

# 开放式复杂航空网络系统的动力学演化\*

钱江海<sup>1)</sup> 韩定定<sup>2)†</sup> 马余刚<sup>1)</sup>

1) (中国科学院上海应用物理研究所, 上海 201800)

2) (华东师范大学信息科学技术学院, 上海 200241)

(2010年11月7日收到; 2010年12月16日收到修改稿)

本文首先实证研究了中国航空网(CAN)与外部经济环境,即国民生产总值(GDP)的相关性,进而揭示CAN的演化方式及其拓扑特征的起源.发现自1998年起网络节点的度与其GDP成线性关系,表明了网络拓扑与经济因素有密切的联系,且CAN的度演化服从类似于多重过程(multiplicative process)的模式,而该模式又强烈依赖于GDP的增长率.这种独特的模式暗示了CAN具有类似适应度模型的微观机理.通过对演化方程的研究发现:1)CAN在增长过程中,节点加入的时刻具有经济上的偏好性,即节点加入的时刻与该节点的初始GDP的对数成负线性关系;2)特别地,发现该演化模式耦合指数增长的节点可以产生双段幂率的度分布.将模拟仿真结果与实际的CAN度分布比较,发现两者基本符合.这表明在GDP增长率驱动下的类似多重过程的动力学方程真实地刻画了CAN的度演化,并且是导致网络出现双段幂率度分布的原因.

**关键词:** 国民生产总值(GDP), 中国航空网(CAN), 双段幂率分布, 多重过程

**PACS:** 89.75.Hc, 89.75.Da, 89.40.Dd

## 1. 引言

近年来,航空网络作为当今最重要的交通工具、同时也作为一个典型的复杂开放式系统,受到了科学家的广泛关注<sup>[1,2]</sup>.其中一个重要的发现就是这类网络是无尺度的,即它们的度分布大部分为幂率分布<sup>[3]</sup>,例如印度航空网,欧洲航空公司网络以及全球航空网<sup>[4-6]</sup>.无尺度的特征同样存在于其他各类网络中<sup>[7-12]</sup>.为了解释这新奇的特征,Barabasi提出了著名的BA模型<sup>[13]</sup>.该模型通过引入关键的“偏好依附”机理解释了实际网络中幂率度分布的存在<sup>[13,14]</sup>.在此之后,许多学者提出各种网络演化模型以更准确细致地再现网络的幂率分布,例如适应度模型和加速网络模型<sup>[15-20]</sup>.除此以外,幂率网络中的动力学机理研究也吸引诸多学者的关注<sup>[21,22]</sup>.然而,并不是所有的网络都遵循严格的幂率度分布.研究表明,中国航空网和美国航空网的度都服从所谓的“双段”幂率分布,而该特征无法由先前提出的各类模型解释<sup>[23]</sup>.这意味着在这类网络的演化中,尚存在一些值得探究的内在机理.

此外,先前的大部分网络研究并没有真正把网络理解成一个开放式复杂系统,环境对网络结构和功能的影响的研究并不多见,比如在航空网络的研究中,没有考虑经济因素与网络之间的相关性.然而,作为一个典型的开放式系统,这种影响往往是不可忽略的.有研究表明在2004年,航空业承担了全球40%的国际旅游和40%的货物运输<sup>[24]</sup>.这显然对于地区的经济有着重要影响.而经济的增长又会刺激该地区航空运输的需求.客流和经济之间的正相关性已被证实,还有学者引入引力模型来估计航空流量<sup>[25]</sup>.然而这些研究往往是定性的.而定量描述这种相关性会对我们理解网络内在的动力学演化大有帮助.

基于以上考虑,我们侧重从经济因素的角度研究中国航空网的动力学演化机理.我们发现CAN和城市GDP都是指数增长的.特别地,从1998年起,节点(城市)的度正比于其GDP.基于这个宏观观测我们得到了度的演化方程.该方程服从类似于多重过程的演化机理.我们发现该机理可以初步解释网络中出现的双段幂率度分布现象.本文对实测数据进行实证分析,主要关注以下三方面,如网络

\* 国家自然科学基金(批准号:11075057,10775167,10979074)资助的课题.

