

А.А. Кочетуров

Центр прикладной оптимизации, Университет Флориды, США

М.В. Бацын

НИУ ВШЭ, Нижний Новгород, Россия

П.М. Пардалос

НИУ ВШЭ, Нижний Новгород, Россия; Центр прикладной оптимизации, Университет Флориды, США

Динамика кластерных структур в сетях фондовых рынков¹

В течение последних пятнадцати лет сетевой анализ активно применялся как инструмент для исследования финансовых рынков. В настоящей работе представлен анализ фондовых рынков США и Швеции, основанный на сетевом подходе. В работе вычисляются и исследуются специальные кластерные структуры в сетях, построенных по корреляционным матрицам доходностей акций фондовых рынков. Кластерная структура сети выделяется с помощью решения задачи о p -медиане, в которой выбирается p центральных акций (медиан) и все акции разбиваются на p кластеров вокруг этих медиан (центров). При этом целевой функцией будет максимизация суммы корреляций между каждой акцией и медианой ее кластера. Полученная кластерная структура представляет собой неориентированный взвешенный граф, компонентами связности которого являются графы-звезды с одной центральной вершиной (медианой) и несколькими конечными вершинами, связанными только с медианой взвешенными ребрами. Наше основное наблюдение заключается в том, что в некризисные периоды времени кластерные структуры сетей фондовых рынков изменяются более хаотично, тогда как во время кризисов они демонстрируют более стабильное поведение и меньшие изменения. Таким образом, повышение устойчивости кластерной структуры сети фондового рынка, полученной с помощью решения задачи о p -медиане, может служить индикатором приближающегося кризиса.

Ключевые слова: *динамика, кластерная структура, фондовые рынки, задача о p -медиане, кластеризация, кризис, сетевой анализ.*

Классификация JEL: C65, G17.

1. Введение

Основная цель работы – проанализировать динамику изменений кластерной структуры для различных фондовых рынков и определить ее связь с кризисными и некризисными периодами. В последнее время подходы, основанные на сетевом анализе, стали одним из современных инструментов для анализа сложных динамических систем, таких как финансовые рынки. Такие подходы позволяют получать полезную информацию, анализируя построенные сети. Например, минимальное остовное дерево в сети фондового рынка отражает некоторую иерархическую структуру, существующую на рынке (Mantegna, 1999), или показывает критически важные банки в межбанковских сетях и оценивает их роль (Mantegna, 1999; Cajueiro, Tabak, 2008; Papadimitriou et al., 2013; Tabak et al., 2014). Клики (Boginski et al.,

¹ Работа выполнена при финансовой поддержке Российского научного фонда (проект 14-41-00039).

2006), объединения (Cajueiro, Tabak, 2008), минимальные остовные деревья (Papadimitriou et al., 2013), p -медианы (Goldengorin et al., 2014) в сети финансового рынка могут быть использованы для сокращения объема анализируемой информации, для выявления критических узлов сети, для визуализации и исследования всей сети.

В ряде работ анализируется динамика финансовых рынков для различных стран (Onnela et al., 2002, 2003; Jung et al., 2006, Huang et al., 2009; Kenett et al., 2010a). Большинство из них используют сетевой подход, предложенный в работе (Mantegna, 1999), в которой описано построение минимального остовного дерева по матрице близости всех акций, входящих в индекс DJIA. Мера близости между двумя акциями вычисляется как оценка коэффициента корреляции Пирсона между доходностями акций. Некоторые авторы предлагают другие сетевые структуры и характеристики для исследования финансовых рынков (Boginski et al., 2006; Cukur et al., 2007; Kenett et al., 2010b, Tabak et al., 2010). В некоторых работах (например, (Tabak et al., 2010)) анализируется поведение в различные периоды времени нескольких сетевых характеристик, не обязательно основанных на каких-либо подструктурах сети.

Необходимо также отметить работу (Buccheri et al., 2013), результаты которой дополнительно подтверждают наши наблюдения для кризисных периодов на фондовых рынках. Авторы (Buccheri et al., 2013) исследовали поведение второго собственного значения матрицы корреляций между 49 индексами акций на фондовом рынке США и обнаружили его резкие изменения во время кризиса доткомов в 2000–2001 гг. и первой фазы мирового финансового кризиса в 2008 г. Перед началом и во время этих кризисов мы также наблюдаем особое поведение кластерных структур – повышение устойчивости.

В настоящей работе предложен новый подход, который позволяет выделять кластерную структуру в сети фондового рынка и исследовать динамику ее изменений с течением времени. Рассматриваются два фондовых рынка США и Швеции. Для вычисления матрицы близости (или матрицы расстояний) акций мы используем цены закрытия (цены на конец торгового дня) для всех акций, торгуемых на этих фондовых рынках. В качестве меры близости применяется оценка корреляции Пирсона между доходностями акций. По матрице близости строится неориентированный взвешенный полный граф, вершинами которого являются акции, а весами ребер – значения меры близости между соответствующими вершинами. С помощью решения задачи о p -медиане для этого графа мы выделяем кластерную структуру в графе, представляющую собой набор из p непересекающихся подграфов-звезд, в которых концевые вершины имеют высокую корреляцию с центральной вершиной своего подграфа (медианой). Решение задачи о p -медиане разбивает граф на p подграфов-звезд так, чтобы суммарный вес всех ребер этих звезд был максимален. Подробный обзор

задачи о p -медиане и методах ее решения можно найти, например, в работе (Reese, 2006).

Все данные о ценах акций делятся на временные интервалы за один календарный год. В среднем для каждой акции за один год получается 251 значение для США и 261 – для Швеции. Затем для каждого интервала времени строится кластерная структура для всех возможных значений $p = 2, \dots, n$ (числа кластеров), где n – число акций на фондовом рынке. Например, на фондовом рынке Швеции рассматриваются $n = 145$ акций и $T = 13$ интервалов времени (13 лет наблюдений). Таким образом, мы должны вычислить $(145 - 1) \times 13 = 1872$ кластерные структуры. Учитывая NP -трудность задачи о p -медиане и большое число вычислений, мы не используем временного окна для изучения динамики изменений кластерной структуры рынка. В этом случае при ширине временного окна в один год для Швеции и общем числе торговых дней, равном 3392, потребовалось бы построить $(145 - 1)(3392 - 261 + 1) = 451\,008$ кластерных структур.

В работе рассматриваются интервалы времени длительностью в 1 год, чтобы число измерений доходностей акций было достаточным для вычисления статистически значимых оценок корреляций между каждой парой акций. Например, для 100 акций необходимо оценить $100 \times 99 / 2 \sim 5000$ попарных корреляций между их доходностями. Если в году примерно 250 торговых дней, получается 25 000 значений доходностей для оценки этих 5000 параметров. При меньшем числе измерений достоверность оценок корреляций будет недостаточной.

Следующим шагом в исследовании динамики изменений кластерной структуры является сравнение кластерных структур, построенных для разных интервалов времени t . Обозначим через $G_{p,t}$ p -кластерную структуру (кластерную структуру из p звезд), полученную для интервала времени t . Для анализа динамики изменений кластерной структуры сети фондового рынка мы попарно сравниваем следующие кластерные структуры: $(G_{p,1}, G_{p,2}), \dots, (G_{p,T-1}, G_{p,T})$ для всех значений $p = 2, \dots, n$. Для сравнения кластерных структур в работе предложены две меры близости между разбиениями графа на кластеры. Обе введенные меры близости демонстрируют специфическое поведение кластерной структуры сети фондового рынка в предкризисные и кризисные периоды: в эти периоды кластерная структура, как правило, ведет себя более устойчиво, чем в некризисные периоды.

