arXiv:2302.14596v1 [physics.soc-ph] 28 Feb 2023

交通系统中准稳态之间的转变：以科隆环城高速公路为例

Shanshan Wang *、Michael Schreckenberg 和 Thomas Guhr

德国杜伊斯堡杜伊斯堡-埃森大学物理学院

2023年3月1日

摘要。 交通系统可以以不同的模式运行。在此前的一项研究中，我们将这些模式识别为相关结构中的不同准稳态。本文分析这些准稳态之间的转变，即系统如何改变其运行模式。从长远来看，这或许有助于预测交通中相关模式随时间的演变。我们以科隆环城高速公路为例，为2015年的每个季度构建一个状态转移网络，并发现交通系统中的这些准稳态具有季节性依赖。我们使用PageRank算法，识别并考察2015年在60天滑动时间窗口内频繁出现的主导状态。据我们所知，这是首次针对交通系统开展此类研究。

关键词：准稳态、交通系统、转移概率矩阵、PageRank、降秩相关矩阵

目录

1 I引言 1

2 识别准平稳 状态. 2
2.1 降秩相关矩阵.2 2.2 准平稳状态.3

3 转移 准平稳状态之间. 5
3.1 PageRank算法.5 3.2 准平稳状态的转移概率.6 3.3 准平稳系统中的主导状态.6 3.4 随机事件对交通系统的影响.9 3.5 准平稳状态的重现.10

4 结论 .11

1 引言

交通系统中的流量和速度时间序列呈现出表明并反映非平稳性的波动。尽管如此，交通系统会形成不同的运行模式，并在

*shanshan.wang@uni-due.de

一定时间跨度内存在。换言之，存在准稳态。最近，我们在相关结构中识别出了这些状态 [1], ，并借鉴了金融市场研究中的方法 [2–11]。为了理解系统动力学，这些状态之间的转移尤其值得关注。在金融领域，这一问题已在参考文献中得到研究 [2–11]。非平稳性和准平稳性也通过不同的观测量得到研究，例如气候领域 [12, 13], 电力系统 [14], 语音识别 [15, 16] 以及其他系统。

大多数关于交通系统的研究都侧重于从宏观或微观视角进行建模与仿真 [17–24]。由于可获得的交通数据数量有限，实证研究 [25–28] 较为少见。在此前的一项研究中 [1], ，我们根据交通流的时间相关结构，将德国科隆环城高速公路的交通系统划分为六种准稳态。六种状态如何随时间演变，以及哪种状态在交通系统中占据主导地位，仍有待研究。取得进展将有助于理解交通行为的演变并进行预测。我们旨在研究准稳态之间的转移，并以科隆环城高速公路为例。我们计算时间窗口内各状态之间的转移概率。结合转移概率矩阵与PageRank算法 [29–31]，我们识别交通系统中的主导状态，并分析随机事件对主导状态的影响。我们还探讨系统通常需要经过多少转移步数才能返回某一状态。

PageRank算法最初是为谷歌搜索引擎开发的，旨在根据网页的链接结构确定哪些网页较为重要 [29–31]。网页的链接结构可以用转移矩阵进行数学表示，而网页则被建模为由该转移矩阵构建的图中的节点。为确保由此得到的重要性分数（称为 PageRank 分数）存在且唯一，需要引入一种外部影响来衡量每个节点的重要性 [32]。凭借其简洁性、通用性、结果存在性与唯一性的保证，以及计算速度快等优点 [33], ，PageRank已被应用于许多领域，包括生物学 [34]、化学 [35], 、神经科学 [36], 、数学 [37], 、经济学 [38–40], 等。具体而言，基于道路网络和城市空间网络，PageRank还被用于预测单条道路的交通流 [41]，以及研究用于道路规划和最优路径规划的马尔可夫链模型 [42]。在本研究中，我们利用PageRank算法评估各准稳态的重要性，以识别所研究交通系统中的主导状态。此外，我们还利用PageRank算法探究随机事件对主导状态的影响。

本文安排如下。在第2节中，我们通过对降秩相关矩阵进行k-均值聚类来识别准稳态。在第3节中，我们介绍 PageRank算法，基于转移概率矩阵将 2015年四个季度的状态转移网络可视化，利用PageRank算法识别主导状态，探讨随机事件对主导状态的影响，并分析准平稳状态的重现。最后，我们在第4节讨论结果。

2 识别准平稳状态

我们开发了相对于主导集体运动分析相关结构的技术，并将其应用于多项研究中 [1, 8, 43–45]。相关应用见参考文献 [2–7,9–11, 46]。然而，在所引用的研究中，研究对象是时间序列的相关性，或者等价地说，是测量时间序列时的位置相关矩阵。本文研究的是位置序列的相关性，即相应的相关矩阵是时间的相关矩阵。为清晰起见，也为方便读者，我们在第2.1节简要概述降秩相关矩阵的构造，当然采用适用于时间相关性的形式。在第2.2节中，我们通过对降秩相关矩阵进行 k-均值聚类，识别科隆环城高速公路交通系统中的准稳态。

2.1 降秩相关矩阵

本研究所使用的交通数据来自德国 $K = 35$ 科隆环城高速公路上可用路段的感应线圈检测器，这些路段属于A1、A3和A4高速公路的一部分。

数据包括每条车道和每个高速公路路段的交通流量和速度。我们在15分钟的时间段内，对一个高速公路路段的多条车道上的交通流量进行聚合。对于35个路段，我们获得了35个时间序列 $F_k(t)$、$k=1,\cdots,K$，每个序列每天的长度为$T=96$ 个数据点，表示聚合交通流量。所有 $F_k(t)$ 都输入 $K\times T$ 数据矩阵 F。重要的是，我们并不关注时间序列 $F_k(t)$ $t=1,\cdots,T$中编码的空间相关性。相反，我们关注通过对位置序列 $F_k(t)$ $k=1,\cdots,K$进行相关性分析而得到的相互时间依赖关系。 F的每一列都是一个位置序列，描述不同位置 $k=1,\cdots,K$在同一时间 t的交通流量。我们通过将每个位置序列的均值归一化为零来

$$A_k(t)=F_k(t)-\langle F_k(t)\rangle_K\ , \tag{1}$$

以获得新的 $K\times T$数据矩阵 A，其中 $\langle\cdots\rangle_K$表示对所有位置 k求平均。归一化数据矩阵 A 通过以下方式得到位置序列的 $T\times T$ 协方差矩阵

$$\Sigma=\frac{1}{K}A^\dagger A\ . \tag{2}$$

下标 † 表示矩阵的转置。我们对协方差矩阵进行谱分解

$$\Sigma=\sum_{t=1}^{T}\Theta_t V_t V_t^\dagger\ , \tag{3}$$

其中 Θ_t 是 t的第 - 个特征值 Σ，而 V_t 是相应的特征向量。特征值按升序排列，使得 Θ_1 是最小的，而 Θ_T 是最大的。公式（3）通过将 t 的范围从 a 设为 b，可以去除某些特征值，