† 通讯联系人. E-mail: ddhan@ee.ecnu.edu.cn

的增长模式、网络的度分布、以及度与 GDP 的相关性;推导度的动力学演化方程并作相应的分析. 我们发现,航空网中,节点的加入存在一种特殊的偏好性. 详细分析了由动力学方程得出的度分布,并证明类似多重过程的演化机理,在耦合了节点的指数增长的情况下可以导致双段幂率分布. 最后将以上结论与实际的航空网度分布作比较.

## 2. 实证分析

中国航空系统可被抽象为一个复杂网络,其中城市代表节点,城市间航班代表边. 为了简化,我们仅考虑无向图,定义节点的度为连接到该节点的总边数. 对于每个城市节点,其自身的 GDP 可以看作是它所处的不断变化的环境. 我们的数据主要包括两个部分:经济信息和网络的度序列. 前者来自于中国城市统计年鉴,收集了每个空港城市从 1988 年到 2007 年的 GDP(1995 和 1996 两年除外). 而度序列是通过收集中国交通年鉴的航空信息并数据挖掘得到,包括 1988 年—2003 年以及 2008 年的数据.

首先,我们研究了 CAN 的总节点数  $N(t)$  和总边数  $M(t)$  的演化,两者都是随时间指数增长的,

$$N(t) \propto e^{c_n t}, \quad (1)$$

$$M(t) \propto e^{c_m t}, \quad (2)$$

其中  $c_n \approx 0.088, c_m \approx 0.154$  分别是  $N(t)$  和  $M(t)$  的指数. 尽管以上两个指数看起来很小,但这种较小的指数并不是中国航空网络独有的. 例如,美国航空网络的客流增长,显示为指数增长,而其相应的指数仅为 0.007<sup>[26]</sup>;文献[27]作者研究了英特网节点和边的增长,确定其为指数增长,而指数分别为 0.029 和 0.034. 这些指数都非常的小,这可能与网络所处的环境和自身特性有关. 我们以上的两个指数满足  $c_n < c_m < 2c_n$ . 由于  $M(t) < \binom{N(t)}{0} \propto e^{2c_n t}$ ,  $c_m < 2c_n$  保证了网络不会变成全连的. 而不等式  $c_n < c_m$  表明了 CAN 是一个加速网络. 先前的加速网络模型通常假设节点随时间线性增加,而边的增长则与时间成某个非线性幂率关系<sup>[16-18]</sup>, CAN 的演化模式显然与该假设完全不同. 在第 4 节我们将看到,节点的指数增长是产生双段幂率的重要因素之一. 值得提出的是,网络的加速增长通常有两个因素引起:非线性增长的内部连边或者是与时间

有关的初始度. 我们分析了从 1988 到 2003 年 CAN 每年新加入节点的初始度,发现 79% 的新节点初始度为 1,最大的初始度也只有 8,且初始度值的分布与时间没有明显相关性. 因此,网络的加速增长现象是由内部非线性增长的连边引起.

对于 CAN 的拓扑结构,我们主要关注其度分布特征. 图 1 画出了 1999 年到 2003 年以及 2008 年的累计度分布. 它们展现出双段幂率特征. 我们发现该特征首次出现在 1999 年. 其第一段指数  $\gamma_1$  大约稳定在 -0.51 而第二段的指数  $\gamma_2$  大约为 -2.7. 而折断点  $k_c$  随时间逐渐增长. 1999 年到 2002 年,  $k_c$  从 18 逐渐增长到 22,尽管这点在图 1 中并不明显,但是到了 2008 年,  $k_c$  的值明显的增加到了 30. 独特的双段幂率的特征与航空网络的内在演化机理密不可分,虽然先前的学者也做了不少实证研究,但该特征的起源依然不清楚.

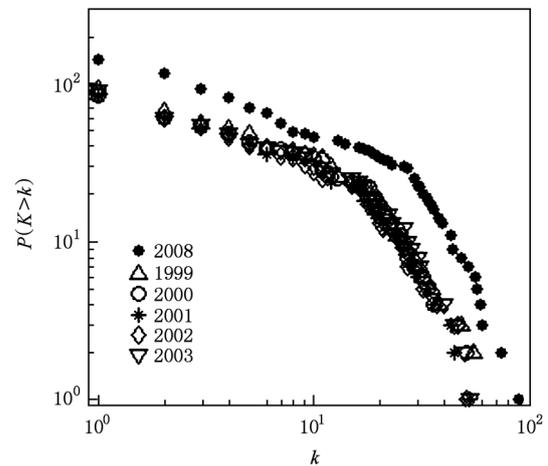


图 1 1999 年至 2003 年和 2008 年中国航空网的累计度分布 (其呈现双段幂率分布. 两段指数的斜率相对稳定,但是折断点却随着时间增加)

