**The complex network of global cargo ship movements**

Pablo Kaluza, Andrea Kölzsch, Michael T. Gastner, Bernd Blasius

## Abstract

运输网络在人类交通、商品交换以及入侵物种的传播中扮演者重要的角色。因为百分之九十的世界商品经由海洋运输，商船的全球网络提供了最重要的运输方式之一。这里我们用在2007年的16363条商船的线路信息建立了一个连接港口的网络。我们发现这个网络具有一些区别于其他运输网络的特点。特别是大多数船只可以分为三类：干散货船（運輸穀物、煤、礦砂、鹽、[水泥](http://wiki.mbalib.com/zh-tw/%E6%B0%B4%E6%B3%A5)等大宗乾散貨物的船舶）、集装箱运货船和油船。这三个类别不仅具有不同的物理特性，而且具有不同的移动模式和网络。集装箱船有定期重复的路线而干散货船和运油船在港口之间的路线是不太可预测的。所有船只的移动网络在港口的连通性和运输链路上的负载与船舶类型之间的系统差异具有重尾分布。本文经过对数据的分析改进了基于重力模型的船舶移动的假设，这是了解全球贸易与生物入侵的重要一步。

## Introduction

旅行的能力、世界贸易和具有空前效率的信息共享是现代全球化经济最本质的特征。在不同的运输方式中，海运是最节能的长途运输大量货物的方式。据估计，世界贸易的百分之九十是由船只完成的。2006年，74亿吨的货物通过全世界的港口装载。船舶贸易量超过30万亿吨-英里，且增长率超过了全球经济。

全球的海事网络在当今的外来入侵物种的传播中也扮演者重要的角色。两个海洋物种入侵的主要方式是船舶压载舱排水和船体污染。甚至陆地物种如昆虫优势也会无意中由集装箱运输。在世界的一些地方，外来物种已经造成了戏剧性地物种灭绝和自然景观改变，进而破坏生态系统，给人类生存、健康和当地经济造成危害。单美国一国每年因为生物入侵造成的经济损失高达1200亿美元。

尽管航运影响着每个人的日常生活，其相比于其他全球运输业总是的到公众更少的关注。因此，它在最近的复杂网络文献中也没有的到多少关注。考虑到当前大众对网络的兴趣，特别是航空、道路、和铁路网络，这种忽视是令人惊讶的。本着当前网络研究的精神，我们以全球航运网络（GCSN）的大规模视角将其视为一个复杂系统，定义为如果一条船的路线经过两个港口，则这两个港口是连通的。

过去类似的研究不得不对在假想网络中的港口对之间的流量做出强假设，以便近似船舶的运动。相比之下，我们的分析是基于全面的真实的船舶运动数据，而这使我们能够构建实际的网络。我们发现，它具有一个小世界拓扑结构，其中在给定港口（以总吨位（GT）计量）的船舶的组合货物容量遵循重尾分布。这个容量与直接连接的端口数量呈超线性。我们确定了网络最中心的港口，并发现了几组高度互联的港口，显示了区域地缘政治和交易区块的重要性。

然而，对于整个网络的高级描述尚未完全捕获网络的复杂性。与先前研究的网络不同，GCSN具有多层结构。概括的来说，GCSN有三类货船---集装箱产、散货船和油船---这三类货船跨越不同的子网。不同类别的货船倾向于联系不同的港口，并且以不同的模式航行。我们分析GCSN中不同货船的轨迹并开发出了提取关于不同航行类型运动特征的技术。通过这些方法，我们可以量化的得出结论，集装箱船的航行比油船或散货船更可预测，经常重复航行路线。我们将经验数据与由重力模型计算的理论交通流进行比较。基于完整的GCSN数据或重力模型的模拟结果与在世界港口间入侵物种传播的人口动态模型相比较差异很大。因此，基于真实网络的预测对于有关世界贸易稳定和降低生物入侵风险的国际政策决策更为有益。

## Data

对全球船舶运动的分析需要详细了解船舶在其港口的到达和离开的事件。近几年这样的数据已经可以使用。从2001年开始，船舶和港口已经开始安装自动识别系统（AIS ）。船上的AIS发射机自动向港口报告到达和离开的时间。该技术主要用于避免冲突和增加港口安全，但到达和离开的记录同时被Lloyd's Register Fairplay基于商业用途记录在其海洋网络数据库作为其一部分。由于AIS设备还没有安装在所有的船只上，因此数据与实际存在一些差别。然而所有主要的港口和大船只都包含在数据库中，故该数据库能够在表大部分货运船只。

我们的研究是基于海洋网络数据库2007年船只进出港记录以及海洋网（sea-web）关于船舶物理特征的数据。我们将研究限制在超过万吨级的货船，占世界货船运输总量的93%。从这些数据中，我们选择了AIS数据可用的16363艘货船，代表全球海运交通和配备AIS接收器的951个港口之间的长途交易（详见电子补充材料）。对于每艘船，我们从数据库获取其轨迹，即按日期排序的船舶所访问的港口列表。2007年，共有490517次不间断旅程，连接了36351个不同的到达和离港的港口。完整的轨迹组合如图1a所示，其中每条路径代表了海上的最短路线，颜色代表经过该路线的船只数量。

每条轨迹可以解释为一个小的有向网络，其中如果船舶直接经过两个港口，那么这两个港口的节点就是连通的。可以通过合并不同船舶的轨迹来形成更大的网络。本文中，我们用四中方式聚合轨迹：所有可用轨迹的聚合网络和集装箱船（3100艘）、散货船（5498艘）和油船（2628）的子网络。这三个子网络组合覆盖了GCSN总吨数（GT）的74%。在所有四个网络中，我们为港口i到j的链路分配了权重wij ，等于在2007年测量的在该链路上航行的所有船舶上可用空间的GT总和。如果一艘船从i到j的旅程不止一次，它对该链路权重的贡献就有多次。

## The global cargo ship network

整个有向网络呈现明显的不对称性，59%的港口对只在一个方向上连通。尽管如此绝大多数港口（951中的935）属于一个单一的强连通组件，即对于该组件中的任意两个港口，在两个方向上都有路径，尽管两个方向可能会访问不同的港口。有趣的是这些路径都很短：网络中从一个港口到另一个港口仅仅需要几步。两个港口之间的最短路径长度l是从起点到目的地之间行进的最少不间断连接数。在GCSN中，所有成对港口的l的平均值而非常小，即<l>=2.5.即使是任意两个港口的最长的最短路径（例如从阿根廷Skagway到意大利的兰佩杜萨岛）只有lmax=8。事实上，所有可能的起始-目的地（52%）中的大部分可以通过两步或者更少连接。

将这个结果和世界范围机场网络（WAN）的结果进行比较，显示出有趣的差异和相似之处。GCSN的高度不对称性在WAN中没有被发现，表明船舶交通在结构上与航空有很大的不同。船舶网络由于不是往返行程的积累而形成的，船只的行程似乎是单向的，通常被安排成圆形的路线。GCSN的最短路径长度也显示出这种最优性。相比之下，在WAN中，平均和最长的最短路径长度分别是<l>=4.4和lmax=15（Guimeràet al。2005），即GCSN中的两倍长。与WAN类似，GCSN是高度集群的：如果港口x连接到港口y和z，则存在y到z的连接的可能性很高。我们计算了其聚类系数C（Watts＆Strogatz 1998），并发现C=0.49，而具有相同数量节点和链路的随机网络平均聚类系数仅为C=0.04。度依赖聚类系数Ck表明聚集程度随节点的度降低而降低。因此，尽管拥有大量的聚类，像WAN这样的GCSN可以被认为是具有短路径长度的小世界网络（Watts＆Strogatz 1998）。然而GCSN的平均度，即到达和离开给定港口的平均连接数（入度加出度）为<k>=76.5,明显高于WAN,其<k>=19.4（Barrat等人，2004）。根据网络的大小（WAN包含3880个节点），这种差异变得更加明显，表明GCSN的密度更高。链路的这种冗余使得网络具有较高的结构稳健性，可以减少维护带来的路线损失。