$$\tilde{\Sigma}=\sum_{t=a}^{b}\Theta_t V_t V_t^\dagger\ . \tag{4}$$

总共有 $T=96$ 个特征值 Θ_t，对应于 Σ。由于 $K<T$， Σ具有 $T-K+1=62$ 个零特征值和 $K-1=34$ 个非零特征值 [43]。按照特征值升序排列，这意味着前62个特征值为零，即 Θ_1、$\cdots$、$\Theta_{62}=0$，而从 Θ_{63}到 Θ_{96}的特征值均为非零。为去除时间上的集体效应，我们排除最大的特征值 Θ_{96}。此外，为减少噪声，我们还从 Σ中去除最小的9个非零特征值，使得 Σ 中保留的非零特征值范围为 Θ_{72} 至 Θ_{95}。因此，我们将 $a=72$ 和 $b=95$ 设为相应值，以获得降秩协方差矩阵$\tilde{\Sigma}$，其定义良好性已在参考文献 [1,8]中得到证明。最后，我们通过以下方式获得降秩相关矩阵

$$\tilde{D}=\tilde{\sigma}^{-1}\tilde{\Sigma}\tilde{\sigma}^{-1}\ , \tag{5}$$

其中 $\tilde{\sigma}$ 为标准差构成的对角矩阵

$$\tilde{\sigma}=\mathrm{diag}\left(\tilde{\sigma}_1,\cdots,\tilde{\sigma}_T\right)\ , \tag{6}$$

$\tilde{\Sigma}$由降秩协方差矩阵 $\tilde{\Sigma}$计算得出。对于每一天，我们都会计算一个降秩相关矩阵 $\tilde{D}$。对于 2015年 数据完整的 362 天，我们总共得到 362 个矩阵 $\tilde{D}$ 。

2.2 准平稳状态

设 $\tilde{D}(i)$ 为第 i 天的降秩相关矩阵， $\tilde{D}_{t_k t_l}(i)$ 为 $\tilde{D}(i)$ 在时刻t_k 和 t_l的元素。两天相关结构之间的相似性度量定义为

$$\eta_{ij}=\left\langle|\tilde{D}_{t_k t_l}(i)-\tilde{D}_{t_k t_l}(j)|\right\rangle_{t_k t_l}\ , \tag{7}$$

其中， $|\cdots|$表示绝对值， $\langle\cdots\rangle_{t_k t_l}$表示所有矩阵元素的平均值。数值$\eta_{ij}$ 生成一个 $N\times N$ 相似性矩阵 η，其中 $N=362$ ，因为2015年有362天可用。

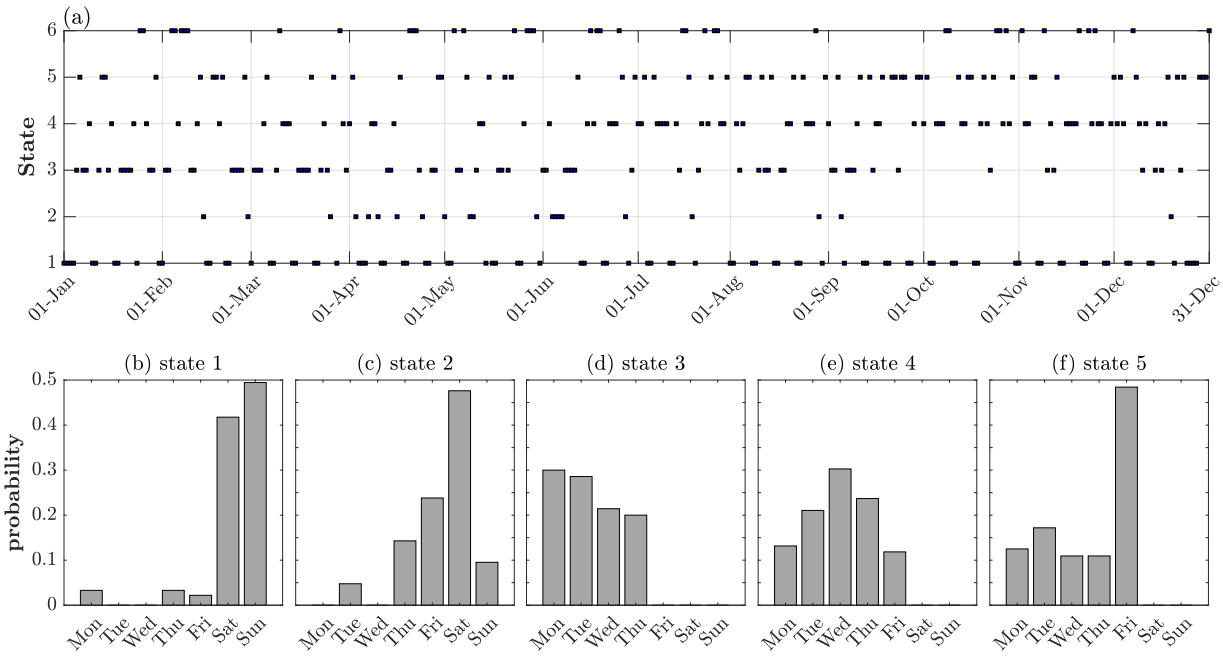

图1： (a) 交通准稳态随时间的演变，其中将缺失数据的日期归类为状态6。(b–f) 各工作日出现给定状态的概率分布。

此外，我们将两个观测值（即降秩相关矩阵）之间的距离定义为平方欧几里得距离

$$E_{ij} = \sum_{n=1}^{N} (\eta_{in} - \eta_{jn})^2 \ . \tag{8}$$

~ ~如果两个降秩相关矩阵 $D(i)$ 和 $D(j)$ 分别与其他矩阵相似，那么这两个矩阵很可能具有相似的结构，因此距离 [1]较短。我们使用平方欧几里得距离矩阵，通过k-均值聚类 [47]对362个降秩相关矩阵进行分类。由于簇内距离平均值的标准差 [48]最小，因此最佳簇数为6。 k-均值聚类是一种基于质心的聚类方法，它会反复将每个观测值分配到距离其最近的质心所属的簇中，直至每个观测值所属的簇不再变化，或迭代次数达到预设的最大值。与其他聚类方法相比，该方法具有复杂度低、计算效率高的优点。

我们将降秩相关矩阵集合中的一个簇称为准稳态。因此，这六个簇对应于交通流时间相关结构中的六个准稳态。有关识别准稳态的更多详情，请参阅文献。 [1]各状态随时间的演变，以及每个工作日出现某一给定状态的概率，展示于图 1中。如文献 [1], 所述，状态1是假日状态，由97%的周末和公共假日组成；而状态3、4和5是工作日状态，包含100%的工作日。具体而言，状态3涵盖周一至周四，状态4涵盖周一至周五，状态5则以周五的数据为主。此外，状态2是混合状态，包含公共假日、周末和工作日。由于相关矩阵中的缺失值会妨碍相似性度量，因此受影响的相关矩阵无法用于聚类算法，我们将所有含缺失值的相关矩阵手动分配至一个额外的簇，即簇6，由此得到状态6。在状态6中，每个相关矩阵内缺失值占所有值的比例从2。07%升至90。87%。