2. Исходные данные и мера близости

Для фондового рынка США в работе рассматриваются американские компании, которые входили в индекс S&P 100 на конец 2012 г. Данные включают 13 лет (с 3 января 2000 г. по 31 декабря 2012 г.) и составляют 3269 значений цены закрытия для каждой из выбранных 90 акций. Используются 90 акций из 100, потому что оставшиеся

10 акций торговались на фондовом рынке менее 80% торговых дней в течение по крайней мере одного из 13 лет. Данные разбиваются на 13 периодов по календарным годам. Такое разбиение дает интервалы времени, достаточно большие для вычисления корреляций между доходностями акций со статистической точки зрения и достаточно маленькие для исследования динамики фондового рынка.

Второй фондовый рынок, изучаемый в работе, – это фондовый рынок Швеции с 266 компаниями. Акции 145 компаний торговались не менее 80% всех торговых дней с 3 января 2000 г. по 31 декабря 2012 г. Этот интервал времени состоит из 3392 торговых дней, или 13 календарных лет.

Отсутствующие данные о цене акции мы заполняем значением ее цены в предыдущий или в следующий торговый день. Если известна цена закрытия $P_{i,t}$ для акции i на день t , а в последующие дни $t+1, \dots, t+k$ эта акция не торговалась или данные о ее цене отсутствуют, то цены $P_{i,t+1}, \dots, P_{i,t+k}$ устанавливаются равными $P_{i,t}$. Если же неизвестны цены закрытия в начале периода $P_{i,1}, \dots, P_{i,k}$, они устанавливаются равными $P_{i,k+1}$.

В качестве меры близости мы используем оценку коэффициента корреляции Пирсона между доходностями акций. Чтобы вычислить матрицы корреляций (оценок корреляции), используется формула

$$\rho_{ij} = E\left(\left(R_i - E(R_i)\right)\left(R_j - E(R_j)\right)\right) / \sqrt{\text{var}(R_i)\text{var}(R_j)}, \quad (1)$$

которая дает коэффициент корреляции между доходностями акций i и j . Здесь R_i – доходность акции i , которая вычисляется в соответствии с формулой

$$R_i(t) = \ln(P_i(t) / P_i(t-1)), \quad (2)$$

$P_i(t)$ – цена закрытия для акции i на конец дня t . Чтобы получить оценку на коэффициент корреляции доходностей вместо математического ожидания E и дисперсии var , в формулу (1) подставляются их оценки – выборочное среднее и выборочная дисперсия, вычисленные по всем ценам за исследуемый период.

3. Кластерная структура сети и задача о p -медиане

Используемый в работе алгоритм кластеризации основан на решении задачи о p -медиане (Reese, 2006). Алгоритм разбивает акции фондового рынка на кластеры, внутри которых доходности акций имеют высокую корреляцию. Для этого вычисляется матрица корреляций $P = [\rho_{ij}]_{n \times n}$ в соответствии с формулой (1). Основная идея алгоритма кластеризации заключается в определении набора S центральных акций, или медиан, заданного размера p так, чтобы максимизировать суммарную близость между всеми акциями и этим множеством медиан.

При этом под близостью $\rho(i, S)$ между акцией i и множеством медиан S понимается максимальная корреляция (оценка корреляции) между доходностью этой акции и доходностями каждой медианы множества S : $\rho(i, S) = \max_{j \in S} (\rho_{ij})$. Таким образом, рассматривается задача оптимизации целевой функции:

$$\max_S \sum_{i=1}^n \rho(i, S) = \max_{S \subset X, |S|=p} \left(\sum_{i=1}^n \max_{j \in S} (\rho_{ij}) \right), \quad (3)$$

где X – множество акций на фондовом рынке, n – число этих акций, p – заданное число кластеров. Заменяя простым преобразованием $d_{ij} = 1 - \rho_{ij}$ корреляционную матрицу P на матрицу расстояний $D = [d_{ij}]_{n \times n}$, получаем эквивалентную (3) целевую функцию

$$\min_{S \subset X, |S|=p} \left(\sum_{i=1}^n \min_{j \in S} (d_{ij}) \right). \quad (4)$$

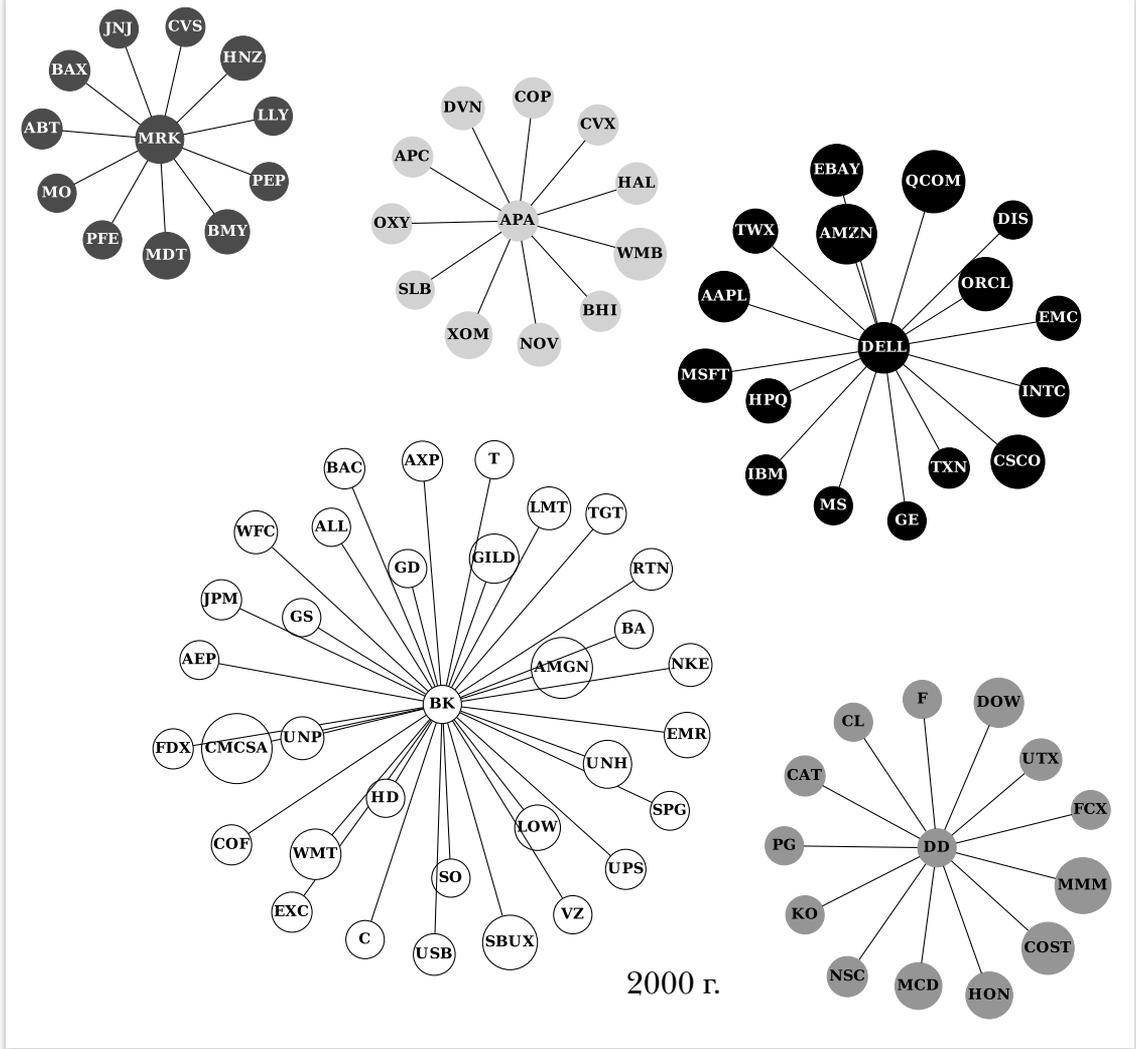


Рис. 1 (а)

5-кластерные структуры для индекса S&P 100 для 2000 г.

