

Санкт-Петербургский Государственный Университет

Математическое обеспечение и администрирование информационных систем

Информационно-аналитические системы

Пашкова Мария Владимировна

Разработка имитационной системы  
моделирования фондового рынка для  
изучения его характеристик

Бакалаврская работа

к. ф.-м. н., доцент кафедры информатики  
Санкт-Петербургского государственного  
университета Григорьев Д. А.

д. т. н., профессор, декан факультета  
информационных технологий и управления  
Санкт-Петербургского государственного  
технологического института Мусаев А. А.

Санкт-Петербург  
2018

SAINT-PETERSBURG STATE UNIVERSITY

Software and Administration of Information Systems  
Analytical Information Systems

Mariya Pashkova

Development of the simulation system for  
modeling the stock market to study its  
characteristics

Bachelor's Thesis

Scientific supervisor:  
Ph.D., assistant professor Dmitry Grigoriev

Reviewer:  
Sc.D, professor Alexander Musaev

Saint-Petersburg  
2018

# Оглавление

<b>Введение</b>	<b>4</b>
<b>1. Финансовый рынок</b>	<b>5</b>
1.1. Определение . . . . .	5
1.2. Классификация . . . . .	6
1.3. Функции . . . . .	7
<b>2. Моделирование фондового рынка</b>	<b>8</b>
2.1. Агентно-ориентированное моделирование . . . . .	9
2.2. Стилизованные факты . . . . .	10
2.3. Гипотеза эффективного рынка . . . . .	11
<b>3. Базовая система</b>	<b>12</b>
3.1. Описание . . . . .	12
3.1.1. Фундаментальные аналитики . . . . .	12
3.1.2. Технические аналитики . . . . .	13
3.2. Тестирование . . . . .	14
<b>4. Внедрение в базовую систему других подходов</b>	<b>18</b>
4.1. Механизм ордеров . . . . .	18
4.2. Гетерогенность внутри групп агентов . . . . .	19
<b>5. Анализ работы итоговой системы</b>	<b>21</b>
<b>6. Практическое применение</b>	<b>25</b>
<b>Заключение</b>	<b>29</b>
<b>Приложение А. Исходный код базовой модели (язык R)</b>	<b>30</b>
<b>Приложение В. Исходный код усовершенствованной модели (язык R)</b>	<b>32</b>
<b>Список литературы</b>	<b>38</b>

# Введение

Понимание поведения инвесторов на финансовом рынке в целом и на фондовом рынке в частности всегда было и остается важной и актуальной проблемой для стабильности современной глобальной экономики. Существует немало примеров, вошедших в историю, когда неправильное представление или игнорирование тех или иных событий влекло к значительным убыткам не только для участников торгов, но и для компаний, акции которых обращаются на биржевом рынке, и даже для людей, которые никаким прямым образом не участвуют в данной системе. Примером может послужить «крах Уолл-стрит» 1929 года, ставший началом Великой депрессии, когда из-за спекулятивного бума образовался экономический пузырь, что привело к потере у Промышленного Индекса Доу — Джонса (DJIA) 40% его стоимости за несколько месяцев. Другой пример: глобальный финансовый кризис 2007-2008 годов, который привел к спаду на рынке жилья, выселениям, падению бизнеса, сокращениям, безработице, а также огромному снижению потребительского благосостояния.

В этой работе рассмотрены основные определения и понятия, связанные с фондовым и финансовым рынком (Раздел 1); основные подходы, используемые в моделировании фондового рынка (Раздел 2); описание, реализация, тестирование базовой системы (Раздел 3), а также описание идей по доработке базовой системы (Раздел 4) и тестирование итоговой модели (Раздел 5).

# 1. Финансовый рынок

В данном разделе будет представлен краткий обзор основных определений и понятий, связанных с финансовым рынком, перечислены его основные функции, а также различные подходы к классификации рынка ценных бумаг. За более подробной информацией можно обратиться к книге [9].