3 准平稳状态之间的转移

现在，我们来研究第2节中识别出的准稳态之间的转移。在第3.1节中，我们介绍 PageRank算法的主要特点，以及用于求解 Google矩阵并获得 PageRank 分数的幂法。在第3.2节中，我们计算准稳态的转移概率，并展示2015年四个季度的状态转移网络。随后，我们在第3.3节中使用 PageRank算法识别交通系统的主导状态，并在第3.4节中探讨随机事件对交通系统的影响。在第3.5节中，我们进一步分析状态重现前通常经历的转移步数。

3.1 PageRank算法

PageRank [29–31] 是一种广为人知的算法，可通过网页的链接结构衡量其重要性。近年来，它被用于评估有向图中节点的重要性、推荐重要节点、预测两个节点之间的链接等 [33]。本文利用该算法，根据准平稳状态的转移概率矩阵来衡量其重要性。

PageRank算法 [29–31, 49] 包含三个重要矩阵：超链接矩阵 H、修正超链接矩阵 S 和 Google矩阵 G。假设有 n 个网页，将每个网页视为一个节点，这些节点之间的连接可构成一个有向网络图，并用 $n \times n$ 超链接矩阵 H表示。如果网页 i 有 l_i ($l_i \geq 1$)个指向其他网页的链接，那么随机冲浪者选择其中任一链接到达网页 j 的概率为 $H_{ij} = 1/l_i$，其中 H_{ij} 是H的一个元素。当随机冲浪者到达一个没有任何链接的悬空网页时，通过选择链接到达其他网页的概率为零，即 $H_{ij} = 0$。此时，人们可以在网络浏览器中输入地址，访问其他网页。选择任一 n 网页的概率为 $1/n$，而不是零。因此，超链接矩阵 H 被修正为一个新的矩阵 S，其方式如下：

$$S = H + \delta w \ , \tag{9}$$

其中，δ 是一个列向量，其元素在 $l_i = 0$ 时为 $\delta_i = 1$，否则为 $\delta_i = 0$；w 是一个均匀行向量 $w = [1/n \ 1/n \ \cdots \ 1/n]$。设标量 α（$0 \leq \alpha < 1$）为阻尼因子，表示随机冲浪者受矩阵 S 引导而移动的概率。此外，冲浪者也有 $1 - \alpha$ 的概率通过在网络浏览器中输入地址，随机跳转到下一个网页，而不考虑当前网页中的链接。因此，矩阵 S 被修改为Google矩阵 G,

$$G = \alpha S + (1 - \alpha)\mathbb{1}\zeta \ . \tag{10}$$

这里，1是一个全1列向量，ζ 是一个行概率分布向量，称为个性化向量。当 α 接近1时，网络的超链接结构会被赋予很高的权重；但当$\alpha = 0$时，超链接结构的作用便会消失。布林和佩奇根据超链接结构的权重以及幂法 [29,30]的收敛速度，选择了 $\alpha = 0.85$ 。他们还指定了一个均匀向量 $\zeta = [1/n \ 1/n \ \cdots \ 1/n]$，以表示随机冲浪者不选择链接、而是浏览任意一个 n 网页的概率。

Google矩阵 G 是一个不可约方阵，因为随机冲浪者可以通过选择链接或在网络浏览器中输入地址，从任意网页前往其他任意网页。 G 中的每个元素表示从一个网页转到另一个网页的概率，取值介于 0 和 1 之间。每一行中各元素的总和等于 1。设 $\pi = [\pi_1 \cdots \pi_n]$ 为一个行概率分布向量，其中 π_n 是该向量的第 n个分量。当行概率分布向量 π 满足 $\pi G = \pi$ [50]时，它就是一个平稳分布，该方程因此有唯一解。估算平稳分布 π 的一种方法是幂法 [31, 51, 52]。矩阵中模最大的特征值称为主特征值。与主特征值对应的特征向量称为主特征向量。幂法是一种用于近似求出方阵主特征向量的技术。给定一个初始向量 $\pi^{(0)}$，可将其设为含有 n 个元素的均匀行向量 $[1/n1/n \cdots 1/n]$，幂法通过连续迭代计算

$$\pi^{(m)} = \pi^{(m-1)}G = \pi^{(0)}G^m \ . \tag{11}$$

当 m 趋于非常大的值，使得 $\pi^{(m)} = \pi^{(m-1)}$时，上述方程得到平稳分布 $\pi = \pi^{(m)}$。如前所述， G 中的元素介于 0 和 1 之间，且 G 每一行中元素的总和等于 1。由于这些特性，Google矩阵 G 是一个行随机矩阵，其主特征值（即 G所有特征值中模最大的值）等于 1 [53]。因此，向量 π 可以视为 G 的主特征向量。作为特征向量分量的每个 π 元素称为 PageRank 分数。网页的 PageRank 分数表示随机冲浪者根据网络的链接结构访问该网页的概率。因此，从 PageRank 的角度来看，该分数可用于估算网页的重要性。对随机冲浪者而言，最重要的网页具有最大的概率，也有最高的 PageRank 分数。

3.2 准平稳状态的转移概率

为估算状态转移的可能性，我们将从状态 i 转移到状态 j 、经过 τ 步的转移概率 $p_{ij}^{(\tau)}$ 定义为：在给定时间段 $\mathcal{T}$ 内，经过 τ 步后，从状态 i 转移到状态 j 的次数 $n_{ij}^{(\tau)}$ ，除以转移到所有状态的总次数 i ，

$$p_{ij}^{(\tau)} = \Pr\Big(X(t+\tau) = j | X(t) = i\Big) = \frac{n_{ij}^{(\tau)}}{\sum\limits_{q=1}^{\mathcal{K}} n_{iq}^{(\tau)}} , \tag{12}$$

其中， X（t）表示时间 t 时的一个状态， $\mathcal{K}$ 表示交通系统中的准稳态总数。此处 $\mathcal{K} = 6$。对于给定的时间段和固定的 τ 步转移， $\sum_{j=1}^{\mathcal{K}} p_{ij}^{(\tau)} = 1$。由于准稳态是在日尺度上识别的， τ 步状态转移表示一个状态在 τ 天后转变为另一个状态。

对于2015年的每个季度，我们计算一步转移的转移概率 $p_{ij}^{(1)}$ ，即经过 $\tau = 1$ 天。根据各项 $p_{ij}^{(1)}$ ，我们得到一个 6×6 转移矩阵 $P^{(1)}$。根据转移矩阵，我们构建准平稳状态之间的有向网络，即状态转移网络。网络中的每个节点代表一个准平稳状态。两个节点之间的每条边都以两个状态之间的转移概率为权重。如果转移概率为零，则两个节点之间不存在边。边的方向表示从一个状态向另一个状态转移的方向。节点的大小表示从其他状态转入该状态的概率总和，即 $\sum_i p_{ij}^{(1)}$ 。如图2所示，四个季度网络结构的差异清楚地揭示了各状态之间转移概率的变化。2015年第一季度，其他状态更可能转移到状态1（即假日状态）和状态3（即工作日状态）。这一时期的交通行为很容易理解：人们通常周一至周四按正常时间工作，周五提前下班，周末休息。到了第二季度，状态2对交通系统也变得重要起来。状态2在工作日和假日出现，其主导地位表明，某些原因改变了交通行为，例如天气状况促使户外活动增加。第三和第四季度，交通系统中从其他状态转移到状态1以及状态4和状态5（即两个工作日状态）的情况更为频繁。尤其是，频繁出现的工作日状态从前两个季度的状态3转变为当前两个季度的状态4和状态5。这表明，某个重大事件（例如道路施工）会对高速公路上的通勤流行为产生影响。此外，在这两个季度中，状态2仅转移到状态1，几乎消失了。交通行为的变化会影响交通流的时间相关结构，导致2015年每个季度的准平稳状态之间形成季节依赖型转移网络。

