

# 上证A股市场的复杂网络特性研究

惠宝锋<sup>1</sup>, 葛志远<sup>2</sup>, 王咏宁<sup>1</sup>

(1.青海民族大学物理与电子信息工程学院, 青海 西宁 810007;

2.北京工业大学, 北京 100124)

**摘要:** 本文利用UCINET社会网络绘图软件, 通过对股票市场网络进行了构建, 发现当相似度 $\rho=0.85$ 时, 呈现清晰的股票网络结构。进而针对这一现象通过度的测量与校验, 运用复杂网络方法, 对复杂网络方法联系股价相似度进行分析。在研究的基础与实验中发现, 上海证券市场的股票网络呈现出明显的无标度特性、根据股票特有的幂分布特征体现出股票之间的相关关系, 表明在以股票为代表的网络中体现出一部分股票, 对上证A股市场中占据着重要影响力。

**关键词:** 相关系数; 复杂网络; 股票市场; 优势股票

**中图分类号:** TP391 **文献标识码:** A

## Study on the Characteristics of the Complex Network of Shanghai A-Share Market

HUI Baofeng<sup>1</sup>, GE Zhiyuan<sup>2</sup>, WANG Yongning<sup>1</sup>

(1. School of Physics and Electronic Information Engineering, Qinghai University for Nationalities, Xining 810007, China;

2. Beijing University of Technology, Beijing 100124, China)

**Abstract:** Through the construction of the stock market network with UCINET software, the study finds that a clear structure of the stock network is showed when the similarity degree is 0.85 ( $\rho=0.85$ ). According to the measurement and verification of the phenomenon, the complex network method is applied to analyze the similarity of the stock price with the complex network method. The study shows that the Shanghai stock market network presents the scale-free characteristic, and the correlations between stocks also reflect the degrees of power law distribution, which indicates that a small part of shares in the network perform an important influence in the stock market.

**Keywords:** correlation coefficient; complex network; stock market; preferable share

### 1 引言(Introduction)

复杂网络中主要是由节点与连接节点之间的边而组成的网络结构, 节点之间和边与节点分别体现了真实系统中的不同个体和个体之间的相互关系<sup>[1]</sup>。金融网络是在具体的经济活动中体现出的由成千上万个账户节点所组成的复杂社会网络体系, 在社会现代化中的一个不可或缺的组成部分就是股票市场, 确定了股票市场中的网络化与网络的密切关系, 网络的许多特性也在股票市场中体现出来。在人们不同理性、偏好、利益及信息的不对称, 使得股票系统复杂特性在人与环境中对系统复杂性的研究富有了挑战性。用计算机技术模拟股票网络, 探索股票走势及预测股票价格, 成为一种新的途径与方法<sup>[2]</sup>。

当下, 应用最小生成法的进行国外股票市场为研究对象已成为主流<sup>[3]</sup>, 进而来对相关股市中股票价格的相关系数所表现出的规律性, 来进行相对应的解释和预测<sup>[4,5]</sup>。Boginski等对美国证券市场的六千多只股票进行研究, 分析发现股票

市场的收盘价格相关性分布呈现出无标度分布的规律<sup>[6]</sup>。Kim等人以股票价格的影响力作为研究对象, 发现股票价格影响力的大小呈现出无标度特性, 对网络起到有影响力的股票只是少数股票<sup>[7]</sup>。L. Bakker以股票价格的变动对股市投资行为的影响力大小作为研究对象, 结果发现现实生活中的信任网络, 可以使得价格网络更加稳定<sup>[8]</sup>。

本文通过复杂网络理论引入股票市场的研究分析中, 对上证A股在工业、商业、房地产、金融及公共事业领域中选择2015年至2017年上证股票的相关数据作对象进行分析, 利用UCINET中的编程技术, 以及统计学的方法, 试图从复杂网络的角度, 来揭示上证A股市场的网络结构特征和分布规律, 通过节点与边的规律进行分析具有影响力的股票市场特征研究。