度分布P(k)表明大部分港口只有几个链接，但是还有一些港口连接了上百个其他港口（图表 2a）。类似的右偏态度分布已经在好多现实世界网络中被观察到。尽管GCSN的度分布不是完全的无标度的，但是链路权重P(w)的分布大致遵循幂率Embedded Image 其中μ= 1.71±0.17（95%可信区间线性回归，图 2b）。通过对到达和离开港口i的链路权重的和进行平均，我们获得节点强度si 强度分布也近似幂率Embedded Image 其中Embedded Image，意味着有少量的港口处理大量的货物（图片2c）。通过线拟合来确定幂率关系受到了强烈批评，所以我们通过Akaike权重模型选择分析了其分布。结果证实了对于P(w)和P(s)幂次定律比指数或对数正态分布更好，但对于P(k)不是。这些结果很好的符合了亚洲提出的货物交通的枢纽-辐射网络。网络中有一些大型的高度连通的港口，所有较小的港口通过这些港口交易。这种无标度的特性使得船舶贸易网络易于外来入侵生物的传播和持续。平均最近邻居度-是网络分散性的度量，再次强调了货船交通的枢纽-辐射属性。

端口的权重和度是根据缩放关系相关的，关系为Embedded Image（95%置信区间标准化主轴回归）。因此，港口的权重通常比其度增长的更快（图表 2d）。换句话说，高度连接的港口不仅具有很多连接，而且他们的连接的权重也比平均权重更高。这个观察结果与以下事实一致：繁忙的港口能够更好的处理大型船舶及大量的货物。相同的结果---Embedded Image 同样在航空中发现，这可能暗示了运输网络的一般模式。在生物入侵方面，这些结果验证了实证结论，大型港口由于越来越大的船舶压载水增加而加剧了生物的入侵。

进一步表现节点重要性的是其中介中心性。港口的中间性是通过该港口的网络中拓扑最短定向路径的数量。在图1b中，我们绘制列出了最中心的港口。一般来说，中心性和度是强相关的（皮尔森相关系数：0.81），但是在个别情况下，其他因素也会发挥一些作用。例如，巴拿马和苏伊士运河是避免南美到非洲长距离航行的捷径。其他港口具有较高的中心性，是因为它们有大量船只访问（例如上海），而其他港口主要通过连接到许多不同的港口（例如安特卫普）而获得其地位。

## The network layers of different ship types

为了比较不同类型货船的移动，为三种主要类型的船只分别构建了单独的网络：集装箱船、散货船和油船。将前一节介绍的网络参数应用到这三个子网中，揭示了一些大规模的区别（表1）。集装箱船网络是非常密集的集群，C = 0.52，有着相当低的平均度，<k> = 32.44，和每条航线很大的平均旅行（任何船只经过的次数）次数，<j> = 24.26。另一方面散货船网络，是低集群的，有着高平均度和更少的平均旅行次数（C = 0.43，<k> = 44.61,<j> = 4.65）。对于油船网络，我们发现其中间值（C = 0.44,<k> = 33.32,<j> = 0.57）。注意到，子网络的平均度<k>远小于完整的GCSN的平均度，表明不同的船只类型使用基本相同的港口但是有不同的路线。

在链路权重的缩放中出现类似的趋势（图2b）。P（w）在每个网络都可以近似为幂率，但是具有不同的指数μ。集装箱船具有最小的指数（μ=1.42），而大型干货船μ=1.93，油轮μ=1.73。相比之下，节点权重分布P（s）在三个子网络中几乎相同，分别为η=1.05，η=1.13，η=1.01.

这些数据给出了不同类型船舶以其独特的移动模式的第一个特征。集装箱船通常遵循按照固定顺序访问多个港口的设定好的时间表，从而提供定期服务。相比之下，干散船似乎不太可预测，因为他们经常根据目前的货物供应和需求在短时间内改变路线。干散船运输网络拥有更多的起点和终点（n=616 个港口，而集装箱船的n=378）解释了给定连接更高德平均度和更少的旅行。油船也遵循短期的市场趋势，但是由于只能装载石油和石油产品，所以可能的目的地（n=505）的数量比干散船的数量更少。

这些差异也被三个网络的中介中心性所强调。尽管一些港口在所有类别中排名靠前（例如苏伊士运河、上海），但其他港口都是专门针对某些船舶类型的。例如，德国的威廉港在世界各地的油轮之间排名第十，但是在干散船中只排名241，集装箱船只排在324。

我们可以通过检查港口的社区结构来进一步了解港口的作用。社区具有许多内部链接的群体，而不同群体之间的链接很少。我们使用针对定向网络的模块化优化方法计算了三个子网络的社区结构（Leicht＆Newman 2008），发现它们与相应的Erdös-Renyi图形的模块性有很大的不同（图3;Guimerà等，2004）。集装箱贸易网络显示12个社区（图3a）。最大的是位于（i）在阿拉伯，亚洲和南非海岸，（ii）在北美东海岸和加勒比海，（iii）在地中海，黑海和欧洲西海岸（ iv）在北欧和（v）在远东和美国西海岸。大宗干货运输显示七组（图3b）。一些可以解释为地理实体（例如北美东海岸，跨太平洋贸易），而其他则分散在多个大陆。特别有趣的是石油运输网络的社区结构，显示六个组（图3c）：（i）欧洲，北非和西非市场，（ii）一个由亚洲，南非和澳大利亚组成的大型社区，（iii）三大西洋市场与委内瑞拉，墨西哥湾，美国东海岸和北欧之间的贸易团体，以及（iv）美洲太平洋海岸。应该指出的是，该网络包括原油运输以及已经成熟的石油产品的商业运输，使得石油生产地区不会以独立的社区出现。这可能是由于模块化优化对较小社区的可检测性的限制（Fortunato＆Barthelemy 2007），但并不影响所揭示的船舶交通社区的相关性。因为根据定义，社区内的运输强度较高，预计同一社区港口之间的生物侵袭性传播将会更重。然而，在图3中，很明显，社区之间没有严格的地理障碍。因此，即使在小型空间尺度上，社区之间的传播也很可能发生在属于不同社区的靠近港口之间的航运或洋流。

尽管三种主要货运船只之间存在差异，但是仍然有一个统一的特征：其主题分布（Milo et al，2002）。像以前的大多数研究一样，我们重点关注三节点主题并计算其归一化Z分数，这是对其网络丰富度的度量（图4）。令人瞩目的是，三种船只的主题分布几乎相同。事实上，其Z分数与万维网和不同的社会网络类似，这些网络被推测形成一个超家族网络（Milo et al。2004）。这个超家族具有许多三元组交互（即，if Embedded Image and Embedded Image then Embedded Image）; 例如，图4中多表示的13个主题中有六个这样的相互作用。不对称的主题，如主题6，是相当罕见的。船舶网络中丰富的传递交互作用表明货物可以直接在港口之间运输，也可以通过几个中间港口运输。因此链路的高聚类和冗余（链路故障的鲁棒性）不仅出现在GCSN中，而且出现在三个子网络中。主题分布于其他人类优化的网络的相似强调了货物贸易在很大程度上取决于人类的互动和信息交流。虽然有利于贸易的鲁棒性，但是以三元组为链接的聚类也具有不必要的副作用：网络的聚类程度越高，外来物种的传播越容易，及时其入侵概率很低（Newman 2003b）。