CAN 的演化并不是孤立的,其中经济因素对航空网络有着重要影响. 然而这方面的研究在国内外仅限于定性描述,以下我们将找出 CAN 拓扑与其 GDP 定量的联系,并以此为基础逐步揭示开放式 CAN 的复杂动力学演化机理.

通过考察  $k_i$  与  $\langle G_i \rangle_k$  的关系,可以得到 GDP 与度的相关性,其中  $\langle \rangle_k$  表示对所有  $k$  求平均. 图 2 给出了结果,我们发现节点度  $k_i$  正比于其  $\langle G_i \rangle_k$ , 即节点度随 GDP 增长而增长,但是简单的线性关系却出乎了我们的预料. 这个关系首次出现在 1998 年,仅仅早于双段幂率所出现一年并且在此后的研究时间段中(即到 2003 年)这种关系一直存在. 这暗示

着双段幂率的涌现与该线性关系可能有着某种内在联系. 如果我们考虑时间演化, 这种相关性可以表示为

$$k_i(t) = D(t) \langle G_i(t) \rangle_k, \quad (3)$$

这里  $D(t)$  是一个与时间相关的系数. (3) 式不仅表明网络的演化和经济因素密切相关, 还刻画了它们是如何相关, 即定量地描述了网络拓扑的演化与时变的外部经济环境的关系. 这种定量的相关性提供了一种新的方式来探测网络的实际演化. 由于航空网络通常是尺度较小, 对其演化过程的直接测量往往带有很大的涨落. 然而 GDP 通常被很好地记录着并且, 如我们后面会看到的, 它通常有着较明确的演化模式. 如果每个城市的  $G_i(t)$  可以被确定, 通过 (3) 式便能确定 CAN 的演化方式.

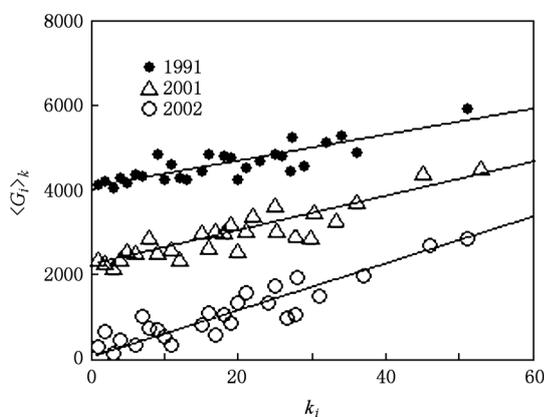


图2 1999年、2001年及2002年节点度和GDP的线性相关性, 并从1998到2003年一直保持这种关系(为了使不同年份的数据看起来清楚, 分别对2001年和1999年的数据向上提升了2000和4000. 线性斜率随着时间的增加而发生变化)

为了研究节点度的演化, 我们有必要先研究GDP的增长. 我们发现90%的空港城市的GDP服从下式的指数增长(图3):

$$G_i(t) = G_i(0)e^{\lambda_i t}, \quad (4)$$

这里  $G_i(t)$  是城市  $i$  在时间  $t$  的GDP,  $\lambda_i$  是其增长率, 它服从均值  $\mu_\lambda = 0.18$ 、标准差  $\sigma_\lambda = 0.02$  的正态分布(图4). 特别的, 总GDP  $G_{\text{total}}(t)$  也服从指数增长, 增长率  $\lambda_{\text{total}} = 0.19$ .