### 2 股票市场网络的构建(Construction of the stock market network)

通过每一个股票代码与之相关的股票利用节点与边的方

法将股票市场绘制成一个网络，其中每一个股票看作节点，利用股票代码作为节点名称，再将与之存在关系的股票代码进行边的连接表示之间的关系。在具体分析中两支股票代码间的相似程度用一个值  $\rho$  来表示，如果两支股票之间的相似程度特别高时，则表示这些股票存在一定的相关关系，在股票市场中股票间存在一个相关系数，而这个系数就是股票之间的相近程度。我们一般利用相关性来度量节点之间的相似程度，选值在[0,1]区间，数值越大表示两者之间相关性越强。

当股票  $x_i$  与  $x_j$  相似程度  $c_{ij} \geq \rho$  时，我们将  $c_{ij}$  称为两支股票间的相互关系，当  $\rho$  值不断变大时，股票之间的相互关系也就越高<sup>[9]</sup>。

我们参照香港理工大学Chi K.Tse和Francis C.M. Lau 等人的研究<sup>[9]</sup>，把股票  $i$  与  $j$  间的相关系数  $c_{ij}$  定义为：

$$c_{ij} = \frac{\sum_t [(x_i(t) - \bar{x}_i)(x_j(t) - \bar{x}_j)]}{\sqrt{\sum_t (x_i(t) - \bar{x}_i)^2} \sqrt{\sum_t (x_j(t) - \bar{x}_j)^2}} \quad (1)$$

其中， $\Sigma$  表示所选时期内相应数值的求和， $x_i(t)$  和  $x_j(j)$  表示股票  $i$  和  $j$  在  $t$  天的股票价格， $\bar{x}_i$  和  $\bar{x}_j$  为股票  $i$  和  $j$  在所选的时间段中的价格平均数。

在2015年1月至2017年12月期间，我们把研究对象锁定为上证A股在工业、商业、房地产、金融及公共事业领域中的共一千多支股票的周收盘价格。

本文通过股票价格的相关系数  $c_{ij}$  搭建了基于股票价格的复杂网络。对于上证A股股票网络中通过股票  $i$  与股票  $j$  的关系和股票  $j$  与股票  $i$  的关系是一致的情况，在进行  $\rho$  的相对设定后，发现具有典型代表的股票之间当  $c_{ij} \geq \rho$  时，表明在现实网络中两者存在相关关系，我们用数字1进行表示，当  $c_{ij} < \rho$  时，表明两者之间不存在相关关系，用数字0进行表示。

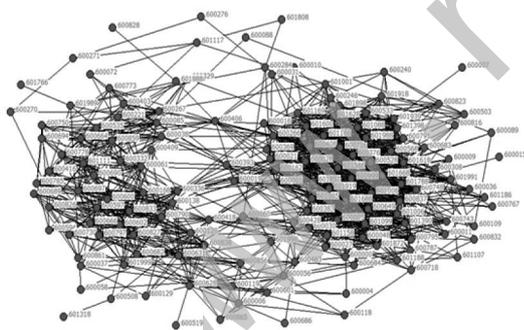


图1 股票网络结构图

Fig.1 Stock network structure diagram

我们将相关系数矩阵导入UCINET进行转化后，得到整个网络结构图。在  $\rho$  值不同的情况下，我们用SPSS软件分别统计了股票节点间的连接数。在实际实验中发现当网络图形处于完全连接时，图形中的节点将会聚集在一起，形成一个点，这时的  $\rho$  值接近或直接等于0。在  $\rho$  不断变化的过程中，体现出当  $\rho$  不断变小时，网络中的股票节点变得杂乱起来，没有任何规律，图形也表现除了离散的状态。反之，随着  $\rho$  的不断变大，原来离散的点逐渐发生集聚，而且网络图中的节点数明显增加，在  $\rho$  值接近0.6时已经表现出了较清晰的图形，当取值到达0.85时，网络图形已经可以清晰的表现出

来，而且彼此相互连通。我们利用UCTNET社会网络绘图工具，在  $\rho=0.85$  取值状态下绘制了基于周收盘价格的股票复杂网络图形。如图1所示。

### 3 复杂网络参数的计算与误差检验(Complex network parameter error test)

进行复杂网络具体实验中，我们首先进行复杂网络的参数设定，股票节点数(Node)、边连接数(Link)、平均最短长度(Short)、网络直径(Dia)、平均聚类系数(Correlation)和平均度(K)。将 在不同取值情况中,对以上股票复杂网络数据进行了参数测量。测量结果如表1所示。