## Network trajectories

超越网络的角度来看，该数据库还提供了关于每个个体船舶（表1）的移动特征数据。每条船经过的不同港口数量<N>在不同类别的船只间差别不大，但集装箱船的停靠港口次数比干散船和油船更频繁。这些差异可由船舶的特点和运行模式来解释。通常情况下，集装箱船速度快（20-25节），并且在港口进行货物装卸用时更少（平均1.9天）。相比之下，大批量的干散船和油船的移动速度较慢（13-17节），并且在港口停留时间更长（干散船平均5.6天，油船平均4.6天）。

然而，航行速度和货物处理速度并不是唯一的差别。轨迹的拓扑结构也有很大的差异。每种船型的特征样本轨迹如图5a-c所示。在研究的时间段内，集装箱船（图5a）在一些链路上行进多次，而干散船（图5b）几乎通过每个链路一次。油船（图5c）在一些港口之间反复航行多次，但总体来说大部分链路也只通过一次。

我们可以根据“规律性指数”p 来表达这些趋势，量化偏离随机网络的链路使用的频率的多少。考虑一艘船的轨迹航行S次，经过N个不同港口并在L条链路上航行。我们在具有相同数量的节点N和港口S的随机轨迹集合中比较每条链路的平均航行次数Embedded Image 和平均链路使用数Embedded Image 。为了量化实际和随机轨迹之间的差异，我们计算Z分数Embedded Image（其中σ是随机集合中f的标准偏差）。如果p=0，则实际轨迹与随机轨迹无法区分，而p值较大则表示运动轨迹更规律。图5d-f显示了不同类型的船只的规律性指数p的分布。对于集装箱船来说，p围绕p≈2广泛分布，从而支持了我们早期的观察，大多数集装箱船为港口间提供定期服务。另一方面，干散船和油船的绝大部分船只轨迹基本上围绕p=0分布。

## Approximating traffic flows using the gravity model

本文中，我们基于船只的详细到达和离开时间将全球船只运动视为一个网络。知道最近，海运贸易的调查都只能依赖少得多的数据：只有一些主要港口的船只到达数量才会公开，而船舶的实际路径不会公开（Zachcial＆Heideloff 2001）。因此，缺少关于旅行频率的信息，必须由合理的假设取代，重力模型是最受欢迎的选择。一般来说，假定旅行在临近的港口之间比在远的港口之间更有可能。如果dij是港口i和j之间的距离，则因距离增加而相互作用下降的威慑函数表达为f（dij）从i到j的旅程次数则采用Embedded Image，其中Oi是从i港口离开的总数，Ij是到达j港口的总数（Haynes＆Fotheringham 1984）。系数ai 和bj 用来确保Embedded Image和Embedded Image。

重力模型近似的实际船舶交通效果如何？我们为威慑函数选择了一个截断幂率函数，Embedded Image。当β=0.59，k=4900km时模型和实际数据相关性最强。初见数据和模型之间的一致性确实令人印象深刻。关于航行距离分布的预测比简单的非空间模型要好很多，非空间模型保留了旅程的总数，但假定是完全随机的起点和终点。

然而，仔细观察重力模型能够发现其局限性。在图6b中，我们计算有Nij次行程的连接的频率将为Fij。在理想情况下，所有数据点将沿着对角线Fij = Nij对其，但是我们发现数据点基本是分散的。虽然选择了参数β和k来最小化分散，但是数据和模型之间的相关性也只是中等的（肯德尔参数T=0.433）。因为一些原因，每年有几千次的预测是偏离的。

最近的研究有使用重力模型来确定入侵物种传播的中心港口和路线（Drake＆Lodge 2004; Tatem et al。2006）。模型的缺点暴露出了这个问题，该模型有多可靠。为此，我们调查了船舶介导的生物入侵模型，其中链路的权重是观察到的交通流量或者重力模型的流量。

我们遵循以前的流行病学研究（Rvachev＆Longini 1985; Flahault et al。1988; Hufnagel et al。2004; Colizza et al。2006）将网络上的传播视为一种复合种群过程，其中节点上的种群数量动态是通过链路上的运输来实现的。在我们的模型中，船只在两个连续访问的港口之间每次行程能够成功运输一个入侵物种的概率只有很小的ptrans = 1%。运输的也只是原港口物种数量的极少一部分s。在抵达新的港口后，物种经历强烈的人口波动，大多数情况下会导致进口物种死亡。然而，如果新物种赢得了这场“生态轮盘赌”稳住了跟脚，其物种数量P会随着随机逻辑方程Embedded Image 及增长率r=1每年和高斯白噪声ξ迅速增长。细节请见电子辅助材料。

从承载能力P=1的单个港口开始，我们将港口之间的接触建模为具有比率为Nij（经验数据）或者Fij （重力模型）的泊松过程。如图7a所示，重力模型高估了传播速率，并且对于那些连接良好但不在中心枢纽的港口而言，这个差异可能会变得特别显著（图7b）。比较被入侵港口的典型序列，我们发现由实际交通流量驱动的生物入侵往往局限于较小区域内的港口，而在重力模型中，入侵快速到达枢纽。船舶旅行的起点和到达港口的总流入和流出在实际上比模型呈现出更加强的正相关（T=0.157 vs 0.047）。因此，重力模型去除了真实网络中存在的分层结构的太多细节。重力模型消除了大多数的相关性，这一点从简单的参数分析也能得出这个结论（详见电子补充材料）。如果实际网络中的相关性是未知的，没有强相关性将使得重力模型成为合适的零假设，但是近期的一些研究表明，相关性在网络传播过程中起着重要的作用（例如，Boguña＆Pastor-Satorras 2002; Newman2002）。因此，如果相关性已知，则不应该忽视它们。

虽然我们观察到AIS数据的传播速度一直比重力模型慢，即使考虑了不同的参数和人口模型，入侵的时间尺度也不太可预测。假设只有一小部分入侵者在本国栖息地之外取得成功，这样的假设才是现实的（Mack et al。2000）。此外，我们对模型中的参数进行了调整，使得每次每条船能够带来入侵的概率大约为4.4\*104 这是Drake＆Lodge（2004）指出的经验值。尽管如此，根据经验，通过足够的精度来确定各个参数也不足以给出超过定性的结果。特别难的是预测潜在入侵者如何对特定地点的环境条件作出反应。根据入侵生物栖息条件的要求，港口之间的增长率肯定会有所不同，这取决于温度或者盐度等因素。因此，我们的结构应该被视作许多不同可能的情形之一。目前我们正在对基于GCSN的生物侵袭风险进行更加详细的研究（H. Seebens＆B. Blasius 2009，未发表的数据）。

## Conclusion

我们已经提出了基于AIS记录的船舶运动研究。将港口视作网络中的节点，通过船舶航行来连接，我们发现全球船舶航行像近年来调查的其他复杂网络一样，拥有小世界属性，也有广泛的度和权重分布。其他的特点，如运河的重要性和将港口分为区域聚类，更具有航船业的特点。该网络的一个重要特征是不同船舶类型的运动模式的差异。干散船和油船往往在港口之间以较不规则的方式运动，而集装箱船则显得更加可预测。这是入侵物种传播的一个重要结论，因为大量干散船和油船经常空船航行，因此会交换大量的压载水。重力模型是预测海洋生物入侵的传统方法，捕捉了一些全球贸易的广泛趋势，但对许多应用而言，其结果太过粗糙。未来遏制物种入侵的策略将不得不考虑更多细节。本文中介绍的网络结构可以看做是朝这个方向迈出的第一步。

## Acknowledgements

我们感谢B. Volk，K.H. Holocher，A.Wilkinson，J.M.Derake和H.Rosenthal激发讨论和他们有益的建议。我们也感谢劳埃德注册的Fairplay公司提供他们的船舶数据库。这项工作获得了德国大众基金会和BMBF的支持。

## References

[1] R. Albert and A.-L. Barab´asi. Statistical mechanics of complex networks. Reviews of Modern

Physics, 74:47–97, 2002.