Формула (4) полностью соответствует формулировке задачи о p -медиане, в которой множество пунктов обслуживания и множество клиентов совпадают. Подробный обзор задачи о p -медиане можно найти в работе (Reese, 2006).

Решение задачи о p -медиане представляет собой набор из p медиан. При этом каждая акция, не являющаяся медианой, относится к кластеру, медиана которого будет ближайшей к ней из всех медиан. В результате получается кластерная структура в виде набора взвешенных графов-звезд. Примеры такой кластеризации показаны на рис. 1 для 2000 и 2001 г. Названия и область деятельности соответствующих компаний для 2000 г. приведены в Приложении.

Для решения задачи о p -медиане мы применяем псевдобулевский подход, предложенный в работах (Hammer, 1968; Береснев,

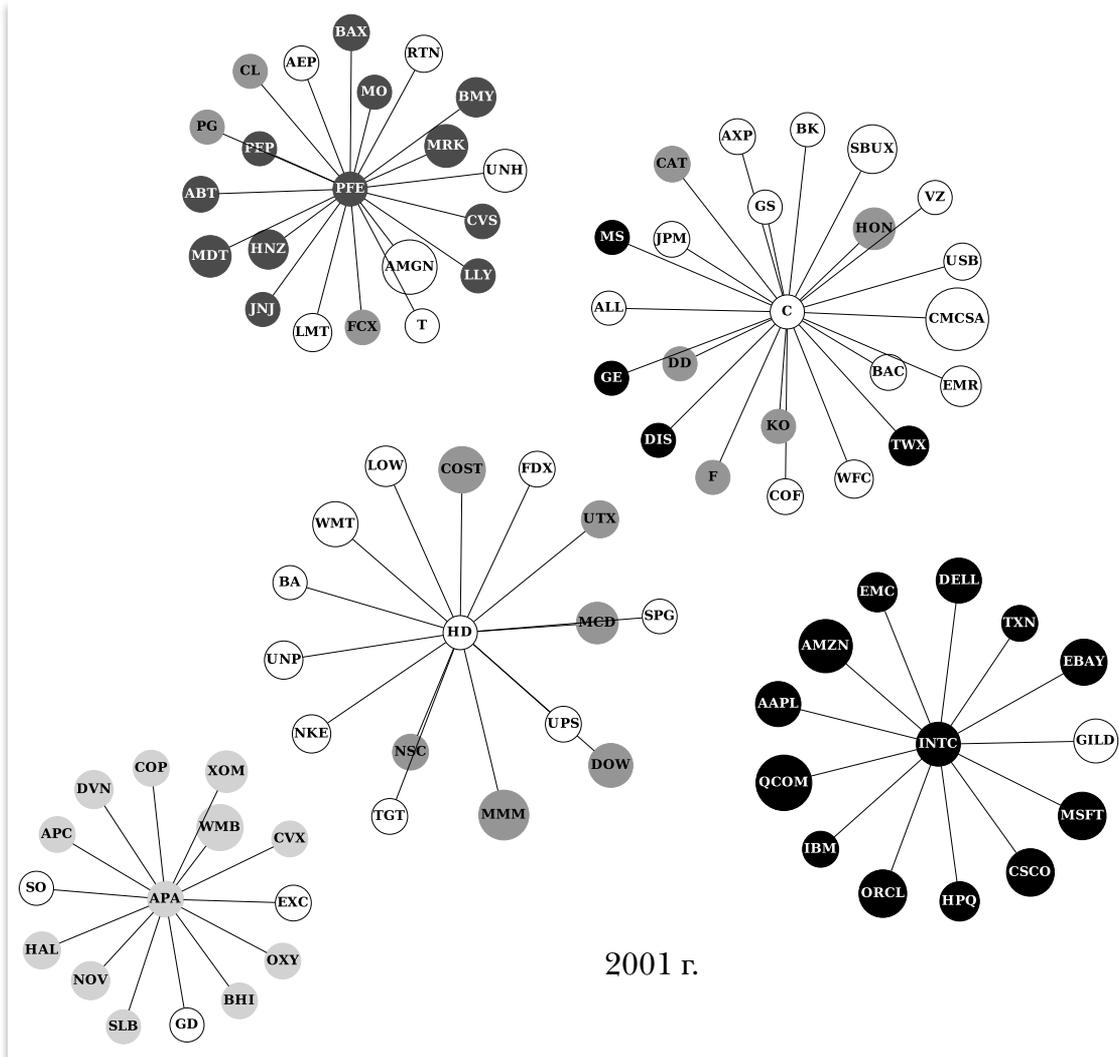


Рис. 1 (б)

5-кластерные структуры для индекса S&P 100 для 2001 г.

1973). Идея использования этого подхода для задачи о p -медианах описана в работе (Goldengorin, Krushinsky, 2011). Получаемая в результате применения этой идеи модель линейного целочисленного программирования легко и быстро решается с помощью стандартного программного обеспечения для тех небольших наборов акций, которые рассматриваются в этой статье. Для этой цели мы используем пакет оптимизации CPLEX.

Чтобы продемонстрировать описанный подход, приведем небольшой пример для 5 акций. Матрица коэффициентов корреляции имеет вид:

$$P_5 = \begin{bmatrix} 1 & 0,9 & 0,8 & 0,3 & 0,2 \\ 0,9 & 1 & 0,7 & 0,4 & 0,1 \\ 0,8 & 0,7 & 1 & 0,3 & 0,2 \\ 0,3 & 0,4 & 0,3 & 1 & 0,9 \\ 0,2 & 0,1 & 0,2 & 0,9 & 1 \end{bmatrix}. \quad (5)$$

Выберем число кластеров p равным 2. Чтобы сформулировать задачу о p -медианах с целевой функцией (4), применим преобразование $d_{ij} = 1 - \rho_{ij}$ и получим матрицу расстояний:

$$D_5 = \begin{bmatrix} 0 & 0,1 & 0,2 & 0,7 & 0,8 \\ 0,1 & 0 & 0,3 & 0,6 & 0,9 \\ 0,2 & 0,3 & 0 & 0,7 & 0,8 \\ 0,7 & 0,6 & 0,7 & 0 & 0,1 \\ 0,8 & 0,9 & 0,8 & 0,1 & 0 \end{bmatrix}. \quad (6)$$

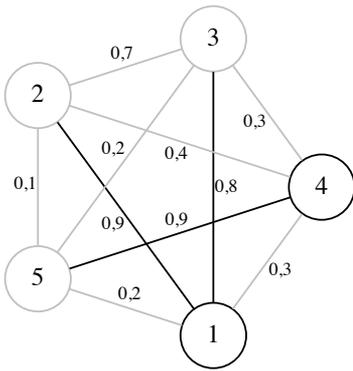


Рис. 2

Пример с 5 акциями и коэффициентами корреляции между ними (медианы и ребра, идущие к ним от концевых вершин кластеров, показаны черным цветом)

Эту задачу нетрудно решить с помощью стандартных моделей целочисленного программирования (см., например, (Reese, 2006)). В нашем примере медианами в оптимальном решении являются вершины 1 и 4, а кластерами – множества вершин $\{1, 2, 3\}$ и $\{4, 5\}$, оптимальное значение целевой функции равно 0,4 (рис. 2).

