ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ

ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ

ВЫСШЕГО ОБРАЗОВАНИЯ

**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ**

**«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

**Факультет информатики, математики и компьютерных наук**

Программа подготовки магистров по направлению

01.04.02 Прикладная математика и информатика

**Интеллектуальный анализ данных**

**КУРСОВАЯ РАБОТА**

Тема:

**«Исследование динамики связей на рыночном графе»**

|  |  |
| --- | --- |
|  | **Выполнил:** студент группы 22 МАГ ИАД  Рухович Игорь Владимирович  **Руководитель:** профессор,  ведущий научный сотрудник  Колданов Александр Петрович |

1. Оглавление

[1. Оглавление 2](#_Toc137248254)

[2. Введение 3](#_Toc137248255)

[2.1. Актуальность выбранной темы 3](#_Toc137248256)

[3. Теоретическая часть 4](#_Toc137248257)

[3.1. Граф рынка 4](#_Toc137248258)

[3.1.1. Описание рыночного графа 4](#_Toc137248259)

[3.1.2. Определение рыночного графа 4](#_Toc137248260)

[3.1.3. Другие варианты определения графа рынка 5](#_Toc137248261)

[3.2. Клики в рыночном графе 6](#_Toc137248262)

[4. Практическая часть 7](#_Toc137248263)

[5. Анализ результатов 8](#_Toc137248264)

[6. Заключение 9](#_Toc137248265)

[7. Список литературы 10](#_Toc137248266)

[8. Приложения 11](#_Toc137248267)

1. Введение
   1. Актуальность выбранной темы

С момента своего появления и по сей день биржи вызывают к себе большой интерес. Для крупных компаний присутствие на бирже – это способ получить инвестиции без лишних обязательств, иметь дополнительное внешнее управление, разделять владение, и даже награждать отличившихся сотрудников без налоговых обременений. Для простых людей, а также для банков, биржи – источник дополнительного, или даже основного заработка, способ сохранить и преумножить капитал. Особенную популярность торговля на биржах приобрела в эпоху цифровых технологий, когда покупать и продавать ценные бумаги стало возможно онлайн, а порог входа на рынок снизился до наличия стабильного доступа в Интернет и подписания специального договора с одним из участников рынка.

Впрочем, не все инвесторы на рынке что-то зарабатывают. Котировки акций на бирже постоянно меняются, моментально реагируя на различные события в компаниях, а также на глобальные политические изменения, курсы валют. Риск потерять вложения всегда довольно существенный, особенно для неопытных игроков, причем сбывается этот риск достаточно часто как у частных инвесторов, так и у крупных инвестиционных компаний.

Анализ рыночных котировок в последнее время привлекает к себе колоссальное внимание всех заинтересованных сторон. Ведь если какой-то человек или компания узнает, пускай с некоторой погрешностью, что будет с курсом той или иной ценной бумаги на заданном интервале времени, он будет обладать существенным преимуществом перед другими игроками, а значит и возможностью увеличить свой капитал.

1. Теоретическая часть
   1. Граф рынка
      1. Описание рыночного графа

Одной из структур, которой активно пользуются аналитики для понимания и описания рынков являются графы. Более формально, графом:

|  |  |  |
| --- | --- | --- |
|  |  | ф. (3.1) |

называют пару – множество вершин и – множество пар из , означающих рёбра в графе. Если за множество вершин взять множество ценных бумаг на каком-либо рынке, а за множество рёбер – индикаторы связей между курсами этих бумаг, то мы получим концепцию рыночного графа [1, 2, 3]. В графе рынка рёбра между вершинами и проводятся при условии, что коэффициент корреляции между этими вершинами не меньше заданного порога .

* + 1. Определение рыночного графа

Более точно, пусть – цена бумаги на конец торгового дня . Тогда:

|  |  |  |
| --- | --- | --- |
|  |  | ф. (3.2) |

называют доходность акции за период в один день с по . Средней доходностью акции за период в дней называют:

|  |  |  |
| --- | --- | --- |
|  |  | ф. (3.3) |

Дисперсией доходности акции за период в дней называют:

|  |  |  |
| --- | --- | --- |
|  |  | ф. (3.4) |

Тогда коэффициент корреляции Пирсона между акциями и за промежуток времени в дней определяется как:

|  |  |  |
| --- | --- | --- |
|  |  | ф. (3.5) |

Посчитав коэффициенты корреляции для каждой пары акций , можно получить матрицу корреляции размера . Эту матрицу будем интерпретировать как матрицу смежности полного взвешенного графа, вершинами которого являются ценные бумаги. Из определения корреляции Пирсона заметим, что матрица симметрична относительно главной диагонали, поскольку корреляция коммутативна. Это значит, что полученный из неё граф будет неориентированным.

Полный взвешенный граф затруднительно анализировать, поэтому из него принято выделять подмножество рёбер, являющееся в той или иной степени самым важным. Возьмём некоторый порог и построим граф

|  |  |  |
| --- | --- | --- |
|  |  | ф. (3.6) |

Такой граф и принято называть графом рынка [4].

* + 1. Другие варианты определения графа рынка

В литературе также встречаются [5, 6] и другие способы определения рыночного графа. Они аналогичны до этапа получения матрицы смежности, но дальнейшую фильтрацию рёбер можно, например, делать с помощью MST (Maximum Spanning Tree).