[2] A.-L. Barab´asi and R. Albert. Emergence of scaling in random networks. Science, 286:509–512,

1999.

[3] A. Barrat, M. Barth´elemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex

weighted networks. Proc. Nat. Acad. Sci., 101:3747–3752, 2004.

[4] M. Barth´elemy and A. Flammini. Modeling urban street patterns. Physical Review Letters,

100:138702, 2008.

[5] M. Bogu˜na and R. Pastor-Satorras. Epidemic spreading in correlated complex networks.

Physical Review E, 66:047104, 2002.

[6] J. Buhl, J. Gautrais, N. Reeves, R. V. Sol´e, S. Valverde, P. Kuntz, and G. Theraulaz. Topological patterns in street networks of self-organized urban settlements. European Physical

Journal B, 49:513–522, 2006.

[7] K. P. Burnham and D. R. Anderson. Model selection and multimodel inference. A practical

information-theoretic approach. Springer, New York, 1998.

[8] J. T. Carlton and J. B. Geller. Ecological roulette: the global transport of nonindigenous

marine organisms. Science, 261:78–82, 1993.

[9] A. Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data.

SIAM Rev., page (in press), 2009.

[10] A. N. Cohen and J. T. Carlton. Accelerating invasion rate in a highly invaded estuary. Science,

279:555–558, 1998.

[11] V. Colizza, A. Barrat, M. Barth´elemy, and A. Vespignani. The role of the airline transportation

network in the prediction and predictability of global epidemics. Proc. Nat. Acad. Sci., 103:

2015–2020, 2006.

[12] J. M. Drake and D. M. Lodge. Global hot spots of biological invasions: evaluating options for

ballast-water management. Proc. Roy. Soc. Lond. B, 271:575–580, 2004.

[13] J. M. Drake and D. M. Lodge. Hull fouling is a risk factor for intercontinental species exchange in aquatic ecosystems. Aquatic Invasions, 2:121–131, 2007.

[14] A. Flahault, S. Letrait, P. Blin, S. Hazout, J. M´enar´es, and A. J. Valleron. Modelling the

1985 influenza epidemic in France. Statistics in Medicine, 7:1147–1155, 1988.

[15] S. Fortunato and M. Barthelemy. Resolution limit in community detection. Proc. Nat. Acad.

Sci., 104:36–41, 2007.

[16] L. C. Freeman. Centrality in social networks I: Conceptual clarification. Social Networks, 1:

215–239, 1979.

[17] T. Gross and B. Blasius. Adaptive coevolutionary networks: a review. Journal of the Royal

Society Interface, 5:259–271, 2008.

[18] R. Guimer`a and L. A. N. Amaral. Modeling the world-wide airport network. European Physical

Journal B, 38:381–385, 2004.

[19] R. Guimer`a, M. Sales-Pardo, and L. A. N. Amaral. Modularity from fluctuations in random

graphs and complex networks. Phys. Rev. E, 70:025101(R), 2004.

[20] R. Guimer`a, S. Mossa, A. Turtschi, and L. A. N. Amaral. The worldwide air transportation

network: Anomalous centrality, community structure, and cities’ global roles. Proc. Nat.

Acad. Sci., 102:7794–7799, 2005.

[21] K. E. Haynes and A. S. Fotheringham. Gravity and spatial interaction models. Sage, Beverly

Hills, 1984.

[22] Y. Hu and D. Zhu. Empirical analysis of the worldwide maritime transportation network.

Physica A, 388:2061–2071, 2009.

[23] L. Hufnagel, D. Brockmann, and T. Geisel. Forecast and control of epidemics in a globalized

world. Proc. Nat. Acad. Sci., 101:15124–15129, 2004.

[24] International Maritime Organization. International shipping and world trade. Facts and figures. Available at http://www.imo.org/, 2006.

[25] V. Latora and M. Marchiori. Is the Boston subway a small-world network? Physica A, 314:

109–113, 2002.

[26] E. A. Leicht and M. E. J. Newman. Community structure in directed networks. Physical

Review Letters, 100:118703, 2008.

[27] L. P. Lounibos. Invasions by insect vectors of human disease. Annual Review of Entomology,

47:233–266, 2002.

[28] R. N. Mack, D. Simberloff, W. M. Lonsdale, H. Evans, M. Clout, and F. A. Bazzaz. Biotic invasions: causes, epidemiology, global consequences, and control. Ecological Applications, 10:

689–710, 2000.

[29] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklo vskii, and U. Alon. Network motifs:

simple building blocks of complex networks. Science, 298:824–827, 2002.

[30] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and

U. Alon. Superfamilies of evolved and designed networks. Science, 303:1538–1542, 2004.

[31] M. E. J. Newman. Assortative mixing in networks. Physical Review Letters, 89:208701, 2002.

[32] M. E. J. Newman. The structure and function of complex networks. SIAM Review, 45:167–256,

2003a.

[33] M. E. J. Newman. Properties of highly clustered networks. Physical Review E, 68:026121,

2003b.

[34] M. E. J. Newman. Who is the best connected scientist? A study of scientific coauthorship

networks. In E. Ben-Naim, H. Frauenfelder, and Z. Toroczkai, editors, Complex Networks,

pages 337–370. Springer, Berlin, 2004.

[35] M. E. J. Newman. Power laws, Pareto distributions and Zipf ’s law. Cont. Physics, 46:323–351,

2005.

[36] T. E. Notteboom. Container shipping and ports: An overview. Review of Network Economics

3:86–106, 2004.

[37] R. Pastor-Satorras and A. Vespignani. Epidemic Spreading in Scale-Free Networks. Phys.

Rev. Lett., 86:3200–3203, 2001.

[38] D. Pimentel, R. Zuniga, and D. Morrison. Update on the environmental costs associated with

alien-invasive species in the United States. Ecological Economics, 52:273–288, 2005.

[39] R. Robinson. Asian hub/feeder nets: the dynamics of restructuring. Maritime Policy &

Management, 25:21–40, 1998.

[40] J.-P. Rodrigue, C. Comtois, and B. Slack. The Geography of Transport Systems. Routledge,

London, 2006.

[41] G. M. Ruiz, T. K. Rawlings, F. C. Dobbs, L. A. Drake, T. Mullady, A. Huq, and R. R. Colwell.

Global spread of microorganisms by ships. Nature, 408:49–50, 2000.

[42] L. A. Rvachev and I. M. Longini, Jr. A mathematical model for the global spread of influenza.

Mathematical Biosciences, 75:3–22, 1985.

[43] H. Seebens and B. Blasius. Modeling marine bioinvasion by global shipping: a probabilistic approach. Unpublished, 2009.

[44] P. Sen, S. Dasgupta, A. Chatterjee, P. A. Sreeram, G. Mukherjee, and S. S. Manna. Smallworld properties of the Indian railway network. Physical Review E, 67:036106, 2003.