## 1.1. Определение

**Финансовый рынок** – организованная или неформальная система торговли финансовыми инструментами. На финансовом рынке происходит мобилизация капитала, предоставление кредита, осуществление обменных денежных операций и размещение финансовых средств в производстве. А совокупность спроса и предложения на капитал кредиторов и заёмщиков разных стран образует мировой финансовый рынок.

**Фондовый рынок** – это организованный и регулируемый финансовый рынок, где продаются и покупаются ценные бумаги (деньги, акции, имущество и другие ресурсы). Сама биржа является лишь торговой площадкой и ее задача – проведение сделок. Если раньше было необходимо физическое взаимодействие трейдеров, то современная биржа оперирует в онлайн-режиме, что зародило понятие *высокочастотной алгоритмической торговли*. Но, несмотря на интеграцию компьютерных технологий в данный процесс, основные концепции и стратегии трейдеров остались прежними за тем лишь исключением, что торги происходят автоматизированно и быстро.

Всех участников рынка ценных бумаг можно разделить на три категории:

1. Основные участники рынка ценных бумаг (эмитенты, инвесторы)
2. Фондовые посредники (брокеры, управляющие компании)

3. Организации инфраструктуры рынка (органы государственного регулирования рынка)

## 1.2. Классификация

Существует множество способов классификации фондовых рынков. Некоторые из возможных критериев:

- По виду ценных бумаг (рынок производных финансовых инструментов, рынок акций, рынок облигаций)
- По территориальному принципу (региональные, национальные и международные рынки)
- По характеру движения ценных бумаг:
  1. Первичный рынок – рынок, на котором происходит первичное размещение вновь выпущенных ценных бумаг.
  2. Вторичный рынок – рынок, на котором производятся сделки с ранее выпущенными и прошедшими процедуру первичного размещения ценными бумагами.
  3. Третий рынок – рынок, на котором происходит торговля не прошедшими процедуру листинга ценными бумагами.
  4. Четвёртый рынок – это электронные системы торговли крупными пакетами ценных бумаг напрямую между институциональными инвесторами.
- По видам сделок (рынок — подразумевающий мгновенное исполнение сделок; форвардный рынок и т. д.).

### 1.3. Функции

Среди множества функций рынка ценных бумаг можно выделить основные, а именно:

- перераспределение денежных средств между странами и территориями, отраслями промышленности и секторами экономики, отдельными предприятиями внутри отдельного сектора.
- фиксация прав владельцев на долю долга или собственности предприятий и на долю долга государства (в случае государственных бумаг). Вознаграждение за участие в финансировании и взятый на свои плечи риск выражаются в виде процентов по облигациям и дивидендов по акциям, а также в росте курсовой стоимости последних.
- централизация капитала, то есть объединение частных индивидуальных капиталов в один общий.
- стимулирующая функция, то есть мотивация юридических и физических лиц стать участниками фондового рынка.

## 2. Моделирование фондового рынка

В рамках данной работы была поставлена задача формирования системы, имитирующей процесс ценообразования на рынке. Такую систему можно использовать для тестирования различных торговых стратегий, а также для проверки различных гипотез по отношению к акциям конкретной компании или же конкретному фондовому индексу (например, DJIA) благодаря возможности гибкого подбора параметров системы.

*Котировки* – это отношение стоимости одного актива к другому. Котировки образуют временной ряд, который используется при построении графиков (Рис. 1). Изменение котировок происходит из-за совершенных торговых операций на финансовом рынке, т.е. цена формируется по закону спроса и предложения.