### 3. CAN 的动力学演化方程及其分析

首先, 我们来估计  $D(t)$  的形式. 令(3)式的等号两边对所有节点求和, 我们得到  $D(t) = \frac{2M(t)}{G_{\text{total}}(t)}$

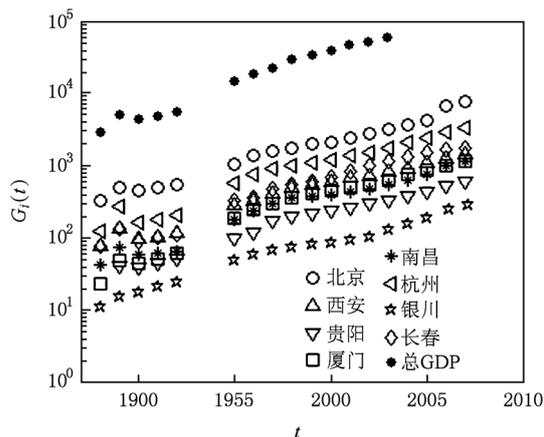


图3 GDP的增长: 各个城市的GDP和所有城市的总GDP都是指数增长的(作为例子, 图中列举了8个城市的GDP演化过程)

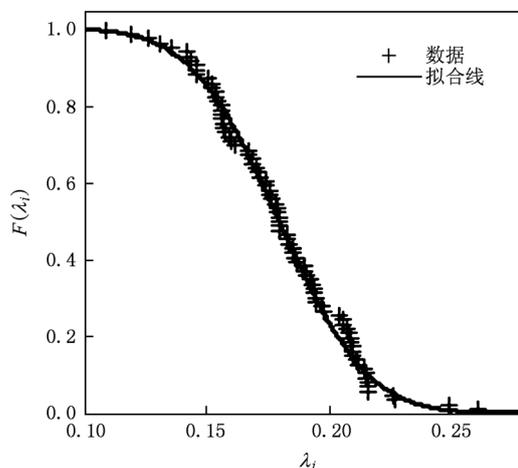


图4 GDP增长率的累计分布(实线表示拟合函数  $F(\lambda_i) = 0.5 \operatorname{erfc}\left(\frac{\lambda_i - 0.18}{0.03}\right)$ , 因此GDP增长率  $\lambda_i$  服从均值为0.18, 标准偏差约为0.02的正态分布)

$\propto e^{-d}$ , 这里  $d = \lambda_{\text{total}} - c_m = 0.036$ . 统计上来看, (3)式可写成  $k_i(t) = G_i(0)e^{(\lambda_i - d)t} = G_i(0)e^{\eta_i t}$ , 其中  $\eta_i = \lambda_i - d$  服从均值  $\mu_\eta = 0.144$ , 标准偏差  $\sigma_\eta = 0.02$  的正态分布. 现在我们可以得出节点度的动力学方程

$$\frac{dk_i}{dt} = \eta_i k_i, \quad (5)$$

(5)式与经典多重过程非常相似, 其离散形式可表示为  $k_i(t+1) = k_i(t)[1 + \varepsilon_i(t+1)]$ ,  $\varepsilon_i(t+1)$  是白噪声<sup>[24]</sup>. 但与经典的多重过程中  $\varepsilon_i(t+1)$  不同的是, (5)式中的  $\eta_i$  是一个不随时间变化的, 并且是正态分布的随机变量. 另一方面, 不随时间变化的  $\eta_i$  的存在

使得(5)式非常类似于适应度模型的演化方程. 等式右边时间因子的缺失是由于节点的增长是指数的而非线性的. 因此该演化方程暗示着 CAN 演化中存在类似适应度模型的微观机理且实际的适应度与 GDP 增长率有关而不是人口和 GDP 本身.

利用初始条件  $k_i(t, t')|_{t=t'} = 1$  ( $k_i(t, t')$  表示  $t'$  时刻加入的节点在  $t$  时刻的度值), 我们很容易得到  $k_i(t, t')$  的表达式

$$k_i(t, t') = e^{\eta_i(t-t')}, \quad (6)$$

它随时间指数增长的形式与 CAN 自身是指数增长相符合. 比较(3)式与(6)式, 可得到  $G_i(0)$  应为

$$G_i(0) = e^{-\eta_i t'}, \quad (7)$$

(7)式揭示了一个有趣的结果: 节点加入的时间随其  $G_i(0)$  成对数衰减的关系. 这种递减的关系是可理解的, 因为一个发达的城市会比其他城市可能更早地建设航空设施. 为了进一步检验(7)式, 我们计算了每个时刻  $t'$  的  $\langle G_i(0) \rangle_{t'}$ , 这里  $G_i(0)$  被假设为 1988 年的 GDP. 如图 5 所示,  $\langle G_i(0) \rangle_{t'}$  确实随  $t'$  成指数衰减, 其衰减指数为 0.155, 与  $\mu_\eta = 0.144$  相近. 这一结果不仅揭示了 CAN 中节点加入的偏好性, 即网络节点的选择也是受经济因素所驱动的, 同时也为(5)式提供了间接的证据.

