的参数值  $\lambda^*$ ，此时取得最低的 RankS 值，即预测精度最高。四种参数的最优参数值及其对应的 RankS 值总结与表 2 中。

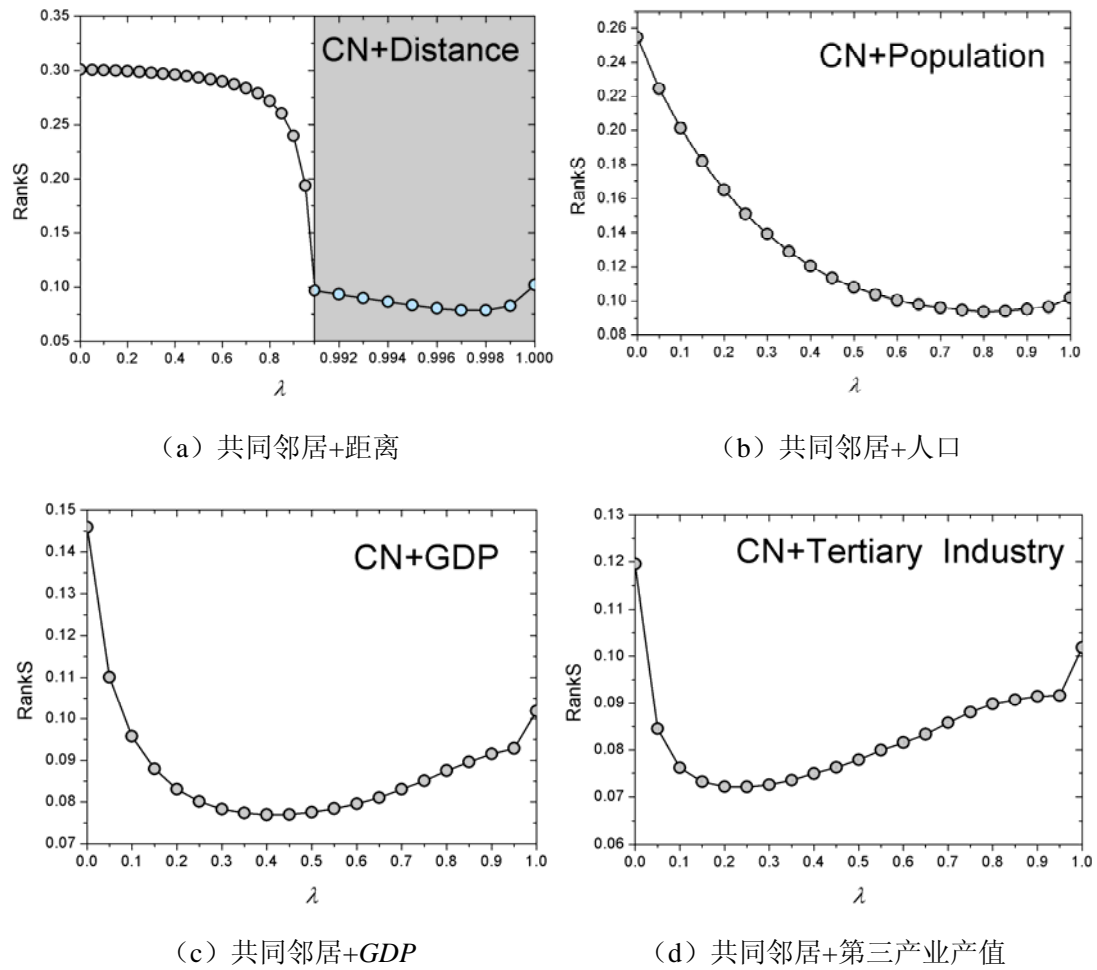


图 1 耦合算法精确度随参数  $\lambda$  变化情况

表 2 耦合算法的预测准确度与比较

耦合算法	$\lambda$ 最优值	RankS	预测提高幅度 (与 CN 比较)	预测提高幅度 (与属性因素比较)
CN+Dis	0.997	0.07844	23%	74%
CN+Pop	0.8	0.09381	8%	63%
CN+GDP	0.4	0.07693	24%	47%
CN+TI	0.2	0.07216	29%	40%