[45] A. J. Tatem, S. I. Hay, and D. J. Rogers. Global traffic and disease vector dispersal. Proc.

Nat. Acad. Sci., 103:6242–6247, 2006.

[46] United Nations conference on trade and development. Review of maritime transport. Available

at http://www.unctad.org/en/docs/rmt2007 en.pdf, 2007.

[47] D. I. Warton, I. J. Wright, D. S. Falster, and M. Westoby. Bivariate line-fitting methods for

allometry. Biological Reviews, 81:259–291, 2006.

[48] D. J. Watts and S. H. Strogatz. Collective dynamics of ’small-world’ networks. Nature, 393:

440–442, 1998.

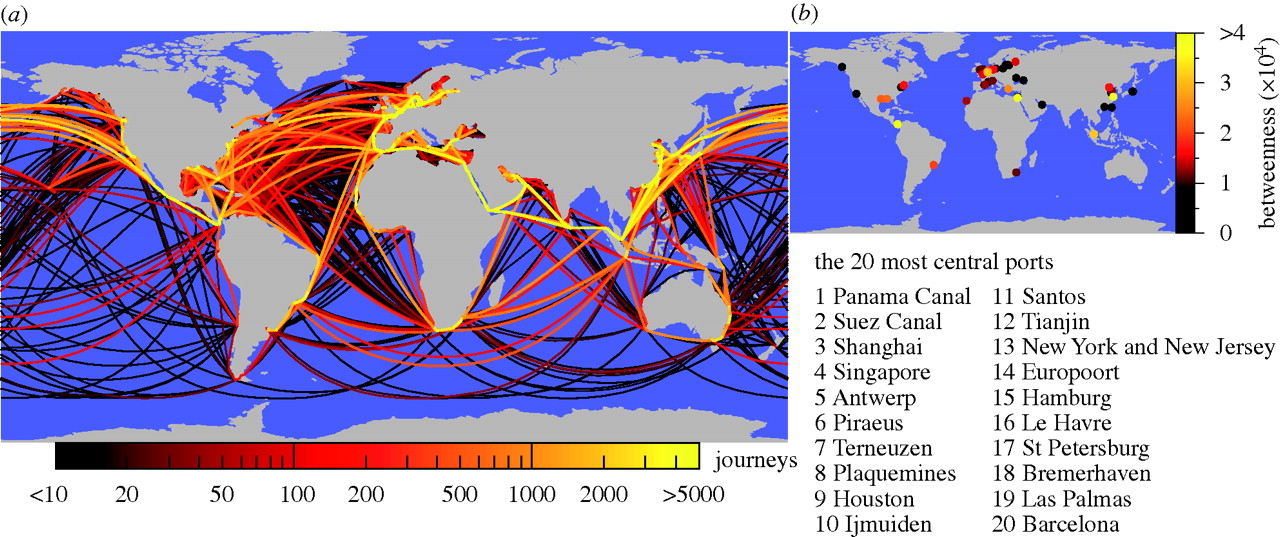
[49] T. Wei, G. Deng, and P. Wu. Analysis of network effect in port and shipping system characterized by scale-free network. In ISKE-2007 Proceedings, Amsterdam, 2007. Atlantic Press.

[50] M. Williamson. Biological Invasions. Chapman and Hall, London, 1996.

[51] M. Zachcial and C. Heideloff. ISL shipping statistics yearbook 2001. Technical report, Institute

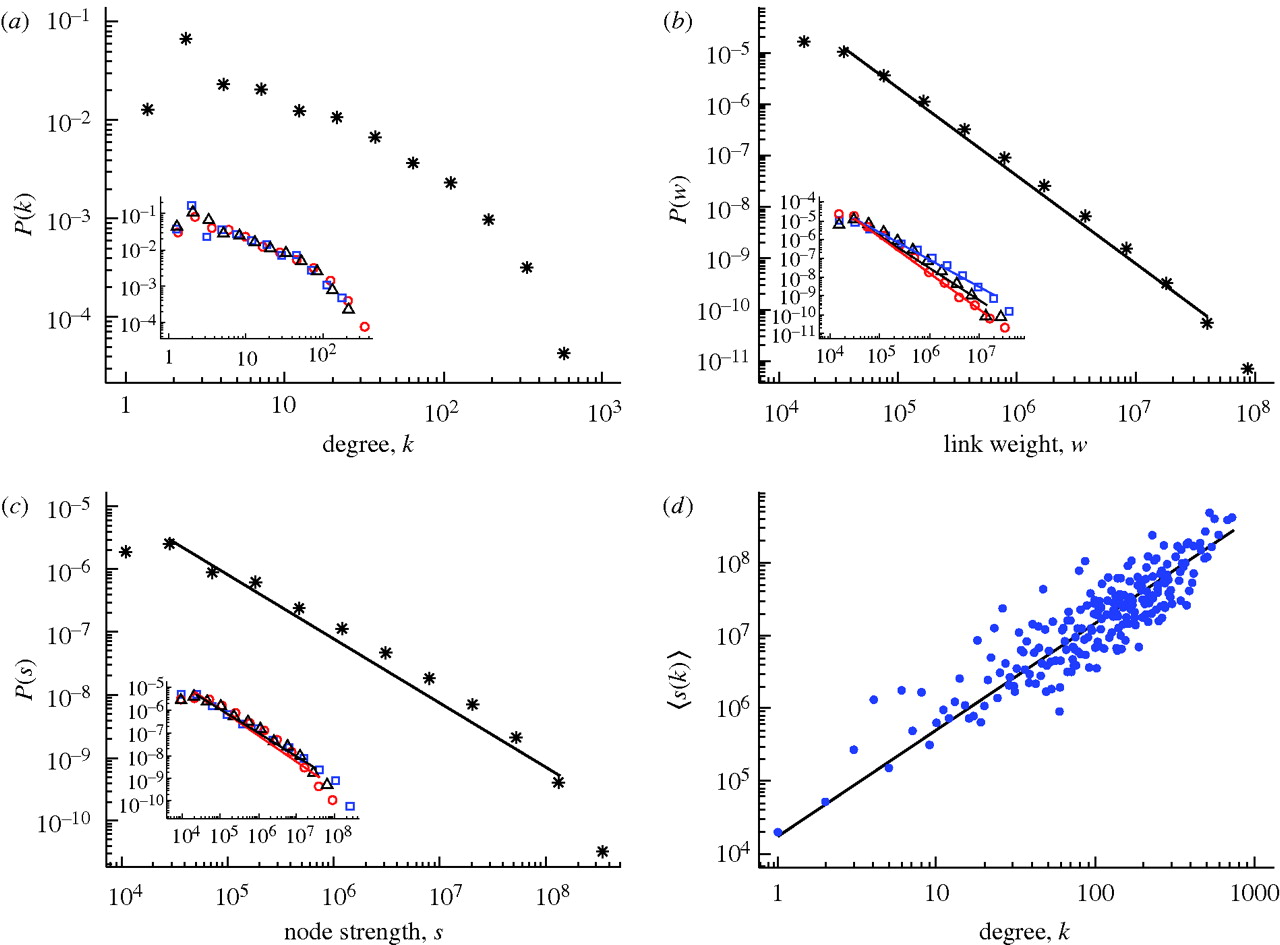
of Shipping Economics and Logistics, Bremen, 2001.