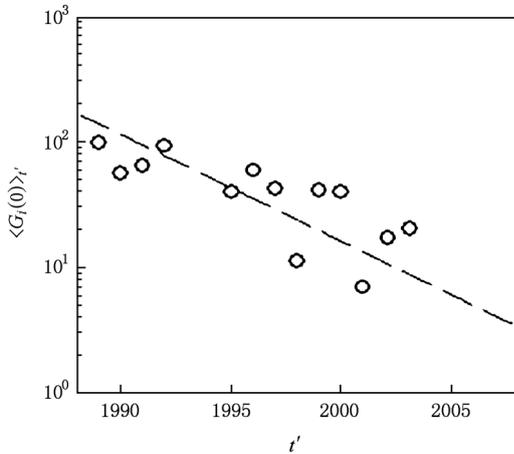


图 5  $\langle G_i(0) \rangle_{t'}$  与  $t'$  的相关性 ( $\langle G_i(0) \rangle_{t'}$  随  $t'$  指数衰减, 其指数大约为 0.155, 接近于  $\mu_\eta = 0.144$ )

#### 4. CAN 的双段幂率度分布

CAN 的度分布, 正如我们第 2 节中讨论的, 服从双段幂率分布, 但先前的网络模型尚没能再现这种分布. 如果(5)式正确描述了实际 CAN 的演化过程, 它应该能够再现这一新奇的特征. 正如大家接

下来将看到的, 类似于多重过程的(5)式耦合节点的指数增长模式会导致一种指数混合的对数正态分布, 而这种分布就是双段幂率分布.

利用公式  $p(k_i(t, t')) dk_i(t, t') = f(\eta_i) d\eta_i$  可以简单得到  $k_i(t, t')$  服从对数正态分布. 由于节点是指数增长的, 则其生存时间  $T$  (定义为  $T = t - t'$ ) 服从指数分布. 因此我们实际观察到的度分布是指数分布  $T$  混合之后的对数正态分布  $k_i(T)$ . 因此实际度分布应表达为

$$p(k) = \int_0^{t_c} \frac{1}{\sqrt{2\pi\sigma_\eta} k T} e^{-\frac{(\log(k) - \mu_\eta T)^2}{2\sigma_\eta^2 T^2}} e^{-c_n T} dT, \quad (8)$$

其中  $t_c$  是我们观察  $p(k)$  的时间. 如果  $\sigma_\eta \rightarrow 0$  则

$$\frac{1}{\sqrt{2\pi\sigma_\eta} T} e^{-\frac{(\log(k) - \mu_\eta T)^2}{2\sigma_\eta^2 T^2}} \rightarrow \delta\left(T - \frac{\log(k)}{\mu_\eta}\right),$$

那么(8)式为

$$p(k) \propto k^{-\left(1 + \frac{c_n}{\mu_\eta}\right)} \zeta(k - e^{\mu_\eta t_c}), \quad (9)$$

其中

$$\zeta(k) = \begin{cases} 1, & k \leq 0, \\ 0, & k > 0. \end{cases}$$

节点度服从带有截断的幂率分布. 截断发生在最大度值处  $k_c = e^{\mu_\eta t_c}$ .