4. Меры близости для кластерных структур

Для исследования динамики изменений кластерной структуры сети мы сравниваем соседние по времени p -кластерные структуры $G_{p,t}$ и $G_{p,t+1}$ для всех интервалов времени и всех значений числа кластеров p . Таким образом, получается n последовательностей $(G_{p,1}, \dots, G_{p,T})$ для $p = 2, \dots, n$ для каждого фондового рынка. Для каждой пары кластерных структур $G_{p,t}, G_{p,t+1}$ мы вычисляем их близость.

Для сравнения кластерных структур в работе используются две характеристики. Первая характеристика $\alpha_{p,t}$ называется индексом Рэнда (Rand, 1971). Она отражает поведение пар акций при переходе от p -кластерной структуры в период времени t к p -кластерной структуре в период времени $t+1$:

$$\alpha_{p,t} = \sum_{1 \leq i < j \leq n} \frac{x_{i,j}^{p,t}}{n(n-1)/2}. \quad (7)$$

Здесь $x_{i,j}^{p,t}$ принимает значение 1, если акции i и j принадлежали одному/разным кластерам в кластерной структуре сети $G_{p,t}$ и остались в одном/разных кластерах в кластерной структуре сети $G_{p,t+1}$. В противном случае $x_{i,j}^{p,t} = 0$. Другими словами, мы считаем число пар акций, которые либо остались вместе в одном кластере, либо оказались в разных кластерах. Это число делится на количество всех возможных пар, чтобы получить значение в пределах от 0 до 1, т.е. долю пар акций, сохранивших свое взаимное расположение в кластерной структуре. Если значение $\alpha_{p,t}$ близко к 1, это означает, что кластерные структуры $G_{p,t}$ и $G_{p,t+1}$ близки, если же значение близко к 0, структуры сильно отличаются друг от друга. Например, значение $\alpha_{p,t}$ равно 5/6 для кластерных структур, изображенных на рис. 3 (серым цветом показана единственная акция, изменившая свое положение в кластерной структуре).

Необходимо отметить, что характеристика $\alpha_{p,t}$ может принимать значения, близкие к 1, не только для близких кластерных структур, потому что она учитывает акции, которые были в разных кластерах

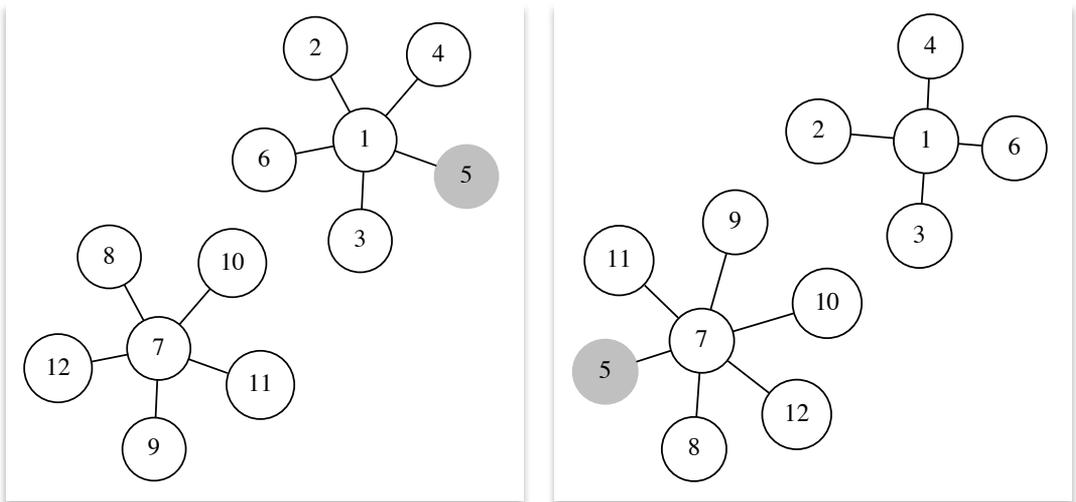


Рис. 3

Две 2-кластерные структуры с мерами близости α и β , близкими к 1

и перешли в разные кластеры, но, возможно, в совсем другие. Поэтому в работе предложена характеристика близости кластерных структур $\beta_{p,t}$, которая отражает поведение групп акций. При переходе от одной кластерной структуры в другую одни акции меняют свой кластер, а другие остаются вместе, образуя устойчивые ядра внутри кластеров. Характеристика $\beta_{p,t}$ вычисляется как доля акций, входящих в такие устойчивые ядра, т.е. как отношение числа таких акций к общему числу акций. Для двух кластерных структур, приведенных на рис. 3, устойчивое ядро первого кластера состоит из 5 акций {1, 2, 3, 4, 6}, а устойчивое ядро второго кластера – из 6 акций {7, 8, 9, 10, 11, 12}. Поэтому характеристика $\beta_{p,t} = 11/12$.

Однако в общем случае необходимо определить, какой кластер в первой кластерной структуре наилучшим образом соответствует какому кластеру во второй кластерной структуре. Для этого требуется оптимальным образом назначить кластеры первой структуры на кластеры второй. Понятно, что чем больше пересечение двух кластеров из разных структур, тем лучше они друг другу соответствуют. Однако один и тот же кластер первой структуры может оказаться наилучшим соответствием сразу для нескольких кластеров второй структуры. Поэтому необходимо сформулировать и решить задачу о назначениях.

Для этого построим матрицу соответствия кластеров $C_{p,t} = [c_{i,j}^{p,t}]$, в которой $c_{i,j}^{p,t}$ будет равно числу акций, принадлежащих одновременно и кластеру i в кластерной структуре $G_{p,t}$, и кластеру j в кластерной структуре $G_{p,t+1}$. Если кластерная структура не изменилась, эта матрица будет диагональной или ее можно будет привести к диагональному виду, перенумеровав кластеры.

Найти оптимальное соответствие между кластерами структуры $G_{p,t}$ и кластерами структуры $G_{p,t+1}$ можно, решив задачу о назначениях с матрицей $C_{p,t}$ и целевой функцией, максимизирующей суммарное соответствие. Для этого в работе используется JV -алгоритм решения задачи о назначениях (Jonker, Volgenant, 1987).

5. Динамика изменений кластерной структуры в сети фондового рынка

Для исследования динамики кластерной структуры сети мы вычисляем характеристики α и β для фондового рынка Швеции и фондового рынка США. На рис. 4–7 приведены графики, отражающие динамику кластерной структуры наблюдаемого фондового рынка (рис. 4–5 иллюстрируют изменения характеристик, измеренных с помощью доли устойчивых пар акций ($\alpha_{p,t}$), а рис. 6–7 – с помощью доли акций в устойчивых ядрах ($\beta_{p,t}$)). Каждый график отражает изменения p -кластерной структуры от $G_{p,1}$ до $G_{p,T}$ для определенного значения p .