3.3 准平稳系统中的主导状态

其他状态更有可能转变为某个主导状态，因此后者在某一时间段内出现的可能性很高。2015年，交通系统中哪些状态占主导地位，

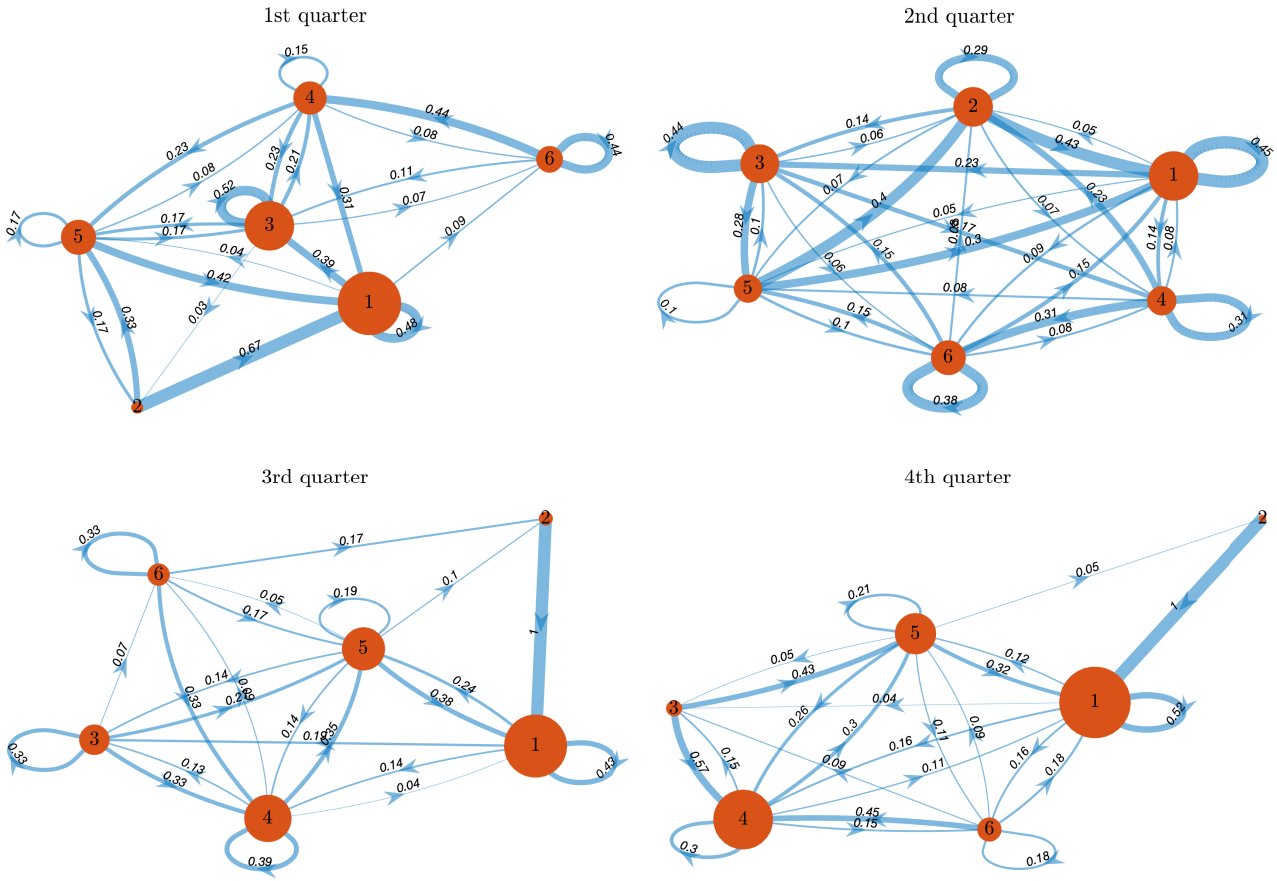

图 2： 2015 年每个季度内六种准平稳状态的转移网络。标有数字的红色节点代表各个状态；节点的大小表示其他状态转移至该状态的概率总和；两个节点之间蓝色连线的宽度表示两个状态之间转移概率的大小；每条蓝色连线上的箭头表示从一个状态转移到另一个状态的方向。

以及随机事件如何影响主导状态，都是值得研究的问题。图2 显示，在较长的时间段内（例如半年或一年），交通系统中准平稳状态之间的转移呈现出非马尔可夫特征。只有在相对较短的时间段内，准平稳状态才可能表现出马尔可夫方式。我们将这个较短的时间窗口设为 $\mathcal{T} = 60$ 天，并始终将该时间窗口向前移动一天。这样，我们便可利用滑动时间窗口计算出一系列转移矩阵 $P^{(\tau)}(d)$，其中每个时间窗口的最后一天标记为时间点 d。

为确定某一时间段内的主导状态，我们将 PageRank 算法应用于转移矩阵 $P^{(\tau)}(d)$。具体而言，准平稳状态相当于网页，而状态转移矩阵 $P^{(\tau)}(d)$ 相当于超链接矩阵 H。常规交通系统中的随机事件，例如交通事故、道路施工、城市中的大型活动或恶劣天气状况，与随机冲浪者通过在网络浏览器中输入地址访问网页、而非点击上一网页上的链接这一随机行为相似。对于每个时间窗口d，我们这里的 6×6 Google 矩阵表示为：

$$G(d) = \alpha\Big(P^{(\tau)}(d) + \delta w\Big) + (1-\alpha)\mathbb{1}\zeta \ . \tag{13}$$

这里，列向量 $\delta = (\delta_1, \ \cdots, \ \delta_6)$，其中 $\delta_k = 1$ 若指向其他状态的链接数为零，即 $l_k = 0$，而 $\delta_k = 0$ 则表示其他情况。均匀行概率向量 w 和 ζ 分别定义为$w = [1/6 \ \cdots \ 1/6]$ 和 $\zeta = [1/6 \ \cdots \ 1/6]$。在我们的案例中，PageRank 分数用于衡量一个准稳态的重要性。与其他状态不同，主导状态是最重要的