对于有限的  $\sigma_\eta > 0$ , 难以从(8)式导出解析结果. 于是我们通过模拟仿真来研究该问题. 仿真是通过如下步骤进行: 在每个时间步中, 节点数是指数增长, 新加入的节点具有初始度 1 并且被分配一个适应度  $\eta_i$ ,  $\eta_i$  选自于某个正态分布; 然后每个节点按照(6)式增加自己的度值. 我们针对不同的  $\sigma_\eta$  模拟了累计度分布, 如图(6). 当  $\sigma_\eta = 0$  时, 累计度分布服从幂率分布并在  $e^{\mu_\eta t_c}$  处发生截断, 与我们先前的分析完全相符. 随着  $\sigma_\eta$  的增加, 曲线的第一部分没有明显的变化, 它与  $\sigma_\eta = 0$  时的幂率曲线基本重合, 即第一部分曲线为(9)式描述的幂率分布; 而曲线分布的第二段随着  $\sigma_\eta$  的增加衰减得越来越慢并且它们的形状在双对数坐标下都是一条直线, 即分布开始展现双段幂律的特征, 且第二段的斜率随  $\sigma_\eta$  的增加而减小. 值得指出的是, 所有的折断点都发生在  $k_c = e^{\mu_\eta t_c}$  附近, 而此处正是  $\sigma_\eta = 0$  时截断发生的地方. 这意味着折断点  $k_c$  可以被  $e^{\mu_\eta t_c}$  很好的估计. 因此, 折断点随着演化时间  $t_c$  的增加而增加, 这与 CAN 的实证数据完全相符. 为了验证(5)式刻画的演化过程确实是网络形成双段幂率的起源, 我们将实际 CAN 的数据代入模拟过程 ( $c_n = 0.088, \mu_\eta = 0.16, \sigma_\eta = 0.01$ ), 并测量了其两段的幂律指数.

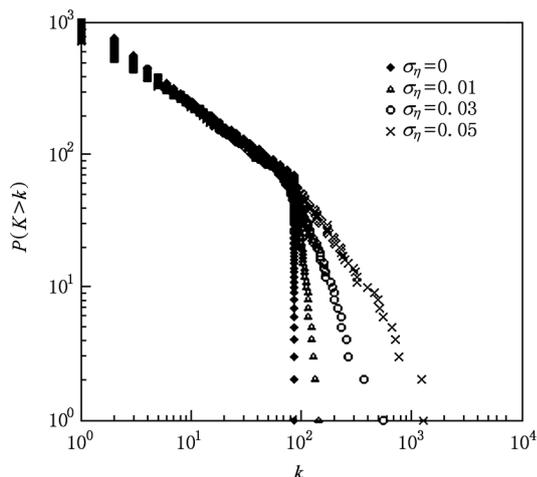


图 6 累计度分布的模拟(该模拟是在  $\mu_\eta = 0.16, c_n = 0.088$  下得到的. 对于其他的  $\mu_\eta, c_n$  有相似的结论)

我们得到  $\gamma_1 = 0.61$ , 与实际的 0.51 接近;  $\gamma_2 = 2.8$ , 与实际测量 2.7 也能很好地符合.

### 5. 讨 论

通过国民生产总值 GDP 与中国航空网 CAN 的相关性实证研究,我们发现 CAN 的演化过程是指数增长的,节点度从 1998 年起展现出与 GDP 的线性关系. 基于这个观测,我们推导出了网络度值的演化方程,该动力学方程暗示了一个类似多重过程的演化方式以及适应度模型的微观机理,而节点的加入具有一种特殊的经济上的偏好性. 进一步地,本文阐述了节点度动力学方程中耦合指数增长的节点可以产生双段幂率分布,且分布的折断点随着演化时间增长而增长,实证数据与模拟仿真结果相符合.