表1 网络参数数据( $\rho$  取值)

Tab.1 Parameters( $\rho$  value) data network

参数	$\rho=0.80$	$\rho=0.85$	$\rho=0.90$	$\rho=0.95$
节点数N	154	154	154	154
连接数L	2138	1376	570	194
平均最短长度S	13.50	10.88	5.27	2.82
直径D	14	17	26	34
平均聚类系数C	0.53	0.47	0.34	0.18
平均度K	24.63	16	6.87	1.18

综合分析表1，发现在具体网络参数中股票之间的相关程度要是越高的那部分股票，其小世界效应就越为凸显和明显，体现为当相似度在取值分别为0.80和0.85时，网络数据中的最短长度都大于10，小世界特征表现的不突出和明显。但在相似度增大的过程中发现当  $\rho \geq 0.9$  时，就表现出了明显的小世界特性，其网络数据中体现出的最短长度分别为5.2和2.8，同时也表现出了这些股票之间存在高度的相关关系。

在分析股票网络节点与股票之间的相似度变化的关系上，绘制了网络关系图进行观察各个节点度数值分布规律，从图像中明显可以看出以下特点：(1)相似度  $\rho=0.85$  时，图像中的股票节点主要分布在[0,40]区间度数。(2)相似度  $\rho=0.90$  时，图像中的股票节点主要分布在[0,30]区间度数。(3)相似度  $\rho=0.95$  时，图像中的股票节点主要分布在[0,5]区间度数，只有少部分散落在5之外，表明节点的分布十分明显。(4)对网络节点进行相关内容的统计分析，发现我们研究的所有上证A股的比例在10%以下。

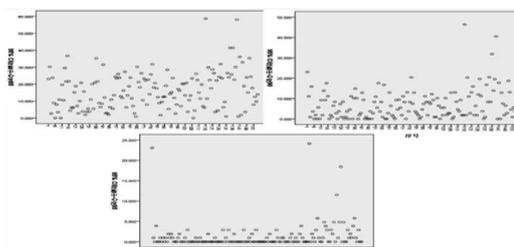


图2 节点度数值k与  $\rho$  值变化的关系图

Fig.2 Relationship between  $k$  and the degree of  $\rho$  node

从数学的概率密度函数角度出发， $k$  的分布就可以看作是一个概率密度函数。这样使得  $p(k)$  成为一个节点  $k$  有多少度的概率，我们就可以通过分布函数的方法对上证A股的周价格网络进行分析它们是怎样进行连接的，再将  $k$  的概率分布与其本身的数值绘制成为一条直线，其特征指数用具体的梯度进行

必要的定义。分别就相似度在0.85和0.95的情况下分别绘制k的分布图。结果如图3所示。

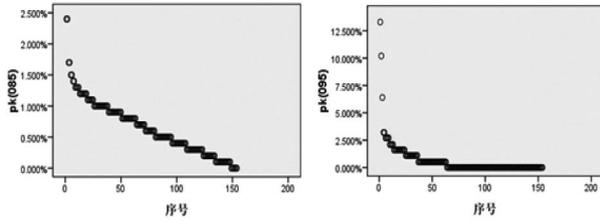


图3  $p(k)$ 对 $k$ 的分布图

Fig.3 Distribution of  $p(k)$  to  $k$

基于图3，我们对 $p(k)$ 和 $k$ 的分布按照幂律分布 $p(k) = ae^{-\gamma k}$ 进行拟合。根据表1中的数据，我们利用最小二乘法对不同相似度的幂律指数和拟合误差，计算结果如表2所示。

表2 幂律分布指数和参数误差值( $\rho$ 取值)

Tab.2 Power-law distribution index and parameter errors( $\rho$  values)

参数	0.80	0.85	0.90	0.95
幂律分布指数 $\gamma$	0.88	1.01	0.95	2.28
参数误差值( $e^{-\gamma}$ )	16.18	7.26	3.88	1.54