从表 2 可以看出,耦合了节点属性信息的算法,其预测准确性比只考虑网络结构的 CN 算法都有所提高,特别是考虑第三产业产值的耦合算法,相比较 CN 算法, RankS 由原来的 0.10185 降为 0.07216,说明预测精确度提高了 29%。若与只考虑第三产业产值的算法比较,预测精度可提高 40%。比较表 2 中最后两列数字,不难看出,共同邻居数量对连边的产生起了相当重要的作用。此外,在其他四个因素中第三产业产值最重要。虽然距离信息和 CN 的耦合也能够产生不错的结果,但其中距离作用不明显( $\lambda$  趋近于 1),仅仅是起到了进一步区分 CN 值相同的节点对的作用。例如北京首都机场与洛阳机场和敦煌机场的共同邻居都是 5 个,由于北京到洛阳的距离小于北京到敦煌的距离,于是我们认为北京和洛阳之间更可能产生连接。耦合算法的预测结果不仅进一步支持了文献<sup>[31]</sup>中的结论,也更细致地刻画出影响航空网络结构的各种因素贡献的大小。

回到网络演化机制的讨论上来,链路预测的结果实际上暗示,如果只能考虑一种驱动因素的话,那么以共同邻居为驱动力的模型可以得到最佳的结果,而在所有的外部影响因素的比较中,以第三产业为驱动力的模型能够产生最佳的结果,这些结论和最近的建模研究<sup>[28,31]</sup>非常吻合。实际上,在所有的外部因素中,只有以第三产业为驱动力的模型可以再现航空网独特的双段幂律分布<sup>[31]</sup>。

### 3. 结论

尽管目前研究网络演化机制的常用方法是直接应用演化模型来推测影响网络演化的因素,但是由于刻画网络结构特征的统计量非常多,这些统计量的表现往往不一致,很难比较不同因素驱动的机制之间孰优孰劣。而用链路预测方法推测网络演化的机制就规避了传统方法的缺陷。

在本文中,我们介绍了基于节点相似性的链路预测方法,分析利用链路预测推测网络演化的机制。在推测中国城市航空网络演化机制检测链路预测方法准确性的例子中,可以看到,与我们之前通过直接建立网络演化模型分析演化的主导因素<sup>[31]</sup>,得到的结论是一致的。这说明,利用链路预测方法分析网络演化机制是一种有效的途径。更为重要的是,与直接建立网络演化模型相比,由于链路预测能够计算预测方法的准确度,能够清晰直观地利用量化结果对各种因素进行辨别,因此,链路预测在分析网络演化机制上比传统方法更为有效。

除此外,在我们以前的工作中<sup>[32]</sup>,从经济学角度将 GDP 分解为三次产业的产值,利用相关分析和偏相关分析讨论出只有第三产业产值和航空客运量具有直接相关关系,在建立网络演化模型中直接摒除了第一产业和第二产业产值,只利用第三产业产值作为驱动因素,得到较好的拟合效果。本文的工作,尽管没有从经济学角度将 GDP 分解为三次产业的产值,而是直接利用了 GDP 和第三产业产值作为链路预测的指标,但是链路预测的方法依然能辨析出第三产业产值的预测准确性更好。这说明,链路预测的方法拨开了数据的表象,能挖掘到数据更深层次的内涵,对那些缺少经济学基础而又期望从这个思路研究网络演化的学者们来说,无疑提供了一种更加便捷的途径。

可以看到, 链路预测方法具有的优势, 使其有望为分析网络演化机制提供一个简单统一且较为公平的比较平台, 量化比较各种不同机制对于真实生长行为的预测能力, 从而推动复杂网络演化模型的理论研究。