[1] Newman M E J 2003 *SIAM Rev.* **45** 167  
 [2] Albert R, Barabási A L 2002 *Rev. Mod. Phys.* **74** 47  
 [3] Han D D, Liu J G, Ma Y G, Cai X Z, Shen W Q 2004 *Chin. Phys. Lett.* **21** 1855  
 Han D D, Liu J G, Ma Y G 2008 *Chin. Phys. Lett.* **25** 765  
 [4] Bagler G 2008 *Phys. A* **387** 2972  
 [5] Han D D, Qian J H, Liu J G 2009 *Phys. A* **388** 71  
 [6] Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A 2004 *Proc. Natl. Acad. Sci. USA* **101** 3747  
 [7] Vazquez A, Pastor-Satorras R, Vespignani A 2002 *Phys. Rev. E* **65** 066130  
 [8] Broder A, Kumar R, Maghoul F, Raghavan P, Rajagopalan S, Stata R, Tomkins A, Wiener J 2000 *Comput. Netw.* **33** 309  
 [9] Redner S 1998 *Eur. Phys. J. B* **4** 131  
 [10] Jeong H, Tombor B, Albert R, Oltvai Z N, Barabasi A L 2000 *Nature* **407** 651  
 [11] Chang H, Su B B, Zhou Y P, He D R 2007 *Phys. A* **383** 687  
 [12] Gao L F, Shi J J, Guan S 2010 *Chin. Phys. B* **19** 010512  
 [13] Barabási A L, Albert R 1999 *Science* **286** 509  
 Barabási A L, Albert R, Jeong H 1999 *Phys. A* **272** 173  
 [14] Jeong H, Néda Z, Barabási A L 2003 *Europhys. Lett.* **61** 567  
 [15] Bianconi G, Barabási A L 2001 *Europhys. Lett.* **54** 436  
 [16] Dorogovtsev S N, Mendes J F F 2006 *Phys. Rev. E* **74** 016117  
 [17] Pan Z F, Wang X F 2006 *Acta Phys. Sin.* **55** 4058 (in Chinese)  
 [潘灶烽,汪小帆 2006 物理学报 **55** 4058]  
 [18] Gagen M J, Mattick J S 2005 *Phys. Rev. E* **72** 016123  
 [19] Zhao H, Gao Z Y 2006 *Chin. Phys. Lett.* **23** 2311  
 [20] Li J, Wang B H, Jiang P Q, Zhou T, Wang W X 2006 *Acta Phys. Sin.* **55** 4051 (in Chinese)  
 [李季,汪秉宏,蒋品群,周涛,王文旭 2006 物理学报 **55** 4051]  
 [21] Qu Z H, Wang P, Song C M, Qin Z G 2010 *Chin. Phys. B* **19** 110504  
 [22] Liu F, Zhao H, Li M, Ren F Y, Zhu Y B 2010 *Chin. Phys. B* **19** 040513  
 [23] Li W, Cai X 2004 *Phys. Rev. E* **69** 04610  
 Chi L P, Wang R, Su H, Xu X P, Zhao J S, Li W, Cai X 2003 *Chin. Phys. Lett.* **20** 1393  
 [24] Ishutkina M A, Hansman R J [http://dspace.mit.edu/bitstream/handle/1721.1/44957/Ishutkina\\_ICAT-2009-2.pdf?sequence=1](http://dspace.mit.edu/bitstream/handle/1721.1/44957/Ishutkina_ICAT-2009-2.pdf?sequence=1) 2010  
 [25] Qian J H, Han D D 2009 *Acta Phys. Sin.* **58** 3028 (in Chinese)  
 [钱江海,韩定定 2009 物理学报 **58** 3028]  
 Qian J H, Han D D 2009 *Phys. A* **388** 4248  
 [26] Gautreau A, Barrat A, Barthélemy M 2009 *Proc. Natl. Acad. Sci. USA* **106** 8847  
 [27] Goh K I, Kahng B, Kim D 2002 *Phys. Rev. Lett.* **88** 108701

## Dynamical evolution of complex airline system \*

Qian Jiang-Hai<sup>1)</sup> Han Ding-Ding<sup>2)†</sup> Ma Yu-Gang<sup>1)</sup>

1) (*Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai 201800, China*)

2) (*School of Information Science and Technology, East China Normal University, Shanghai 200241, China*)

(Received 7 November 2010; revised manuscript received 16 December 2010)

### Abstract

As a typical open complex system, airline network usually interplays with the economic environment during its self-organized evolution. In this paper, we investigate empirically the correlation between gross domestic product (GDP) and the Chinese airline network (CAN). It is found that the degree of the node has formed a linear correlation with its corresponding GDP since 1998 while a year later the CAN began to exhibit double-power-law degree distribution. Based on the observation, the dynamical equation of CAN is derived. A multiplicative process-like evolution pattern, which strongly depends on the GDP growth rate, is obtained. We demonstrate that such a process coupled with the exponential increase of nodes can yield a double-power-law distribution. The simulation is applied to CAN and we find that the simulation result is in agreement with real CAN.

**Keywords:** gross domestic product (GDP), Chinese airline network (CAN), double power-law distribution, multiplicative process

**PACS:** 89.75.Hc, 89.75.Da, 89.40.Dd

---

\* Project supported by the National Natural Science Foundation of China (Grant Nos. 11075057, 10775167, 10979074).

† Corresponding author. E-mail: ddhan@ee.ecnu.edu.cn