由表2可知，在数据取得方面如何误差越小，那么 $\rho$ 取值应越大。在这种特殊情况下当相似度足够大的时候，股票网络的度是不断变小的，也就是说随着相似度不断增大的情况下，根据其度分布所展现出的无标度网络的特点，体现除了股票网络中也只有少数股票之间有这较大的相似，大多股票之间没有相关关系。反之，当我们在相似度在一定取值以下时，幂律分布也随之变得不清晰，网络参数的误差也随之增大。

#### 4 具有重要市场影响力的股票探索分析(Analysis of stock market influence)

在股票网络模型分析的基础上，主要针对相关度数高的股票进行筛选，最终选择了相对度数最高的前的十分之一的股票，因为在实际股市中它们都存在一种现象就是有着最大的连接数量，因此反映出在研究所给定的相应时间内股票之间的收盘价格具有的相关性也非常高，而且呈现无标度状态分布。这部分对应的股票信息内容符合完整性要求，对这些具体的股票进行了基于度的指数公式的定义：

$$\text{基于度的指数} = \frac{\sum [\text{股价}_i \times \text{交易股数}_i]}{\text{基期股票总市值}} \quad (2)$$

主要对应股票在所规定的时间范围中的收盘价格，将相似度设定为可清晰呈现出股票网络的相互连接的图形时的值0.85，而且这个对应的股票连接度正好反映了该股票与其他股票的相互连接数量关系，并将在网络图中呈现出的相连接度最高的那部分10%的股票，并将与之对应的股票信息数据利用式(2)进行再次分析。

通过分析以周为特定时期的度指数结果，并结果结合同期上证指数、沪深300指数进行可视化比较，分别绘制出基于不同指数的股票网络价格收盘价变化趋势图，来证明基于相

似度的指标实用性及可利用性，结果如图4所示。

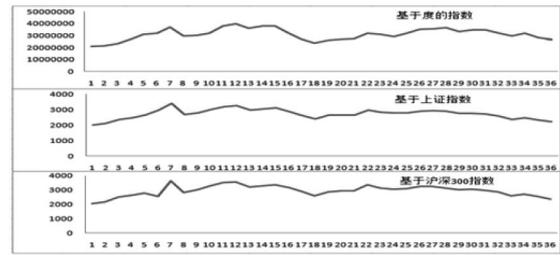


图4 指数对比图

Fig.4 Index comparison

从对比图中可看出，三个图形虽然采用的是不同定义的指标体系，但在股市具体价格变化趋势角度，表现出类似的图形曲线走势图。也正因此，反映出了整个股票市场变化情况方面那些具有高度连接相似的股票所具有的代表性，并且相对应的度数也很高，能够充分反映具体股市的变化，具有一定的实用性和说服力。

总的来说，通过分析和对比，在实际的股票市场中的价格受那些少数且连接度高的股票进行影响的，为了更好的进行研究，我们在那些相连接度高的股票中进行了二次筛选和利用，将这些所选出的股票再次进行度视角下的指数股票与上证指数指数进行了分析与对比，得到的具体数据如表3。

表3 度视角下的度指数与上证指数的股票份额对比图(相似度大于0.85)

Tab.3 Stock share comparison between index and Shanghai stock index( $\rho=0.85$ )

分类	基于上证指数	基于度指数(收盘价格网)
工业	10.12%	7.02%
商业	17.60%	21.30%
金融	42.10%	53.60%
房地产	4.06%	6.72%
公共事业	6.33%	5.25%
综合	19.79%	6.11%
总计	100.00%	100.00%

利用表3所得出的具体数据可以给我们清楚的展现出在度的视角下所反映出的一些结论，其中上证指数中的金融体系的股票在整个网络中具有绝对的影响力，其他股票的影响力相对薄弱许多，在我们设定的度指数视角下，所占具体份额达到了53.6%，相对于上证指数中的42.1%的份额，突显出了更强的地位与影响力。在对比中还表现出商业体系类股票分别在这两个指数中所占份额较大，具体的指数度分别表现为17.6%和21.3%，但就其对股票市场的影响力和地位就远低于金融体系股票。

#### 5 结论(Conclusion)

本文研究了基于相关系数的股票网络市场模型。针对上海证券市场的近九百只股票，我们建立了其基于相关系数的