## Pictures and tables



**图1.**

GCSN中的路线，港口和中介中心。 （a）2007年期间所有货船大于10 000 GT的轨迹。色标表示每条路线的行程数。 假定船舶沿水上最短（测地线）的路线行驶。 （b）最高中心性的50个港口和20个最中心港口的排名列表。



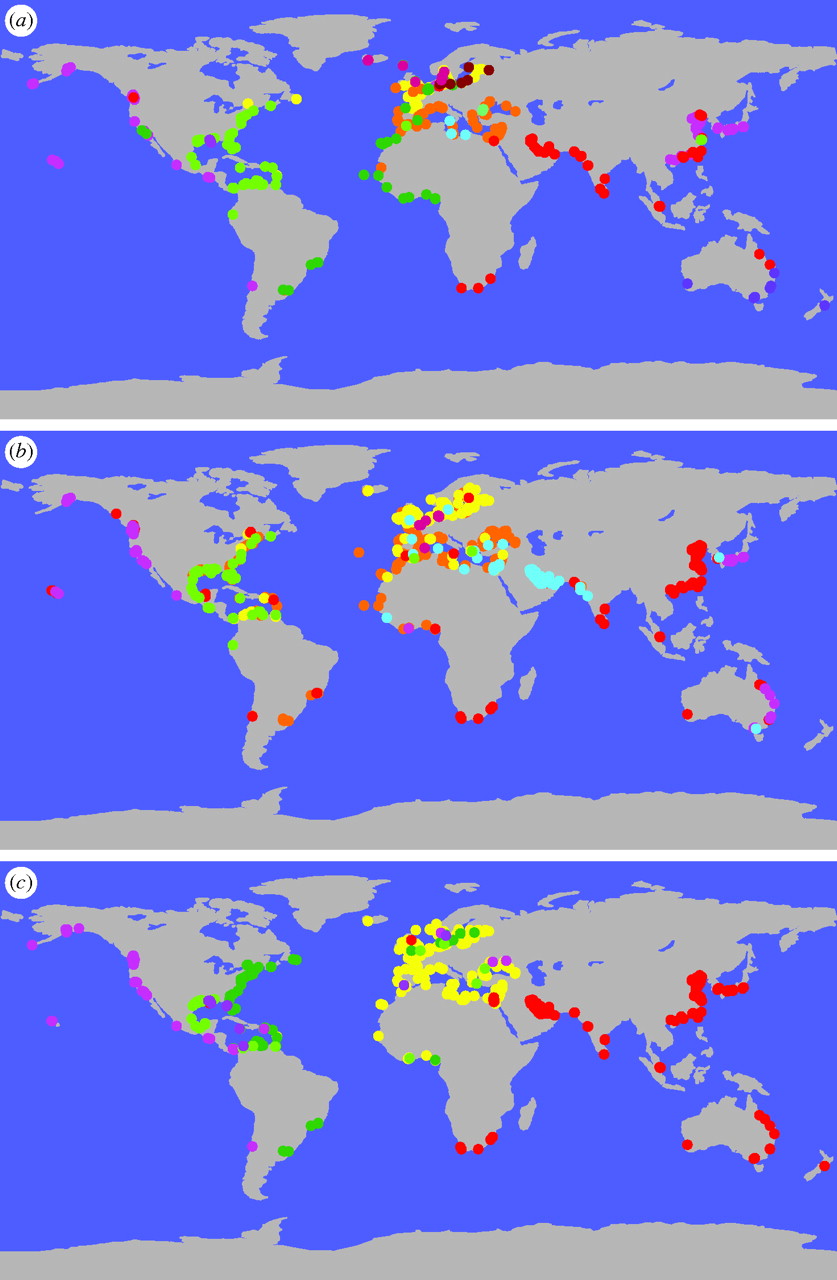
**图 2.**

GCSN的度和权重。（a）度分布P（k）对于GCSN及其子网都是右倾的，而不是幂律分布。此处的k度定义为入射角和外倾角之和，因此k = 1是相当罕见的（星号，所有船舶类型;方形，集装箱船;圆形，散装干运输车;三角形，油轮）。 （b）链路权重分布P（w）显示了GCSN和三个子网络的明确的幂律关系，其中不同船舶类型的运动模式的特征μ（星号，μ= 1.71±0.14;方形，μ= 1.42±0.15;圆，μ= 1.93±0.11;三角形，μ= 1.73±0.25）。 （c）节点强度分布P（s）也是重尾，显示幂律关系。通过线性回归计算95％置信区间（通过最大似然估计得到类似结果，见电子补充材料）（星号，η= 1.02±0.17;方形，η= 1.05±0.13;圆，η= 1.13 ±0.21;三角形，η= 1.01±0.16）。 （d）节点<s（k）的平均强度>其程度<s（k）>αk1.46±0.1超线性地表示，表示高度连接的端口平均具有较高权重的链路。

**表1.**

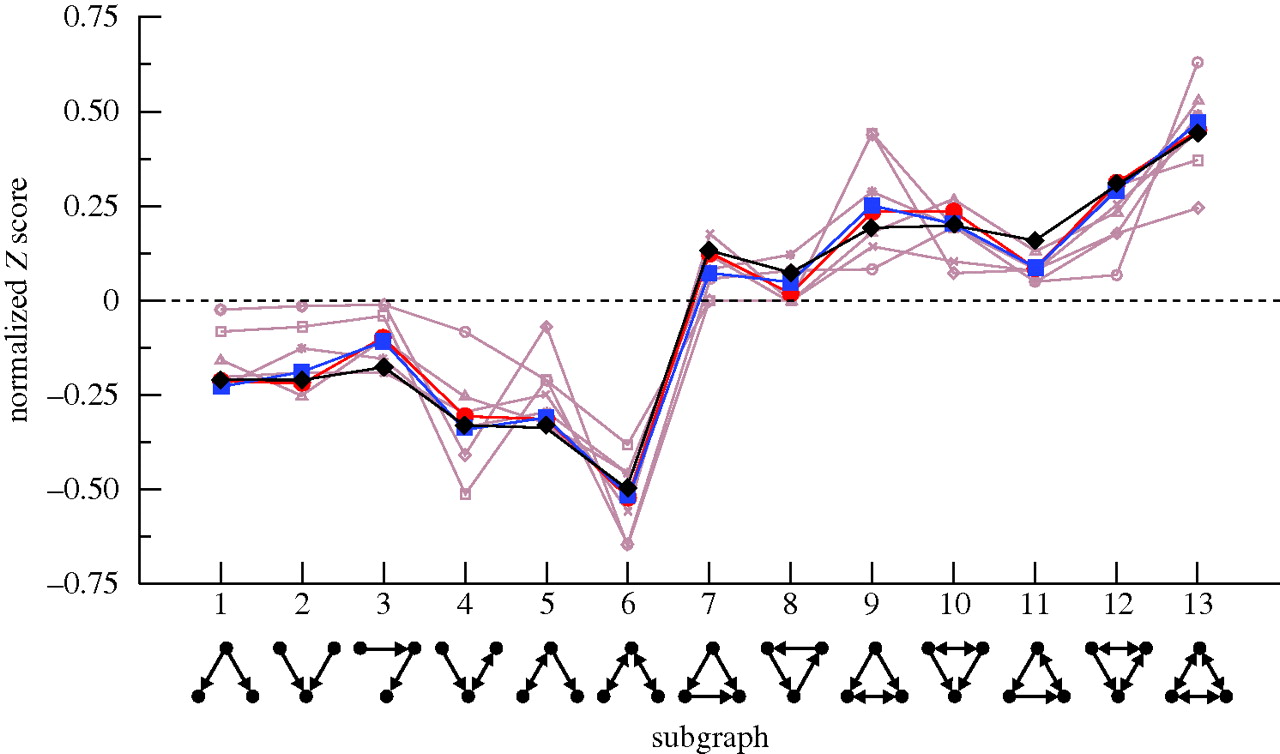
不同子网的表征。 每个子网中的船数，总GT（106 GT）和港口数n; 连同网络特征：平均度<k>，聚类系数C，平均最短路径长度<l>，每个链路的平均行程<J>，幂律指数μ和η; 和轨迹属性：不同港口的平均数量<N>，链接<L>，每个船舶的访问港口<S>和规则性索引<p>。 一些值得注意的值以粗体突出显示。