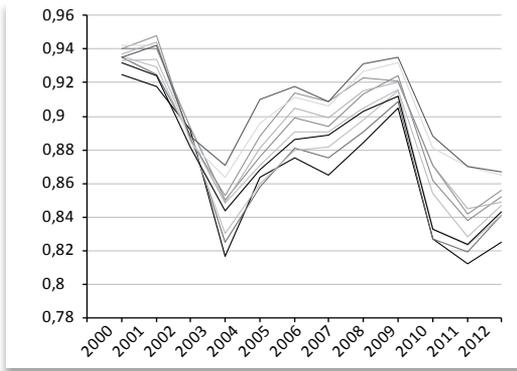


Рис. 4

Динамика изменений кластерной структуры для индекса S&P 100 (для $\alpha_{p,t}$, $p = 16, \dots, 25$)

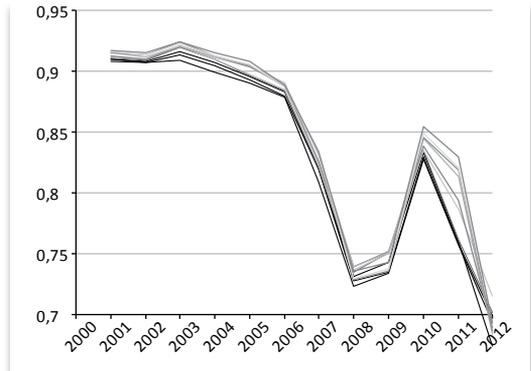


Рис. 5

Динамика изменений кластерной структуры для фондового рынка Швеции (для $\alpha_{p,t}$, $p = 26, \dots, 35$)

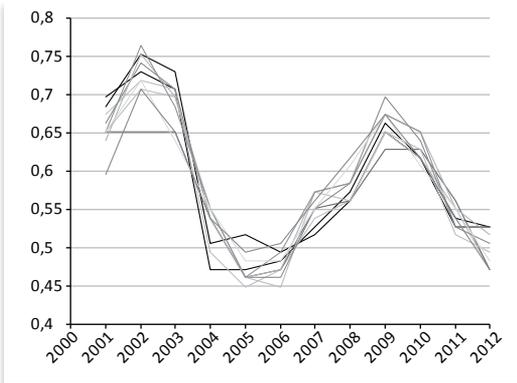


Рис. 6

Динамика изменений кластерной структуры для индекса S&P 100 (для $\beta_{p,t}$, $p = 16, \dots, 25$)

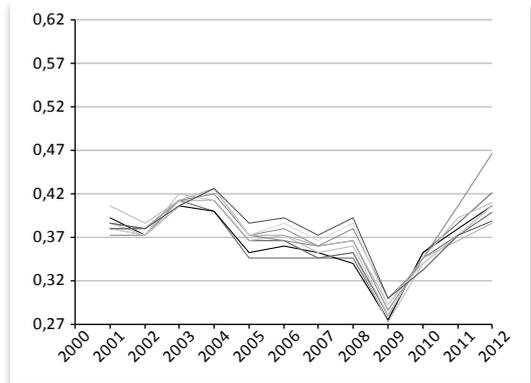


Рис. 7

Динамика изменений кластерной структуры для фондового рынка Швеции (для $\beta_{p,t}$, $p = 26, \dots, 35$)

Мы не приводим графиков характеристик α и β для небольшого числа кластеров p , потому что на них трудно увидеть какое-либо устойчивое поведение соответствующих кластерных структур. Это может быть объяснено большим размером таких кластеров и, как следствие, меньшей близостью между акциями внутри них. Мы также исключили графики для больших значений p , при которых значения $\alpha_{p,t}$ и $\beta_{p,t}$ всегда близки к 1 и разница между кластерными структурами $G_{p,t}$ и $G_{p,t+1}$ всегда мала. Это происходит, потому что при больших значениях p появляется большое число кластеров размера 1.

На рис. 4 и 6 можно увидеть два периода времени – 2000–2002 и 2008–2009 гг., – в которые значения $\alpha_{p,t}$ и $\beta_{p,t}$ имеют локальные максимумы для всех графиков индекса S&P 100. Эти периоды соответ-

ствуют двум известным кризисам, охватившим почти все экономики мира: кризису доткомов в 2000–2002 гг. и начальной фазе мирового финансового кризиса в 2007–2009 гг. Эти результаты показывают, что в периоды кризисов кластерная структура сети фондового рынка ведет себя более устойчиво, или, другими словами, меньшее число акций меняет свои кластеры. В некризисные периоды кластерная структура ведет себя более хаотично. При этом во время кризиса 2000–2002 гг. доходности акций S&P 100 менялись вместе с доходностью самого индекса S&P 100 (акция OEX) – средняя корреляция доходностей акций с доходностью индекса была высока (рис. 8). Однако во время мирового финансового кризиса средняя корреляция принимала низкие значения. Поскольку кластерная структура в этот период имела высокую устойчивость (большинство акций оставались в своих кластерах), это означает, что доходности центральных акций (медиан) кластеров в кластерной структуре имели низкую корреляцию с доходностью индекса S&P 100.

Для индекса S&P 100 графики характеристики $\beta_{p,t}$ (см. рис. 6) повторяют поведение графиков $\alpha_{p,t}$ (см. рис. 4), но для фондового рынка Швеции эти графики отличаются друг от друга. В 2007–2008 гг. все графики $\alpha_{p,t}$ имеют глобальный минимум (см. рис. 5), а все графики $\beta_{p,t}$ имеют локальный максимум (см. рис. 7). Такая ситуация может иметь следующее объяснение. Во время начала мирового финансового кризиса акции, входящие в устойчивые ядра кластеров, оставались в своих кластерах, но другие акции начали хаотично двигаться по отдельности (их корреляции с другими акциями хаотично изменялись). Другими словами, компании, которые входили в устойчивые ядра кластеров, выбрали консервативную стратегию, и поэтому кор-

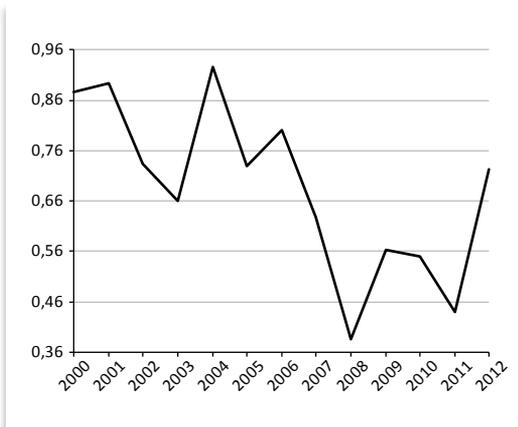


Рис. 8

Средняя корреляция доходностей акций индекса S&P 100 с доходностью самого индекса S&P 100 (акция OEX)

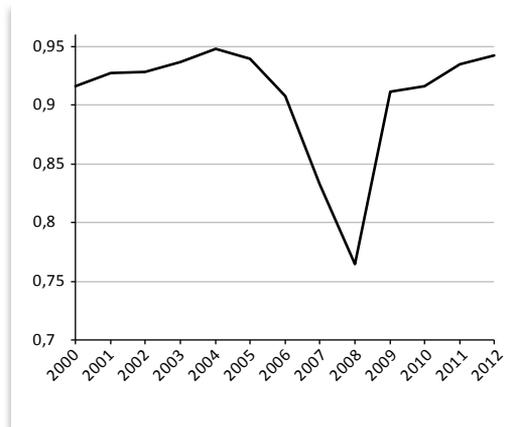


Рис. 9

Средняя корреляция доходностей акций фондового рынка Швеции с доходностью индекса рынка OMXSPI

реляции их акций с другими акциями несильно менялись. Доходности акций других компаний и их корреляции менялись более хаотично. При этом, как и для акций S&P 100, средняя корреляция доходностей акций с доходностью индекса OMXSPI фондового рынка Швеции имеет минимальное значение в 2008 г. (рис. 9).