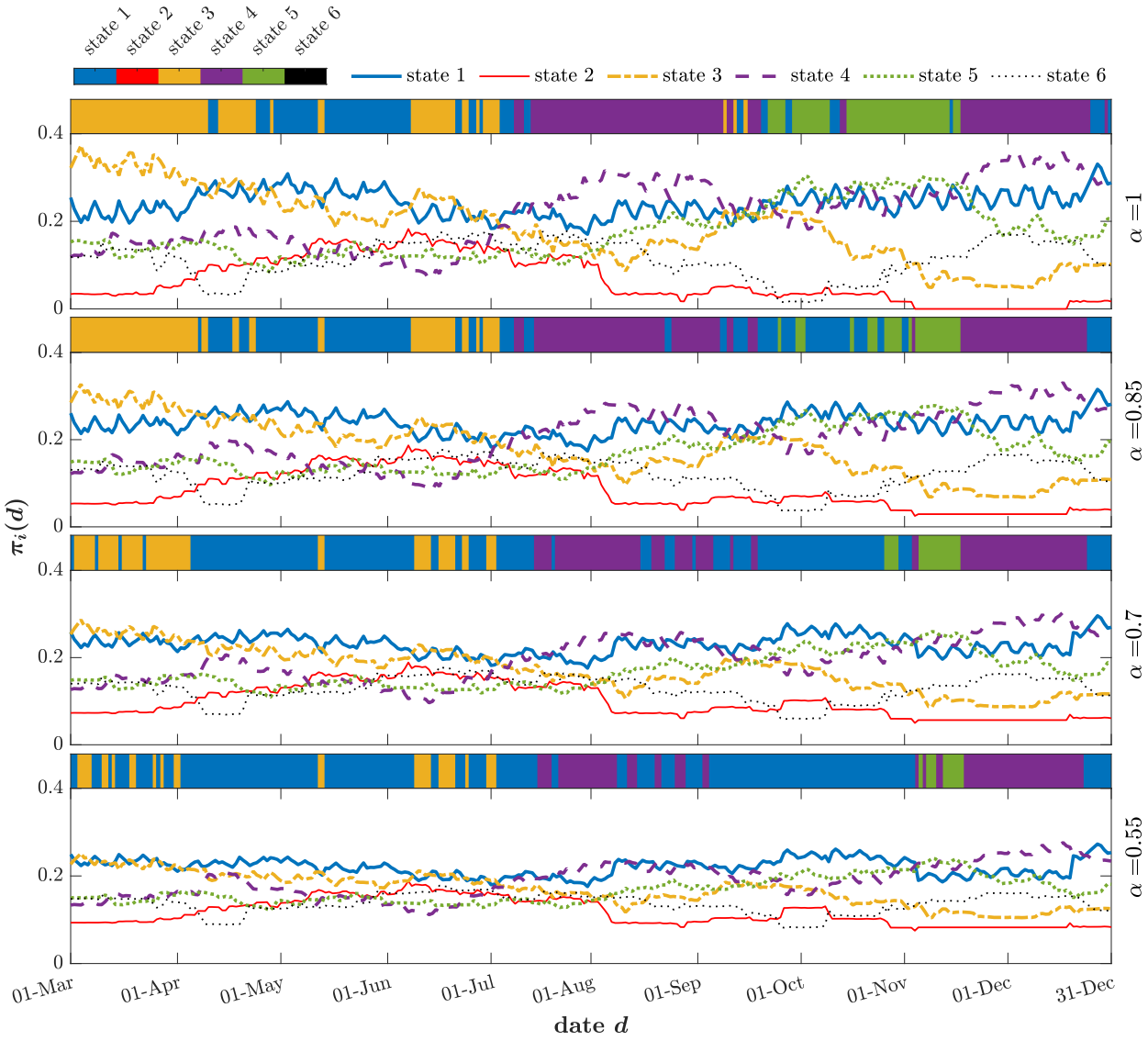

图 3： 阻尼因子分别为 $\alpha = 1$、0.85、0.7、0.55 时，六个准平稳状态的主导状态（以长色条表示）和 PageRank 分数 $\pi_i(d)$（以线条表示）随时间的演变。由于时间窗口长度为 60 天，2015 年前两个月的结果未显示。

状态，即 PageRank 分数最高的状态。方程（$\pi(d)G(d) = \pi(d)$）的解为式（11）的结果，其中包含时间窗口 $\pi(d)$ 内六种状态的 PageRank 分数 d。

对于一步转移，即 $\tau = 1$ 天，当阻尼因子为 $\alpha = 1$ 且列向量为$\delta = (0, \cdots, 0)$，这表示任何状态都可能转移到任何其他状态，方程（13）的解为

$$\pi^{(m)} = \pi^{(0)}\left(P^{(1)}(d)\right)^{m}, \tag{14}$$

其 m 非常大。在此我们设定 $m = 1000$。通过这一设定，我们获得了纯经验性的 PageRank 分数结果，这些结果对六种状态均存在且唯一。对于每个时间窗口 d，PageRank 分数最高的状态即为主导状态，其主要取决于季节。具体而言，从一月到三月，状态3（即周一至周四出现的工作日状态）主导交通系统。从四月到六月，状态1占据主导地位，这表明与其他时间段相比，此期间有更多人外出度假。从七月到九月，状态4最为重要，意味着更多人结束假期返回，并形成周一至周五的常规通勤。从十月到十二月，状态5和状态4（即两种工作日状态）主导其他状态。圣诞节假期状态（即状态1）在十二月底出现，这一点并不意外。此外，各状态的重要性