| **ship class** | **ships** | **MGT** | ***n*** | **〈*k*〉** | ***C*** | **〈*l*〉** | **〈*J*〉** | ***μ*** | ***η*** | **〈*N*〉** | **〈*L*〉** | **〈*S*〉** | **〈*p*〉** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| whole fleet | 16 363 | 664.7 | 951 | **76.4** | 0.49 | 2.5 | 13.57 | 1.71 | 1.02 | 10.4 | 15.6 | 31.8 | 0.63 |
| container ships | 3100 | 116.8 | **378** | 32.4 | 0.52 | 2.76 | **24.25** | **1.42** | 1.05 | 11.2 | **21.2** | **48.9** | **1.84** |
| bulk dry carriers | 5498 | 196.8 | 616 | 44.6 | 0.43 | 2.57 | 4.65 | 1.93 | 1.13 | 8.9 | 10.4 | 12.2 | 0.03 |
| oil tankers | 2628 | 178.4 | 505 | 33.3 | 0.44 | 2.74 | 5.07 | 1.73 | 1.01 | 9.2 | 12.9 | 17.7 | 0.19 |



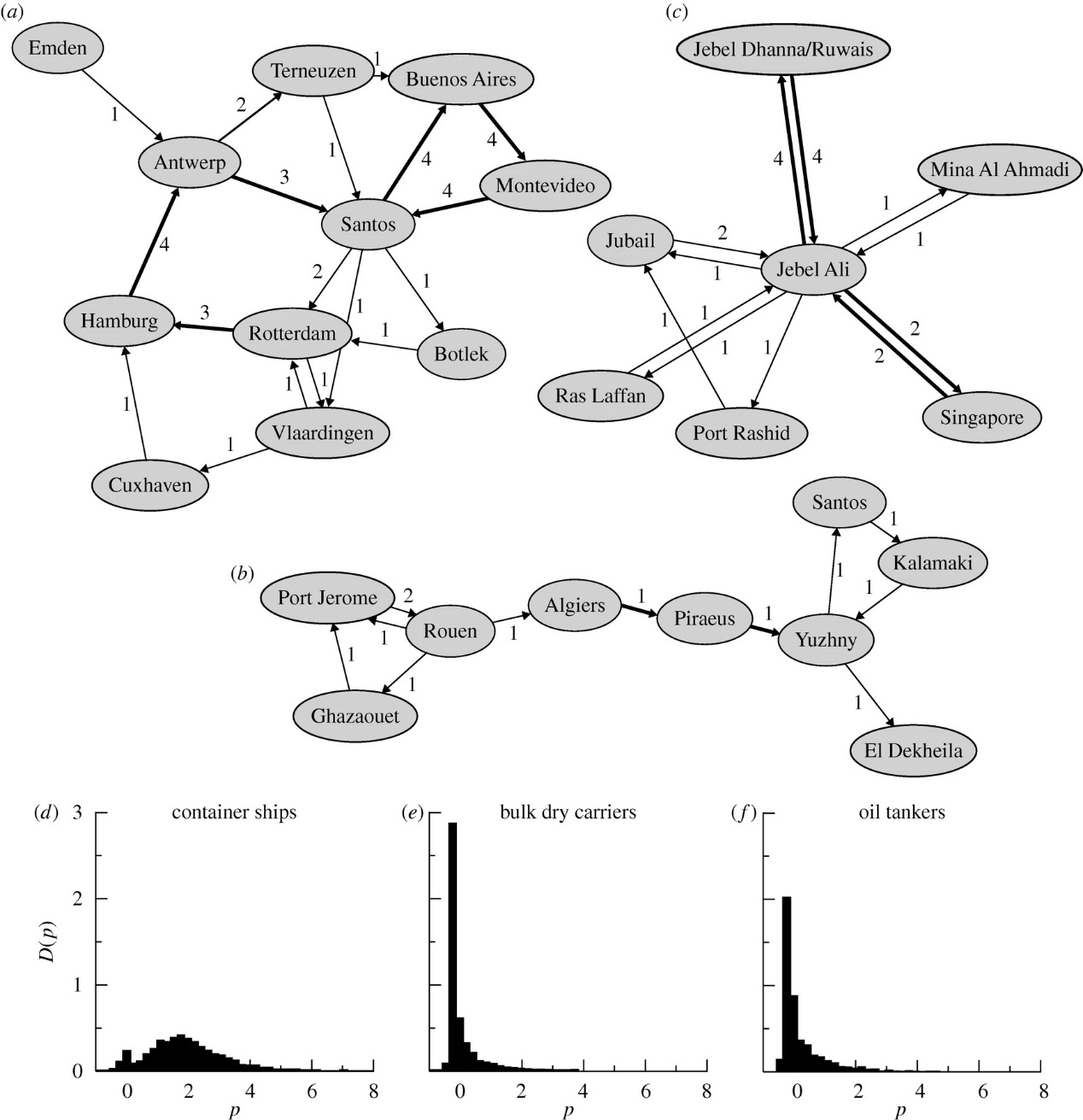
**图 3.**

**港口在三个货船子网络的社区。 社区是港口的集群，可以最大限度地提高组内的链接数量，而不是群体之间的模块化Q（Leicht＆Newman 2008）。 在每张地图中，不同颜色代表不同交易社区的货物运输，其中（a）集装箱船（c = 12，Q = 0.605），（b）散货船（c = 7，Q = 0.592）和（c ）油罐车（c = 6，Q = 0.716）。 被检查网络的所有模块Q与Erdös-Renyi图中相同大小和链接数的图形显着不同（Guimerà等，2004）。 对应于（a），（b）和（c）的网络，值分别为QER = 0.219，QER = 0.182和QER = 0.220。**