На рис. 9 характеристики $\beta_{p,t}$ кроме локального максимума в 2007–2008 гг., можно наблюдать и другие максимумы устойчивости кластерной структуры в сети фондового рынка Швеции: в 2003–2005 и 2011–2012 гг. (см. рис. 7). Первый максимум не связан с каким-либо кризисом, однако второй относится к значительному снижению темпов роста экономики Швеции с 6,6% в 2010 г. до 0,9% в 2012 г., связанному с общим экономическим спадом в Еврозоне. В целом, низкие значения характеристики $\beta_{p,t}$ (см. рис. 7) для фондового рынка Швеции говорят о том, что и мировой финансовый кризис, и последующий спад экономики Европы не так сильно повлияли на фондовый рынок Швеции и не привели к серьезным изменениям его кластерной структуры.

Мы также проверили чувствительность полученных результатов, удаляя случайным образом небольшой процент акций и заново вычисляя характеристики $\alpha_{p,t}$ и $\beta_{p,t}$. Эта процедура повторялась десять раз для 5 и 10% удаляемых акций, и каждый раз результаты оставались качественно теми же, хотя и немного изменялись количественно.

6. Кластеризация фондового рынка на основе функции модулярности

Так как в работе рассматривается кластеризация фондового рынка, естественно будет использовать и другие меры кластеризации для исследования динамики кластерной структуры в сети фондового рынка и сравнить полученные результаты. Одной из современных и успешных мер кластеризации является функция модулярности Q , введенная М. Ньюманом (Newman, 2004):

$$Q_t = \frac{1}{2sum} \sum_i \sum_j \left(\rho_{ij} - \frac{\rho_i \rho_j}{2sum} \right) \delta(i, j), \quad (8)$$

где $2sum = \sum_i \sum_j \rho_{ij}$, $\rho_i = \sum_k \rho_{ik}$, t – период времени, а $\delta(i, j) = 1$, если акции i и j принадлежат одному кластеру, и $\delta(i, j) = 0$ в противном случае.

Для каждого календарного года t и всех полученных для него кластерных структур $G_{p,t}$ мы определяем максимальное значение Q . Нам важно и максимальное значение модулярности, и число кластеров, при котором оно достигается, полученные для каждого года (рис. 10–13). Этот подход позволяет выявить кризис доткомов на обоих фондовых рынках. Самые высокие значения модулярности и относительно небольшое число кластеров соответствуют этому кризису. Однако мировой финансовый кризис никак не выделяется на графиках.

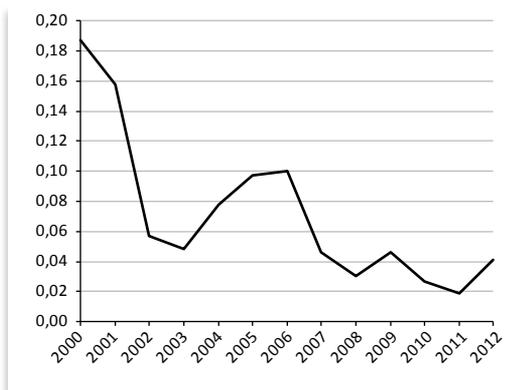


Рис. 10

Значения функции модулярности Q для индекса S&P 100

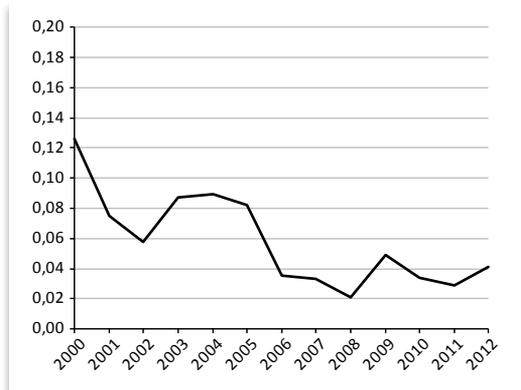


Рис. 11

Значения функции модулярности Q для фондового рынка Швеции

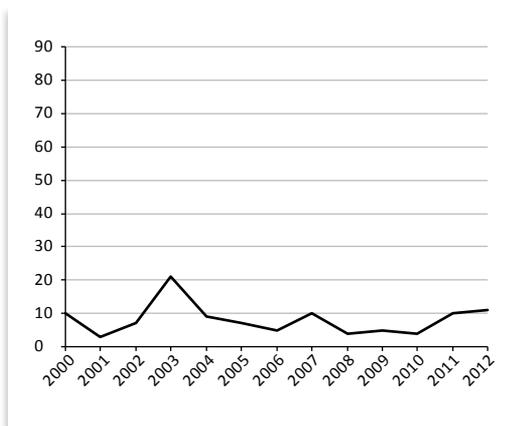


Рис. 12

Число кластеров для индекса S&P 100

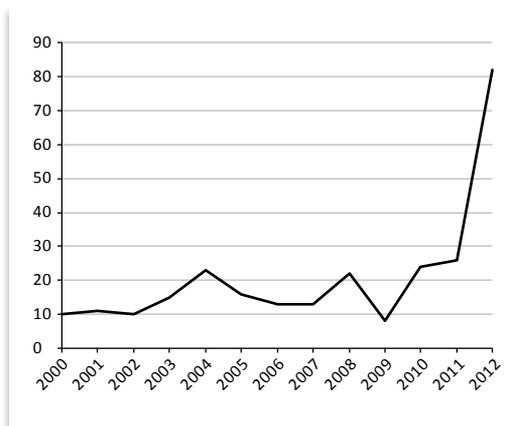


Рис. 13

Число кластеров для фондового рынка Швеции

7. Заключение

В работе изучается динамика кластерных структур в сетях фондовых рынков США и Швеции. Предложен оригинальный способ построения кластерной структуры сети. Авторы обнаружили, что в кризисные периоды, такие как кризис доткомов, ипотечный кризис в США, мировой финансовый кризис, построенная кластерная структура сети ведет себя более устойчиво. Таким образом, увеличение устойчивости кластерной структуры в сети фондового рынка, полученной с помощью решения задачи о p -медиане, может служить индикатором приближающегося кризиса.

Мы обнаружили интересное поведение акций на фондовом рынке Швеции во время ипотечного кризиса США в 2007–2008 гг.:

ядра кластеров были более устойчивы, чем в другие периоды, в то время как акции, не принадлежащие ядрам, активно перемещались между кластерами. Это одно из отличий между фондовыми рынками США и Швеции.

Другим отличием является то, что значения характеристики $\beta_{p,t}$ для индекса S&P 100 гораздо выше, чем для фондового рынка Швеции. Это говорит о том, что фондовый рынок Швеции достаточно спокойно пережил все кризисные периоды, постоянно сохраняя достаточно высокую хаотичность (низкую устойчивость $\beta_{p,t}$) своей кластерной структуры, характерную для некризисных периодов. Отчасти это также может быть связано с тем, что акции из индекса S&P 100 принадлежат крупнейшим компаниям, торгующимся на американском фондовом рынке. Поэтому эти акции подвержены меньшим изменениям, и динамика соответствующих кластерных структур является более устойчивой.