Алгоритм построения максимального (или минимального) остовного дерева довольно прост – рёбра графа сортируются по невозрастанию (или неубыванию) весов, а затем поочередно добавляются в пустой граф по условию: добавляем только те рёбра, которые уменьшают число компонент связности. По-другому – если добавление нового ребра создаёт в графе цикл – такое ребро не попадает в MST.

С помощью остовных деревьев можно так же анализировать зависимости между ценными бумагами: “тяжелые” рёбра – высокая зависимость, рёбра “лёгкие” – зависимость между курсами этой пары акций низкая. Если поочерёдно удалять из MST наиболее лёгкие рёбра, можно получить разделение ценных бумаг на связные кластеры.

Дальнейшее развитие этой теории привело [6] к пониманию, что, оставляя от полного графа лишь его остов, мы теряем много информации. Поэтому было предложено добавлять в MST оставшиеся рёбра в порядке невозрастания их весов пока граф остаётся планарным. Планарный граф – граф, рисунок которого убирается на плоскость без пересечения рёбер. Более формально – граф не содержит подграфов, стягивающихся в или [7] (теорема Понтрягина – Куратовского). Такой граф стали называть PMFG (Planar Maximally Filtered Graph).

В данной работе будет рассмотрен рыночный граф только в классическом его представлении – на основе порогов.

* 1. Клики в рыночном графе

Для того, чтобы получить полезную информацию из рыночного графа, предлагается рассматривать его максимальные клики. Клика в графе – это подмножество его вершин, такое, что между любой парой вершин имеется ребро. Максимальная клика – это клика графа, которая при добавлении любой новой вершины, перестаёт быть кликой. К примеру, в PMFG максимальный размер клики по определению планарного графа равен 4 (при наличии в исходном графе хотя бы четырех вершин). Клик максимального размера в графе может быть несколько. Задача поиска максимальной по включению клики в графе относится к классу NP-полных.

Максимальные клики – это то, на что обычно обращают внимание при анализе рыночного графа. На клику можно смотреть как на множество акций, стоимость которых изменяется похожим образом. При этом важно понимать, что хотя попарная корреляция между каждой парой вершин достаточно велика, ещё не значит, что соответствующие им случайные величины зависимы вместе. Для проверки этого факта требуется одновременное выполнение гипотез о попарной зависимости этих величин с заданным уровнем значимости. Впрочем при анализе рыночных графов этот факт часто опускают.

1. Практическая часть
   1. Введение

Экспериментальная часть работы была частично вдохновлена статьей [8], опубликованной в 2014 году. В ней рассматриваются рыночные графы, построенные на данных об акциях Московской Биржи за промежуток времени с середины 2007 по середину 2011 года. Это время интересно тем, что в середине отрезка (2008 год) произошел глобальный финансовый кризис, безусловно повлиявший на котировки многих акций.

В нашей же работе появился интерес сравнить полученные исследователями данные с ситуацией на данный момент и попытаться отследить похожие моменты и динамику.

* 1. Получение данных

Для нашей работы информацию было решено брать из Интернет-сервиса Tinkoff Invest API [9]. В сервисе доступны данные по всем акциям, торгующимся на Московской бирже за период, начиная с [начало периода].

[добавить как получали данные]

* 1. Преобразование данных

[добавить про разбиение на периоды, подсчет корреляций]

* 1. Поиск и выделение максимальных клик

[добавить про построение графов, поиск клик, объединение клик в подмножества с весами]

1. Анализ результатов
2. Заключение
3. Список литературы
4. Boginski, V., Butenko, S., Pardalos, P.M.: Mining market data: a network approach. Comput. Oper. Res. 33(11):3171–3184, (2006)
5. Boginski, V., Butenko, S., Pardalos, P.M.: On structural properties of the market graph. Innov. Financ. Econ. Netw. 29–45, (2003)
6. Boginski, V., Butenko, S., Pardalos, P.M.: Statistical analysis of financial networks. Comput. Stat. Data Anal. 48(2):431–443, (2005)
7. Mantegna R.N., Stanley H.E.: An introduction to econophysics: corrleations and complexity in finance. Cambridge Universiy Press, Cambridge, (2000)
8. Kalyagin V.A., Koldanov A.P., Koldanov P.A., Pardalos P.M.: Statistical analysis of graph structures in random variable networks, Springer Brief in Optimization, Springer, (2020)
9. Tumminello, M., Aste, T., Di Matteo, T., Mantegna, R.N.: A tool for filtering information in complex systems. Proc. Natl. Acad. Sci. U.S.A. 102(30):10421–10426, (2005)
10. Kuratowski, K.: Sur le problème des courbes gauches en topologie, Fund. Math. Т. 15: 271–283, (1930)
11. Vizgunov, A.N., Goldengorin, B., Kalyagin, V.A., Koldanov, A.P., Koldanov, P., Pardalos, P.M.: Network approach for the Russian stock market. Comput. Manag. Sci. 11:45–55 (2014)
12. Tinkoff Invest API // Сервис взаимодействия с биржей через брокера Tinkoff. [Электронный ресурс]. – Режим доступа: <https://www.tinkoff.ru/invest/open-api/> (Дата обращения: 27.05.2023)
13. Приложения