****

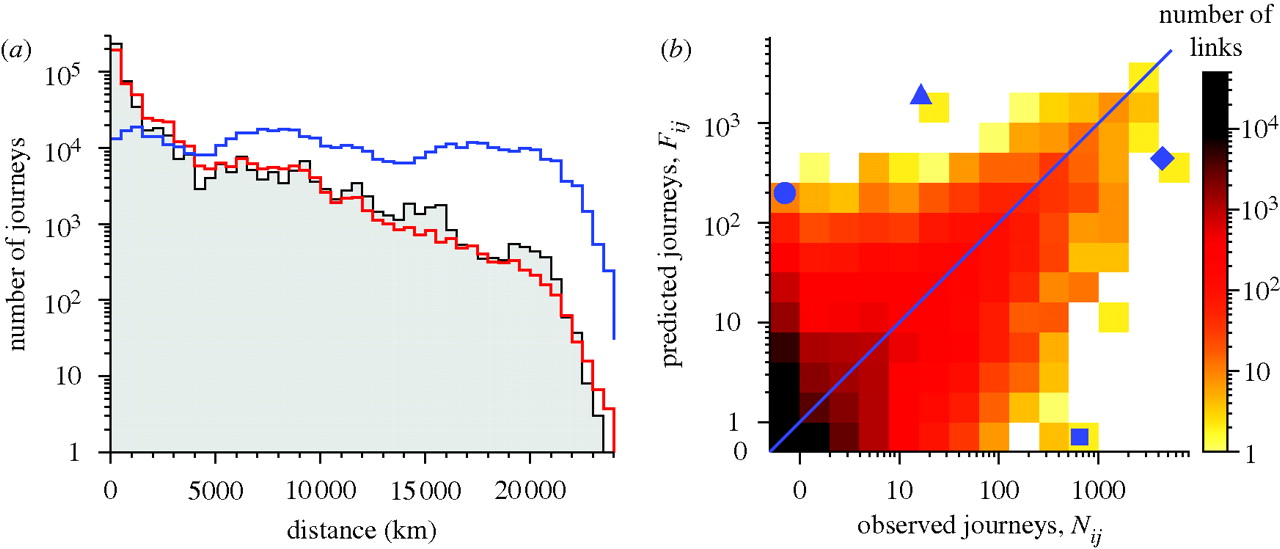
**图 4.**

三主要货运船队的主题分布。 正（负）归一化Z分数表示在真实网络中，主题比具有相同程度序列的随机网络中的主题更多（更少）频繁。 为了比较，我们覆盖了万维网和社交网络的Z分数。 该协议表明，船舶网络属于同一个网络超家族（Milo et al。2004）。 即使25,55和75％的最弱连接被去除，船队的主体分布也保持不变。 线上有开放的圈，www-1; 线与方块，www-2; 线和菱形，www-3;线和三角形，社会-1; 线与十字，社会2; 线带星号，社交3; 线和红色圆圈，散装干运输车; 线配蓝色方块，集装箱船; 线和黑色菱形，油罐车。

****

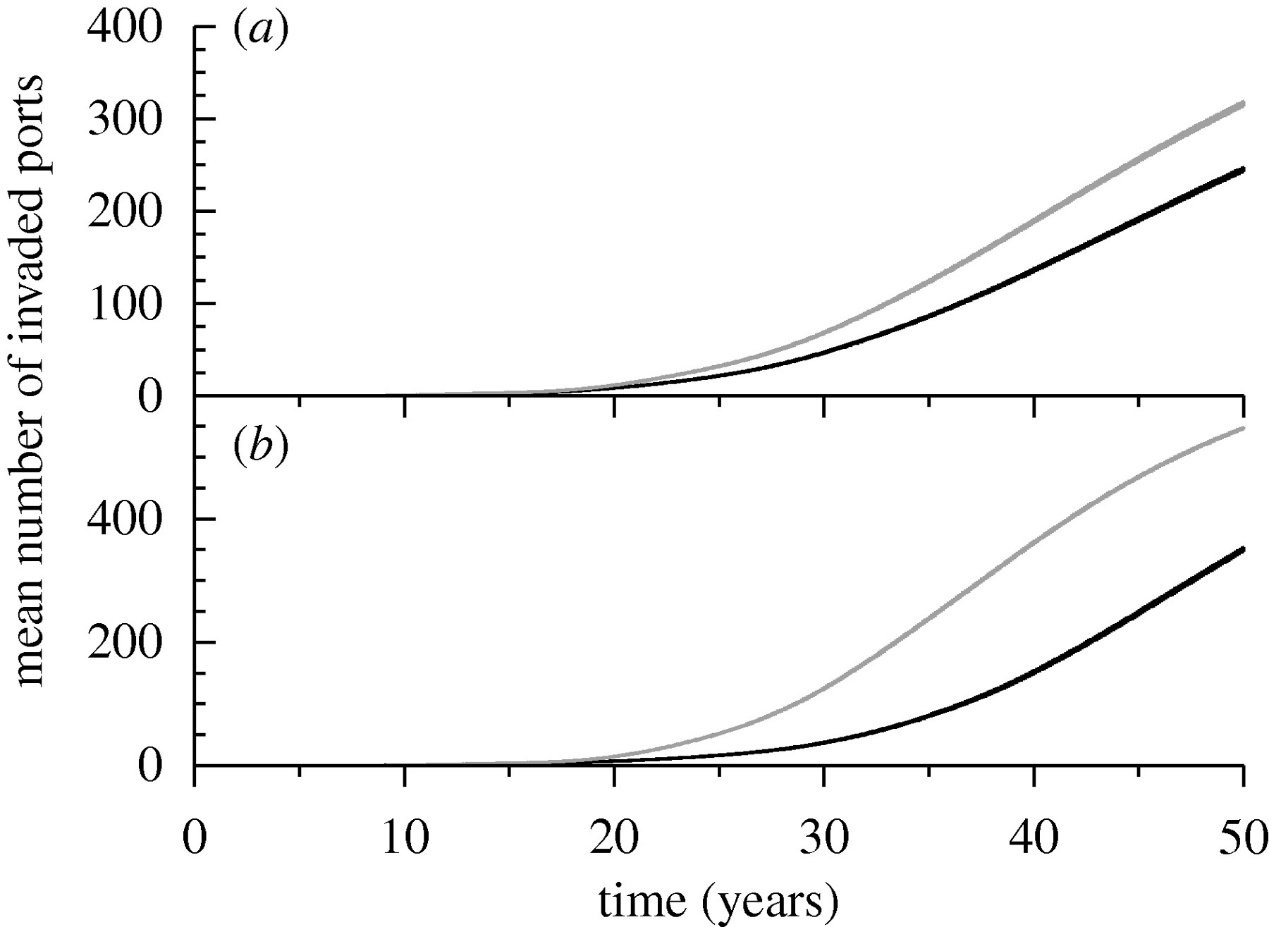
**图 5.**

（a）具有规则指数p = 2.09的集装箱船的样本轨迹，（b）散装干运输船，p = 0.098和（c）油罐车，p = 1.027。 在三个轨迹中，数字和线条厚度指示每个路线上的行程频率。 （d-f）三个主要船舶的p的分布。

****

**图 6.**

（a）GCSN（可通航的距离在大陆周围如图1所示）中的港口到港口距离的直方图。 我们重叠两个不同模型的预测。 基于关于港口和总港口访问之间的距离的信息的引力模型（红色）比简单的模型（蓝色）更适合，其仅修复总行程数（黑线，观察到的；红线，重力模型 ;蓝线，随机流量）。 （b）与Nij观察到的港口对数量和Fij预测的旅程。 流量Fij用重力模型计算（四舍五入到最接近的整数）。 一些最差的异常值以蓝色突出显示。 圆形，安特卫普到加来（Nij = 0 vs Fij = 200）; 三角形，荷兰Hook到欧罗波特（16对1895）; 菱形，加来到多佛（4392对443）; 方形，哈里奇到荷兰的胡克（644对0）。

****

**图 7.**

港口入侵物种传播随机人口模型的结果。 （a）入侵从一个单一随机选择的港口开始。 （b）初始港口作为卑尔根（挪威）固定，这是一个连接良好的港口（度k = 49）的例子，不在中央枢纽之间。从i到j每年的旅程被假定为Nij（来自GCSN的实际流量）或Fij（重力模型）。每个旅程的一小部分物种从原地运送到目的地的可能性很小。调整参数（每年r = 1，ptrans = 0.01和s = 4×10-5），以产生大约4.4×10-4的每次访问的入侵概率（Drake＆Lodge 2004;见电子补充材料详情）。绘制的是（a）14 000和（b）1000次模拟运行（标准误差等于线宽度）的入侵港口（人口数大于承载能力的一半）的累积数量。黑线，实际交通流量;灰线，重力模型。