ПРИЛОЖЕНИЕ

Таблица

Рассматриваемые компании из индекса S&P 100

Акция	Название компании	Сфера деятельности
AAPL	Apple Inc.	Information Technology
ABT	Abbott Laboratories	Health Care
AEP	American Electric Power	Utilities
ALL	Allstate Corp	Financials
AMGN	Amgen Inc	Health Care
AMZN	Amazon.com Inc	Consumer Discretionary
APA	Apache Corp	Energy
APC	Anadarko Petroleum Corp	Energy
AXP	American Express Co	Financials
BA	Boeing Co	Industrials
BAC	Bank of America Corp	Financials
BAX	Baxter Intl Inc	Health Care
BHI	Baker Hughes Inc	Energy
BK	The Bank of New York Mellon Corp	Financials
BMJ	Bristol-Myers Squibb	Health Care
C	Citigroup Inc	Financials
CAT	Caterpillar Inc	Industrials
CL	Colgate-Palmolive Co	Consumer Staples
CMCSA	Comcast Corp	Consumer Discretionary
COF	Capital One Financial	Financials
COP	ConocoPhillips	Energy

Продолжение таблицы

Акция	Название компании	Сфера деятельности
COST	Costco Wholesale Corp	Consumer Staples
CSCO	Cisco Systems Inc	Information Technology
CVS	CVS Caremark Corp.	Consumer Staples
CVX	Chevron Corp	Energy
DD	DuPont	Materials
DELL	Dell Inc	Information Technology
DIS	Walt Disney Co	Consumer Discretionary
DOW	Dow Chemical	Materials
DVN	Devon Energy Corp	Energy
EBAY	eBay Inc.	Information Technology
EMC	EMC Corp	Information Technology
EMR	Emerson Electric Co	Industrials
EXC	Exelon Corp	Utilities
F	Ford Motor Co	Consumer Discretionary
FCX	Freeport McMoRan Copper & Gold	Materials
FDX	FedEx Corp	Industrials
GD	General Dynamics	Industrials
GE	General Electric Co	Industrials
GILD	Gilead Sciences Inc	Health Care
GS	Goldman Sachs Group Inc	Financials
HAL	Halliburton Co	Energy
HD	Home Depot Inc	Consumer Discretionary
HNZ	Heinz	Consumer Staples
HON	Honeywell Intl Inc	Industrials
HPQ	Hewlett-Packard Co	Information Technology
IBM	Intl Business Machines Corp	Information Technology
INTC	Intel Corp	Information Technology
JNJ	Johnson & Johnson	Health Care
JPM	JP Morgan Chase & Co	Financials
KO	Coca-Cola Co	Consumer Staples
LLY	Lilly	Health Care
LMT	Lockheed Martin	Industrials
LOW	Lowe's Cos Inc	Consumer Discretionary
MCD	McDonald's Corp	Consumer Discretionary
MDT	Medtronic Inc	Health Care
MMM	3M Co	Industrials

Окончание таблицы

Акция	Название компании	Сфера деятельности
MO	Altria Group Inc	Consumer Staples
MRK	Merck & Co Inc	Health Care
MS	Morgan Stanley	Financials
MSFT	Microsoft Corp	Information Technology
NKE	NIKE Inc B	Consumer Discretionary
NOV	National Oilwell Varco Inc	Energy
NSC	Norfolk Southern Corp	Industrials
ORCL	Oracle Corp	Information Technology
OXY	Occidental Petroleum	Energy
PEP	PepsiCo Inc	Consumer Staples
PFE	Pfizer Inc	Health Care
PG	Procter & Gamble	Consumer Staples
QCOM	QUALCOMM Inc	Information Technology
RTN	Raytheon Co	Industrials
SBUX	Starbucks Corp	Consumer Discretionary
SLB	Schlumberger Ltd	Energy
SO	Southern Co	Utilities
SPG	Simon Property Group	Financials
T	AT&T Inc	Telecommunications Services
TGT	Target Corp	Consumer Discretionary
TWX	Time Warner Inc	Consumer Discretionary
TXN	Texas Instruments Inc	Information Technology
UNH	Unitedhealth Group Inc	Health Care
UNP	Union Pacific Corp	Industrials
UPS	United Parcel Service Inc B	Industrials
USB	US Bancorp	Financials
UTX	United Technologies Corp	Industrials
VZ	Verizon Communications Inc	Telecommunications Services
WFC	Wells Fargo & Co	Financials
WMB	Williams Cos Inc	Energy
WMT	Wal-Mart Stores	Consumer Staples
XOM	Exxon Mobil Corp	Energy

ЛИТЕРАТУРА

- Береснев В.Л.** (1973). Об одной задаче математической теории стандартизации
// *Управляемые системы*. № 11. С. 43–54.

- Boginski V., Butenko S., Pardalos P.M.** (2006). Mining Market Data: a Network Approach // *Computers & Operations Research*. No. 33. P. 3171–3184.
- Buccheri G., Marmi S., Mantegna R.N.** (2013). Evolution of Correlation Structure of Industrial Indices of U.S. Equity Markets // *Physical Review E*. No. 88(1). 012806.
- Cajueiro D.O., Tabak B.M.** (2008). The Role of Banks in the Brazilian Interbank Market: Does Bank Type Matter? // *Physica A: Statistical Mechanics and its Applications*. No. 387(27). P. 6825–6836.
- Cukur S., Eryigit M., Eryigit R.** (2007). Cross Correlations in an Emerging Market Financial Data // *Physica A*. No. 376. P. 555–564.
- Goldengorin B., Kocheturov A., Pardalos P.M.** (2014). A Pseudo-Boolean Approach to the Market Graph Analysis by Means of the p-Median Model. In: “*Clusters, Orders, Trees: Methods and Applications*” Aleskerov F., Goldengorin B., Pardalos P.M. (eds.). N.Y.: Springer. P. 77–89.
- Goldengorin B., Krushinsky D.** (2011). Complexity Evaluation of Benchmark Instances for the p-Median Problem // *Mathematical and Computer Modelling*. No. 53(9). P. 1719–1736.
- Hammer P.L.** (1968). Plant Location: A Pseudo-Boolean Approach // *Israel Journal of Technologies*. No. 6. P. 330–332.
- Huang W.-Q., Zhuang X.-T., Yao S.** (2009). A Network Analysis of the Chinese Stock Market // *Physica A*. No. 388. P. 2956–2964.
- Jonker R., Volgenant A.** (1987). A Shortest Augmenting Path Algorithm for Dense and Sparse Linear Assignment Problems // *Computing*. No. 38(4). P. 325–340.
- Jung W.-S., Chae S., Yang J.-S., Moon H.-T.** (2006). Characteristics of the Korean Stock Market Correlations // *Physica A*. No. 361. P. 263–271.
- Kenett D.Y., Shapira Y., Madi A., Bransburg-Zabary S., Gur-Gershgoren G., Ben-Jacob E.** (2010a). Dynamics of Stock Market Correlations // *AUCO Czech Economic Review*. No. 4. P. 330–340.
- Kenett D.Y., Tumminello M., Madi A., Gur-Gershgoren G., Mantegna R.N., Ben-Jacob E.** (2010b). Dominating Clasp of the Financial Sector Revealed by Partial Correlation Analysis of the Stock Market // *PLoS ONE*. No. 5(12). e15032.
- Mantegna R.N.** (1999). Hierarchical Structure in Financial Markets // *European Physical Journal B*. No. 11. P. 193–197.
- Newman M.E.J.** (2004). Analysis of Weighted Networks // *Physical Review E*. No. 70. 056131.
- Onnela J.-P., Chakraborti A., Kaski K., Kertesz J.** (2002). Dynamic Asset Trees and Portfolio Analysis // *European Physical Journal B*. No. 30. P. 285–288.
- Onnela J.-P., Chakraborti A., Kaski K., Kertesz J.** (2003). Dynamic Asset Trees and Black Monday // *Physica A*. No. 324. P. 247–252.
- Papadimitriou T., Gogas P., Tabak B.M.** (2013). Complex Networks and Banking Systems Supervision // *Physica A: Statistical Mechanics and its Applications*. No. 392(19). P. 4429–4434.
- Rand W.M.** (1971). Objective Criteria for the Evaluation of Clustering Methods // *Journal of the American Statistical Association*. No. 66(336). P. 846–850.