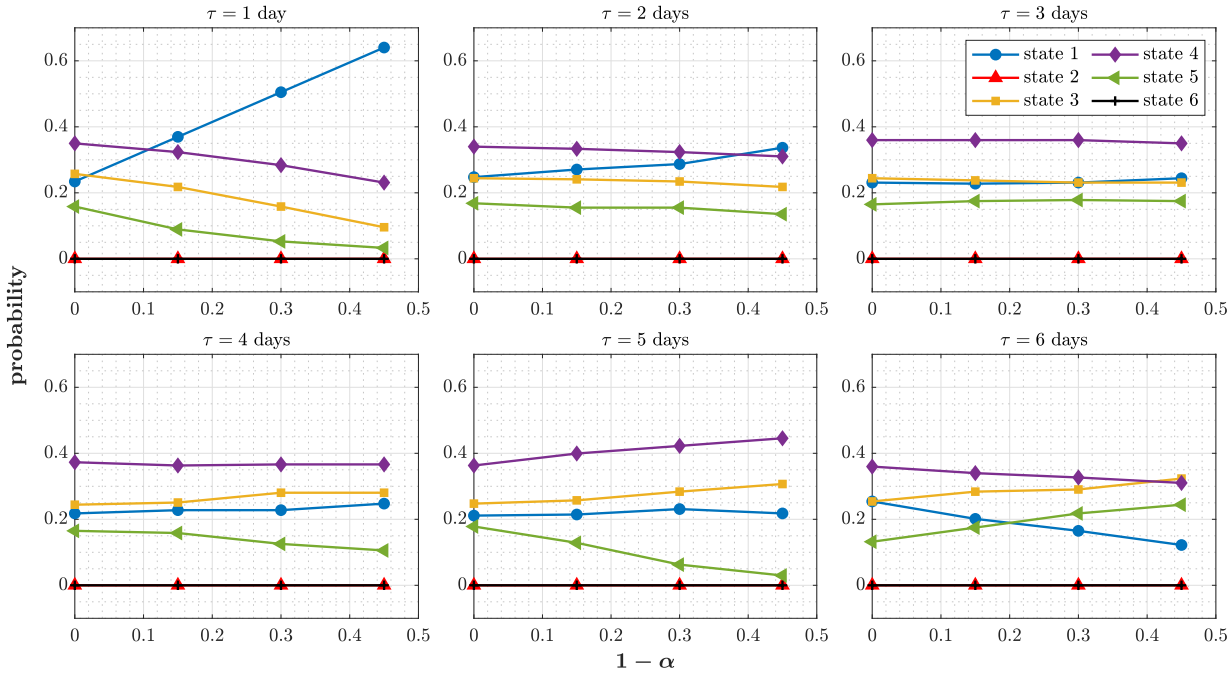

图4： 2015年以60天为时间窗口时，主导状态的发生概率与随机事件的发生概率之间的关系 $1-\alpha$

在六月和七月，六种状态的重要性趋于接近。在此期间，主导状态频繁变化，状态2的 PageRank 分数也接近较高水平，而其他时段则处于较低水平。这种剧烈变化表明，事件对常规交通系统有很大影响，进而导致交通流相关性的调整。追溯至2015年，我们发现A3高速公路在科隆-米尔海姆与勒沃库森之间的道路扩建于六月开始，¨ 2015 [54]。这一事件在一定程度上降低了相应高速公路路段的通行能力，并改变了科隆环城高速公路交通流的分布。

3.4 随机事件对交通系统的影响

为探究随机事件（如2015年A3 高速公路上上述道路施工）对总体交通行为的影响，我们将阻尼因子 α 分别从1降低至0.85、0.7和0.55。阻尼因子的降低对应于交通系统中随机事件发生概率的增加 $1-\alpha$。如图3所示，随机事件的增加延长了由假日状态主导的时间段，同时缩短了由工作日状态主导的时间段。回顾2020年德国爆发的冠状病毒疫情。2020年期间，大多数人被鼓励居家办公或被建议居家隔离 [55]。另一个类似的例子与出现极端天气的时期有关 [56], ，例如暴雨、龙卷风和洪水。这类极端天气限制大多数人进行户外活动。因此，工作日繁重的交通负担会因这些随机事件而得到缓解，交通系统的表现如同假日时一样。

最后一个结论适用于一步状态转移。如果状态之间的转移需要两步或更多步，会怎样？为研究这一问题，我们计算了不同转移步数下主导状态的概率与随机事件的概率，即 $1-\alpha$， $\tau = 1$、2、$\cdots$、6天。主导状态的发生概率，是指在全年中某一状态主导交通系统的次数，占全年交通系统中所有主导状态总发生次数的比例。对于一步转移，图4中主导状态发生概率的趋势，从定量上印证了图3中对主导状态持续时间的定性评估。对于两步至四步转移，尽管随机事件的概率不断增加，主导状态发生概率的整体趋势仍保持稳定。然而，

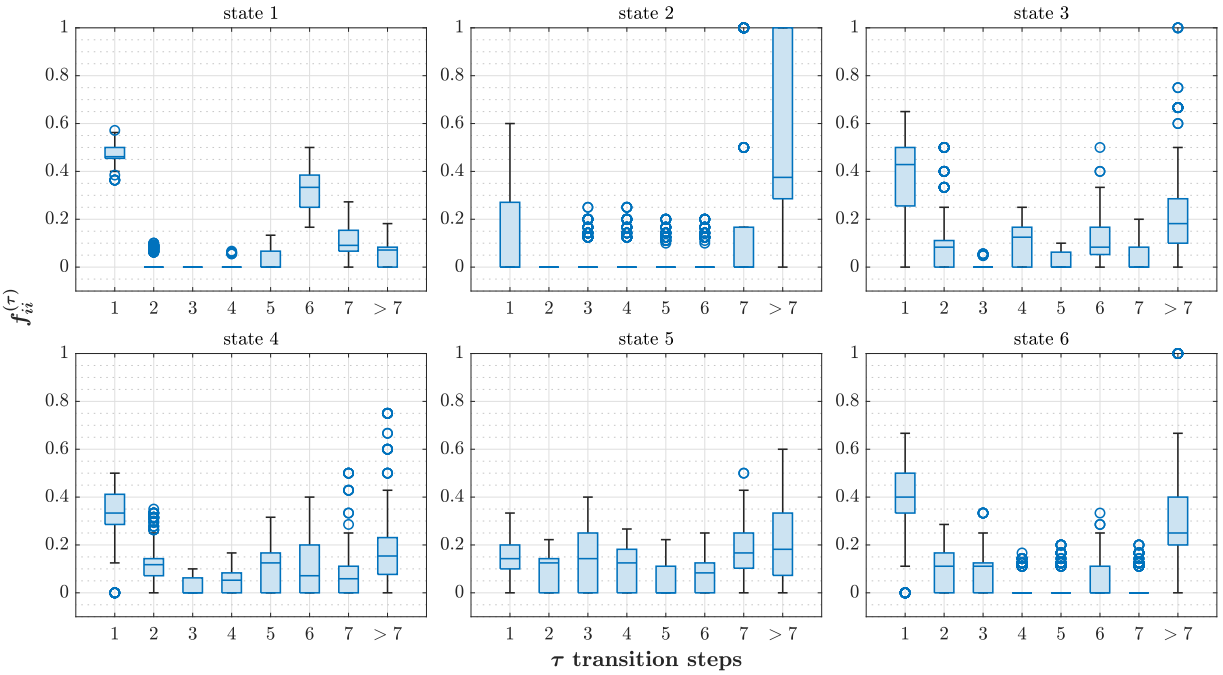

图5： 2015年时间窗口为60天时，经过不同转移步数 τ 后，状态 i 的概率 $f_{ii}^{(\tau)}$

在五步至六步转移后，状态5的发生概率分别明显下降和上升。看来，初始时刻随机事件的增加，使得五天后状态5成为主导状态的可能性降低，却使其在六天后成为主导状态的可能性增大。状态1也呈现出这种相互矛盾的变化。如前所述，随机事件增加时，一步转移后状态1更有可能主导交通系统。然而，六步转移后，状态1成为主导状态的可能性则会降低。我们可以设想，某一天发生随机事件（例如冬季风暴）后，次日高速公路上的交通负荷会减轻，因为许多人倾向于在家工作，或因这一事件请假。次日之后或接下来的几天，这种情况会有所改变；但由于管理部门采取了措施，这种情况持续长达六天的可能性较低。

3.5 准平稳状态的重现

上述研究重点关注经过 τ 步后从一个状态转移到相同或不同状态。接下来，我们将把研究范围限定在相同状态之间的转移，并探讨某个状态在经过一定步数后首次重现的可能性。

我们将状态 i 在 τ步转移中首次返回自身的概率定义为

$$f_{ii}^{(\tau)} = \Pr\Big(X(t+\tau)=i, X(t+\tau')\neq i \text{ for } \tau'=1,\cdots,\tau-1 | X(t)=i\Big) = \frac{n_{ii}^{(\tau)}}{\sum_{\tau=1}^{\infty} n_{ii}^{(\tau)}}, \qquad (15)$$