#### 参考文献

- 1 Watts D J, Strogatz S H. Collective dynamics of 'small-world' networks. *Nature*, 1998, 39: 440-442.
- 2 Albert R, Jeong H, Barabási A-L. Diameter of the world wide web. *Nature*, 1999, 401: 130-131.
- 3 Adamic L A, Huberman B A. Power-law distribution of the world wide web. *Science*, 2000, 287: 2115.
- 4 Newman M E J. The structure of scientific collaboration networks. *Proc Natl Acad Sci U.S.A.*, 2001, 98(2): 404-409.
- 5 Albert R, Barabási A-L. Statistical mechanics of complex networks. *Rev Mod Phys*, 2002, 74(1): 47-97.
- 6 Dorogovtsev S N, Mendes J F F. Evolution of networks. *Adv Phys*, 2002, 51: 1079-1187.
- 7 Barabási A-L, Albert R. Emergence of scaling in random networks. *Science*, 1999, 286: 509-512.
- 8 Garlaschelli D, Capocci A, Caldarelli G. Self-organized network evolution coupled to extremal dynamics. *Nat. Phys.*, 2007, 3: 813-817.
- 9 Valverde S, Cancho R F i, Sole R V. Scale-free networks from optimal design. *Europhys. Lett.*, 2002, 60(4): 512-517.
- 10 Kim B J, Trusina A, Minnhagen P, *et al.* Self organized scale-free networks from merging and regeneration. *Eur. Phys. J. B*, 2005, 43: 369-372.
- 11 Perotti J I, Billoni O V, Tamarit F A, *et al.* Emergent self-organized complex network topology out of stability constraints. *Phys. Rev. Lett.* , 2009, 103: 108701.
- 12 Zhou S, Mondragón R J. Accurately modelling the Internet topology. *Phys.Rev. E*, 2004, 70: 066108.
- 13 Carmi S, Havlin S, Kirkpatrick S, *et al.* A model of Internet topology using k-shell decomposition *PNAS*, 2007, 104: 11150-11154.
- 14 Getoor L, Diehl C P. Link mining: A survey. *ACM SIGKDD Explorations Newsletter*, 2005, 7(2): 3-12.
- 15 吕琳媛. 复杂网络链路预测. *电子科技大学学报(自然科学版)*, 2010, 38(5).
- 16 Lü L, Zhou T. Link prediction in complex networks: A survey. to be published.
- 17 Lin D. An Information-Theoretic Definition of Similarity. in *Proceedings of the 15th International Conference on Machine Learning*. July, 1998. Madison, Wisconsin.
- 18 Liben-Nowell D, Kleinberg J. The link prediction problem for social networks. *J. Am. Soc. Inform.Sci. & Technol.* , 2007, 58: 1019.
- 19 Adamic L A, Adar E. Friends and neighbors on the Web. *Social Networks*, 2003, 25(3): 211-230.
- 20 Zhou T, Lü L, Zhang Y-C. Predicting Missing Links via Local Information. *Eur. Phys. J. B*, 2009, 71: 623-630.
- 21 Pan Y, Li D H, Liu J G, *et al.* Detecting community structure in complex networks via node



- similarity. *Physica A*, 2010, 389(14): 2849-2857
- 22 Wang Y L, Zhou T, Shi J J, *et al.* Empirical analysis of dependence between stations in Chinese railway network. *Physica A*, 2009, 388(14): 2949-2955.
- 23 Lü L, Jin C-H, Zhou T. Similarity index based on local paths for link prediction of complex networks. *Phys Rev E*, 2009, 80: 046122.
- 24 Liu W, Lü L. Link prediction based on local random walk. *EPL*, 2010, 89: 58007.
- 25 Hanley J A, McNeil B J. The meaning and use of the area under a receiver operating characteristic(ROC) curve. *Radiology*, 1982, 143: 29-36.
- 26 刘宏鲲, 周涛. 中国城市航空网络的实证研究与分析. *物理学报*, 2007, 56(1): 106-112.
- 27 Lorrain F, White H C. Structural equivalence of individual in social networks. *J. Math. Sociol.*, 1971, 1: 49-80.
- 28 崔爱香, 傅彦, 尚明生, *et al.* 复杂网络局部结构的涌现: 共同邻居驱动网络演化. *物理学报*, 已接收.
- 29 Guimerà R, Amaral L A N. Modeling the world-wide airport network. *Eur. Phys. J. B*, 2004, 38: 381-385.
- 30 Guimerà R, Mossa S, Turttschi A, *et al.* The world-wide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proc. Natl. Acad. Sci. USA*, 2005, 102(22): 7794-7799.
- 31 刘宏鲲, 张效莉, 曹崑, *et al.* 中国城市航空网络航线连接机制分析. *中国科学 G 辑*, 2009, 39(7): 935-942.
- 32 Liu H K, Zhang X L, Zhou T. Structure and External Factors of Chinese City Airline Network, *Physics Procedia*, 2010, 3: 1781-1789.