- Reese J.** (2006). Solution Methods for the p-Median Problem: An Annotated Bibliography // *Networks*. No. 48(3). P. 125–142.
- Tabak B.M., Serra T.R., Cajueiro D.O.** (2010). Topological Properties of Stock Market Networks: The Case of Brazil // *Physica A: Statistical Mechanics and its Applications*. No. 389(16). P. 3240–3249.
- Tabak B.M., Takami M., Rocha J.M.C., Cajueiro D.O., Souza S.R.S.** (2014). Directed Clustering Coefficient as a Measure of Systemic Risk in Complex Banking Networks // *Physica A: Statistical Mechanics and its Applications*. No. 394(C). P. 211–216.

REFERENCES (with English translation or transliteration)

- Beresnev V.L.** (1973). On one Problem of Mathematical Theory of Standardization. *Controllable Systems* 11, 43–54 (in Russian).
- Boginski V., Butenko S., Pardalos P.M.** (2006). Mining Market Data: a Network Approach. *Computers & Operations Research* 33, 3171–3184.
- Buccheri G., Marmi S., Mantegna R.N.** (2013). Evolution of Correlation Structure of Industrial Indices of U.S. Equity Markets. *Physical Review E* 88(1), 012806.
- Cajueiro D.O., Tabak B.M.** (2008). The Role of Banks in the Brazilian Interbank Market: Does Bank Type Matter? *Physica A: Statistical Mechanics and Its Applications* 387(27), 6825–6836.
- Cukur S., Eryigit M., Eryigit R.** (2007). Cross Correlations in an Emerging Market Financial Data. *Physica A* 376, 555–564.
- Goldengorin B., Kocheturov A., Pardalos P.M.** (2014). A Pseudo-Boolean Approach to the Market Graph Analysis by Means of the p-Median Model. In: “*Clusters, Orders, Trees: Methods and Applications*” Aleskerov F., Goldengorin B., Pardalos P.M. (eds.). N.Y.: Springer, 77–89.
- Goldengorin B., Krushinsky D.** (2011). Complexity Evaluation of Benchmark Instances for the p-Median Problem. *Mathematical and Computer Modelling* 53(9), 1719–1736.
- Hammer P.L.** (1968). Plant Location: A Pseudo-Boolean Approach. *Israel Journal of Technologies* 6, 330–332.
- Huang W.-Q., Zhuang X.-T., Yao S.** (2009). A Network Analysis of the Chinese Stock Market. *Physica A* 388, 2956–2964.
- Jonker R., Volgenant A.** (1987). A Shortest Augmenting Path Algorithm for Dense and Sparse Linear Assignment Problems. *Computing* 38(4), 325–340.
- Jung W.-S., Chae S., Yang J.-S., Moon H.-T.** (2006). Characteristics of the Korean Stock Market Correlations. *Physica A* 361, 263–271.
- Kenett D.Y., Shapira Y., Madi A., Bransburg-Zabary S., Gur-Gershgoren G., Ben-Jacob E.** (2010a). Dynamics of Stock Market Correlations. *AUCO Czech Economic Review* 4, 330–340.
- Kenett D.Y., Tumminello M., Madi A., Gur-Gershgoren G., Mantegna R.N., Ben-Jacob E.** (2010b). Dominating Clasp of the Financial Sector Revealed by Partial Correlation Analysis of the Stock Market. *PLoS ONE* 5(12), e15032.

- Mantegna R.N.** (1999). Hierarchical Structure in Financial Markets. *European Physical Journal B* 11, 193–197.
- Newman M.E.J.** (2004). Analysis of Weighted Networks. *Physical Review E* 70, 056131.
- Onnela J.-P., Chakraborti A., Kaski K., Kertesz J.** (2002). Dynamic Asset Trees and Portfolio Analysis. *European Physical Journal B* 30, 285–288.
- Onnela J.-P., Chakraborti A., Kaski K., Kertesz J.** (2003). Dynamic Asset Trees and Black Monday. *Physica A* 324, 247–252.
- Papadimitriou T., Gogas P., Tabak B.M.** (2013). Complex Networks and Banking Systems Supervision. *Physica A: Statistical Mechanics and Its Applications* 392(19), 4429–4434.
- Rand W.M.** (1971). Objective Criteria for the Evaluation of Clustering Methods. *Journal of the American Statistical Association* 66(336), 846–850.
- Reese J.** (2006). Solution Methods for the p -Median Problem: An Annotated Bibliography. *Networks* 48(3), 125–142.
- Tabak B.M., Serra T.R., Cajueiro D.O.** (2010). Topological Properties of Stock Market Networks: The Case of Brazil. *Physica A: Statistical Mechanics and its Applications* 389(16), 3240–3249.
- Tabak B.M., Takami M., Rocha J.M.C., Cajueiro D.O., Souza S.R.S.** (2014). Directed Clustering Coefficient as a Measure of Systemic Risk in Complex Banking Networks. *Physica A: Statistical Mechanics and Its Applications* 394(C), 211–216.

Поступила в редакцию 23 апреля 2015 года

A.A. Kocheturov

Center for Applied Optimization, University of Florida, USA

M.V. Batsyn

NRU HSE, Nizhny Novgorod, Russia

P.M. Pardalos

NRU HSE, Nizhny Novgorod, Russia; Center for Applied Optimization, University of Florida, USA

Dynamics of Cluster Structures in Stock Market Networks

In recent 15 years network analysis has been actively applied for studying financial markets. In this paper we present a network-based analysis of stock markets of USA and Sweden. We extract and study special cluster structures of networks built from correlation matrices of stock returns for these stock markets. A cluster structure of a network is extracted by solving the p -median problem which chooses p central stocks (medians) and partitions all stocks into p clusters around these medians – centers. The objective function maximizes the sum of correlations between each stock and the median of its cluster. The obtained cluster structure is represented by an undirected disconnected weighted graph, which components are star-graphs with one central vertex (median) and several leaf vertices connected only with the median by weighted edges. Our main observation is that in non-crisis periods cluster structures of stock market networks change more chaotically, while during crises they demonstrate more stable behavior and smaller changes. Thus an increase in stability of the cluster structure for a stock market network obtained by means of the p -median problem solution can serve as an indicator of a coming crisis.

Keywords: *dynamics, cluster structure, stock markets, p -median problem, clustering, crisis, network analysis.*

JEL Classification: C65, G17.