使得 $\sum_{\tau=1}^{\infty} f_{ii}^{(\tau)} = 1$。这里，$n_{ii}^{(\tau)}$表示状态 i经过 τ步转移后首次重现的次数。图5展示了 $f_{ii}^{(\tau)}$在每种步数 τ下各状态的概率。状态1是假日状态，主要出现在星期六和星期日。因此，状态1首次重现通常发生在一步或六步转移之后。状态2由工作日和假日组成，在图3所示的六月前后特殊时期才略显重要。因此，从一年尺度来看，状态2通常要在一周以上才会首次回到自身。状态3和状态4只出现在工作日，因此它们很可能在一步转移后首次重现。尽管状态5只包含工作日，但它在星期五出现的频率高于其他工作日

。因此，它要么在一周后再次出现，要么在一周以上后再次出现。我们将所有含缺失值的相关矩阵归为状态6。这些缺失值是由感应线圈检测器故障或道路施工等事件造成的，这些情况会妨碍交通数据的检测。状态6首次重现大多发生在一步转移后，也常发生在七步以上的转移后。这表明，如果发现交通数据检测故障，那么很可能在一天后再次发现此类故障。故障也可能在一周以上后再次出现，但其频率不及一天后再次出现的情况。与图1 相比，图5 量化了各状态首次重现所需步数的概率，尤其是状态2和状态6。

4 结论

我们以科隆环城高速公路为例，研究了交通系统时间相关结构中准平稳状态之间的转移。我们发现了五种准平稳状态，分别代表交通流降秩相关矩阵中的五种不同结构。在降秩相关矩阵中，交通的集体行为和噪声信息已被去除。我们还考虑了另一种包含带有缺失值的相关矩阵的准平稳状态。这六种准平稳状态共同构成了2015年的状态时间序列。因此，我们据此计算了状态之间的转移概率。

根据转移概率矩阵构建的状态转移网络揭示了状态的季节性依赖。具体而言，在2015年前两个季度，交通系统更有可能进入状态1（即假日状态）和状态3（即工作日状态）；而在后两个季度，则更有可能进入状态1和状态4（即工作日状态）。此外，状态2（即同时包含工作日和假日的状态）在2015年第二季度出现得更频繁。

在交通系统中，最重要的状态是某一时间段内占主导地位且出现概率较高的状态。我们利用PageRank算法，识别了六种准稳态的重要性；所用时间窗口为 60 天，并逐日向前滚动。主导状态再次揭示了季节性依赖。对于一步转移，PageRank算法模拟的随机事件发生概率增加，会使交通系统的行为更像假日而非工作日。相反，在六步转移中，随着随机事件增加，假日状态在交通系统中的主导地位逐渐减弱。

准稳态首次重现所需的转移步数，与该状态的特征密切相关。假日状态，即主要出现在星期六和星期日的状态1，首次通过一步或六步转移返回自身的概率很高；而两种工作日状态，即常在星期一至星期五出现的状态3和状态4，则经常在一天后首次重现。另一种工作日状态，即常在星期五出现的状态5，通常经过较长时间后才重现，至少需要一周。状态6的首次重现可用于估算感应线圈检测器发生故障的时间。检测器通常似乎会在两个不同的时间尺度上发生故障：要么一天后，要么在一周以上后。

对准平稳状态转移的实证研究，有助于揭示传统方法无法呈现的交通系统隐藏信息。这些信息，例如某个准平稳状态的出现或重现，反过来可能有助于阐明交通行为和运行模式，以及相关模式在空间或时间上的发展。

致谢

我们衷心感谢德国研究联合会（DFG，418382724）通过资助项目“Korrelationen und deren Dynamik in Autobahnnet-zen”提供经费。感谢 Strassen.NRW 提供实测交通数据。我们还感谢 Sebastian Gartzke 进行了富有成效的讨论。

作者贡献

T.G. 和 M.S. 提出了这项研究。S.W. 和 T.G. 制定了分析方法。S.W. 完成了所有计算。所有作者均同等参与了结果分析、论文撰写和审阅。

参考文献

[1] Shanshan Wang、Sebastian Gartzke、Michael Schreckenberg 和 Thomas Guhr。《交通系统时间相关性中的准稳态：以科隆环形高速公路为例》。*J.Stat. Mech.: Theory Exp.*，2020(10):103404，2020年。[2] Michael C M¨unnix、Takashi Shimada、Rudi Sch¨afer、Francois Leyvraz、Thomas H Seligman、Thomas Guhr 和 H Eugene Stanley。《识别金融市场状态》。*Sci. Rep.*，2:644，2012年。[3] Desislava Chetalova、Rudi Sch¨afer 和 Thomas Guhr。《深入观察市场状态》。*J. Stat.Mech.: Theory Exp.*，2015(1):P01029，2015年。[4] Desislava Chetalova、Marcel Wollschl¨ager 和 Rudi Sch¨afer。《市场状态的依赖结构》。*J. Stat. Mech.: Theory Exp.*，2015(8):P08012，2015年。[5] Thomas Guhr。《金融市场的非平稳性：市场状态动力学与一般特征》。*Acta Phys. Pol. B*，46(9)，2015年。[6] Philip Rinn、Yuriy Stepanov、Joachim Peinke、Thomas Guhr 和 Rudi Sch¨afer。《准稳态系统的动力学：以金融为例》。*EPL*，110(6):68003，2015年。[7] Yuriy Stepanov、Philip Rinn、Thomas Guhr、Joachim Peinke 和 Rudi Sch¨afer。《准稳态的稳定性与层级结构：以金融市场为例》。*J. Stat. Mech.: TheoryExp.*，2015(8):P08011，2015年。[8] Anton J Heckens、Sebastian M Krause 和 Thomas Guhr。《揭示相关结构相对于市场集体运动的动力学》。*J. Stat. Mech.: Theory Exp.*，2020(10):103402，2020年。[9] Hirdesh K Pharasi、Eduard Seligman、Suchetana Sadhukhan 和 Thomas H Seligman。《市场状态动力学与风险评估》。*arXiv:2011.05984*，2020年。[10] Hirdesh K Pharasi、Eduard Seligman 和 Thomas H Seligman。《市场状态：一种新的理解》。*arXiv:2003.07058*，2020年。[11] Hirdesh K Pharasi、Suchetana Sadhukhan、Parisa Majari、Anirban Chakraborti 和 Thomas H Seligman。《相关矩阵空间中的市场状态动力学及其在金融市场中的应用》。*arXiv:2107.05663*，2021年。[12] Michael E Mann。《关于平滑可能非平稳的气候时间序列》。*Geophys. Res.Lett.*，31(7):L07214，2004年。[13] Linyin Cheng、Amir AghaKouchak、Eric Gilleland 和 Richard W Katz。《气候变化背景下的非平稳极值分析》。*Clim. Change*，127(2):353–369，2014年。[14] Arturo Roman Messina。《电力系统中的区域间振荡：非线性与非平稳视角》。Springer Science & Business Media，2009年。[15] Israel Cohen 和 Baruch Berdugo。《非平稳噪声环境下的语音增强》。*Signal Process.*，81(11):2403–2418，2001年。

[16] Sundarrajan Rangachari 和 Philipos C Loizou。《适用于高度非平稳环境的噪声估计算法》。Speech Commun.，48(2):220–231，2006年。[17] Kai Nagel 和 Michael Schreckenberg。《用于高速公路交通的元胞自动机模型》。J. Phys. I，2(12):2221–2229，1992年。[18] Andreas Schadschneider 和 Michael Schreckenberg。《元胞自动机模型与交通流》。J. Phys. A: Math. Theor.，26(15):L679，1993年。[19] Gunnar G Løv˚as。《行人交通流的建模与仿真》。Transp. Res. B: Methodol.，28(6):429–443，1994年。[20] Michael Schreckenberg、Andreas Schadschneider、Kai Nagel 和 Nobuyasu Ito。《交通流的离散随机模型》。Phys. Rev. E，51(4):2939，1995年。[21] Serge P Hoogendoorn 和 Piet HL Bovy。《车辆交通流建模的最新进展》。Proc. Inst. Mech. Eng., I: J. Syst. Control Eng.，215(4):283–303，2001年。[22] GCK Wong 和 SC Wong。《多类别交通流模型——对含异质驾驶员的 LWR模型的扩展》。Transp. Res. Part A Policy Pract.，36(9):827–841，2002年。[23] Martin Fellendorf 和 Peter Vortisch。《微观交通流模拟器 VISSIM》。载于《交通仿真基础》，第63–93页。Springer，2010年。[24]Martin Treiber 和 Arne Kesting。《交通流动力学：数据、模型与仿真》。Springer-Verlag Berlin Heidelberg，2013年。[25] Boris S Kerner。《高速公路瓶颈处时空交通模式的宏观经验特征》。Phys. Rev. E，65(4):046138，2002年。[26] Robert L Bertini 和 Monica T Leal。《高速公路车道减少处交通特征的实证研究》。J. Transp. Eng.，131(6):397–407，2005年。[27] Martin Sch¨onhof 和 Dirk Helbing。《拥堵交通状态的经验特征及其对交通建模的启示》。Transp. Sci.，41(2):135–166，2007年。[28] Boris S Kerner。《交通物理学：高速公路模式的经验特征、工程应用与理论》。Springer，2012年。[29] Sergey Brin 和 Lawrence Page。《大型超文本网络搜索引擎的剖析》。Comput. Netw. ISDN Syst.，30(107-117)，1998年。[30] Lawrence Page、Sergey Brin、Rajeev Motwani 和 Terry Winograd。《PageRank引文排名：为网络梳理秩序》。技术报告，Stanford InfoLab，1999年。[31] Amy N. Langville 和 Carl D. Meyer。《Google的PageRank及其拓展：搜索引擎排名科学》。普林斯顿大学出版社，2011年。[32] Paolo Boldi、Massimo Santini 和 Sebastiano Vigna。《PageRank：函数依赖》。ACM Trans. Inf. Syst.，27(4):1–23，2009年。[33] David F Gleich。《网络之外的PageRank》。SIAM Rev.，57(3):321–363，2015年。[34] Julie L Morrison、Rainer Breitling、Desmond J Higham 和 David R Gilbert。《Generank：利用搜索引擎技术分析微阵列实验》。BMC Bioinform.，6(1):233，2005年。[35] Barbara Logan Mooney、L Ren´e Corrales 和 Aurora E Clark。《Molecularnetworks：一种综合图论与数据挖掘工具，用于探索分子模拟中的溶剂组织结构》。J. Comput. Chem.，33(8):853–860，2012年。

[36] Jonathan J Crofts 和 Desmond J Higham。《用谷歌探索大脑：发现层级和非对称网络结构及其在神经科学中的应用》。Internet Math., 7(4):233–254, 2011年。[37] Klaus M Frahm、AD Chepelianskii 和 DL Shepelyansky。《整数的 PageRank》。J. Phys. A Math. Theor., 45(40):405101, 2012年。[38] Leonardo Ermann 和 Dima L Shepelyansky。《多产品世界贸易网络的谷歌矩阵分析》。Eur. Phys. J. B, 88(4):1–19, 2015年。[39]C´elestin Coquid´e、Leonardo Ermann、Jos´e Lages 和 Dima L Shepelyansky。《基于联合国商品贸易统计数据库数据的约化谷歌矩阵分析：石油和天然气贸易对欧盟经济体的影响》。Eur. Phys. J. B, 92(8):1–14, 2019年。[40] C´elestin Coquid´e、Jos´e Lages 和 Dima L Shepelyansky。《世界贸易网络中的危机传染》。Appl. Netw. Sci., 5(1):1–20, 2020年。[41] Bin Jiang、Sijian Zhao 和 Junjun Yin。《用于预测交通流的自组织自然道路：一项敏感性研究》。J. Stat. Mech.: Theory Exp., 2008(07):P07008, 2008年。[42] Arieh Schlote、Emanuele Crisostomi、Stephen Kirkland 和 Robert Shorten。《电动汽车交通建模框架》。Int. J. Control, 85(7):880–897, 2012年。[43] Shanshan Wang、Michael Schreckenberg 和 Thomas Guhr。《识别大型高速公路网络中的次主导集体效应》。J. Stat. Mech.: Theory Exp., 2022(11):113402, 2022年。[44] Anton J Heckens 和 Thomas Guhr。《金融协方差与相关性的新集体性度量》。Physica A, 第127704页, 2022年。[45] Anton J Heckens 和 Thomas Guhr。《识别市场相关结构中内生性金融危机长期前兆的新尝试》。J. Stat. Mech.: Theory Exp., 2022(4):043401, 2022年。[46]Shanshan Wang、Sebastian Gartzke、Michael Schreckenberg 和 Thomas Guhr。《北莱茵-威斯特法伦州高速公路网络中的集体行为》。J. Stat. Mech.: Theory Exp., 2021(12):123401, 2021年。[47] Stuart Lloyd。《脉冲编码调制中的最小二乘量化》。IEEE Trans. Inf. Theory, 28(2):129–137, 1982年。[48] Hirdesh K Pharasi、Kiran Sharma、Rakesh Chatterjee、Anirban Chakraborti、Francois Leyvraz 和 Thomas H Seligman。《利用相关模式识别金融市场崩盘的长期前兆》。New J. Phys., 20(10):103041, 2018年。[49] Rebecca S Wills。《谷歌 PageRank：搜索引擎背后的数学》。Math. Intell., 28(4):6–11, 2006年。[50] Thomas M Cover。《信息论基础》。John Wiley & Sons, 1999年。[51] Richard Bronson 和 Gabriel Costa。《矩阵方法：应用线性代数》。EBL-Schweitzer。Elsevier Science, 2008年。[52] Gene H Golub 和 Charles F Van Loan。《矩阵计算》。约翰·霍普金斯大学出版社, 第4版，2013年。[53] Taher Haveliwala 和 Sepandar Kamvar。《谷歌矩阵的第二特征值》。技术报告 2003-20，Stanford InfoLab，2003年。[54]德国北莱茵-威斯特法伦州州府交通部新闻办公室。《科隆与勒沃库森之间的A3高速公路于5月全面扩建为八车道》。网址：https://www.land.nrw/de/pressemitteilung/a3-zwischen-koeln-und-leverkusen-im-mai-komplett-mit-acht-fahrstreifen，2017年5月8日。

[55] 维基百科。《德国2019冠状病毒病疫情》。https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Germany，2022年9月。[56] 维基百科。《2021年天气》。https://en.wikipedia.org/wiki/Weather_of_2021，2022年9月。