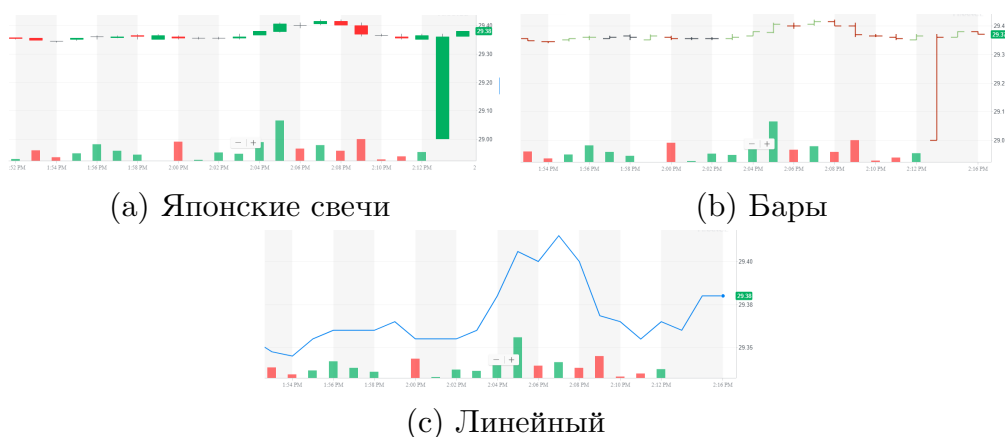


Рис. 1: –Основные типы графиков, которые используют участники рынка

В этой главе мы рассмотрим самые главные концепции, наблюдаемые во многих статьях, посвященных тематике моделирования фондового рынка. С подробным описанием и обоснованием данных подходов можно ознакомиться в [4],[6] и [8].



## 2.1. Агентно-ориентированное моделирование

В последнее время все чаще в литературе по моделированию поведения трейдеров на финансовом рынке встречаются такие понятия, как **агентно-ориентированная модель (АОМ)** или же **мультиагентная система (МАС)**. Действительно, рассмотрение участников финансового рынка как ограниченно-рациональных адаптивных агентов является гибкой методологией моделирования таких сложных систем и их поведения. С АОМ финансовый рынок может быть смоделирован снизу вверх, создавая макроструктуру за счет разработки микроструктуры, принимая во внимание эволюцию и взаимодействие агентов поэтапно. Данный подход, действительно, с интуитивной точки зрения можно назвать естественным, т.к. он принимает во внимания неоднородные особенности трейдеров.

В литературе, среди моделей финансового рынка на основе агентов, модели двух типов агентов является простейшим видом по отношению гетерогенности. В таких моделях финансовые агенты следуют одной из двух общих систем веры: *фундаментальной* и *технической*.

*Фундаментальные аналитики* принимают решения, предполагая, что цена актива возвращается к своей фундаментальной ценности в долгосрочной перспективе и делают выводы исключительно на основе новостей и авторитетности компании, в то время как *технические аналитики* (также называемые чартистами от англ. chart – график) в основном обеспокоены трендом и закономерностями, наблюдаемыми в прошлых ценах. Доказательства существования этих двух типов трейдеров на реальных рынках можно проследить в данных опроса [2].

## 2.2. Стилизованные факты

Целью исследования характеристик современного финансового рынка является перечисление так называемых *стилизованных фактов*, применимых к фондовому рынку.

**Стилизованный факт** - это термин, используемый в экономике для обозначения эмпирических данных (статистик), которые настолько согласованы (например, по широкому спектру инструментов, рынков и периодов времени), что воспринимаются за действительность. Из-за их общности они используются как качественные характеристики при тестировании финансовых моделей.

Стоит отметить, что далее будут перечислены характеристики, относящиеся не к прямым значениям цены актива на временном интервале, а к *логарифмической доходности* актива, которая рассчитывается как разность натуральных логарифмов последовательностей цены:

$$r_t = \ln(p_t) - \ln(p_{t-1}) \quad (1)$$

Из работы [6] можно выделить некоторые из них, по которым будет проводиться тестирование системы:

- **Зависимость:** Автокорреляция самого ряда незначительно мала, в то время как автокорреляция квадратов или абсолютных значений ряда может быть существенна для большого числа шагов.
- **Распределение:** приблизительно симметричное, коэффициент эксцесса положительный, ”тяжелые хвосты”
- **Кластеризация волатильности:** т.е. большие изменения во временном ряде следуют за большими изменениями, а малые – за малыми

- *Среднее значение и стандартное отклонение* положительные, принимают незначительно маленькие значения

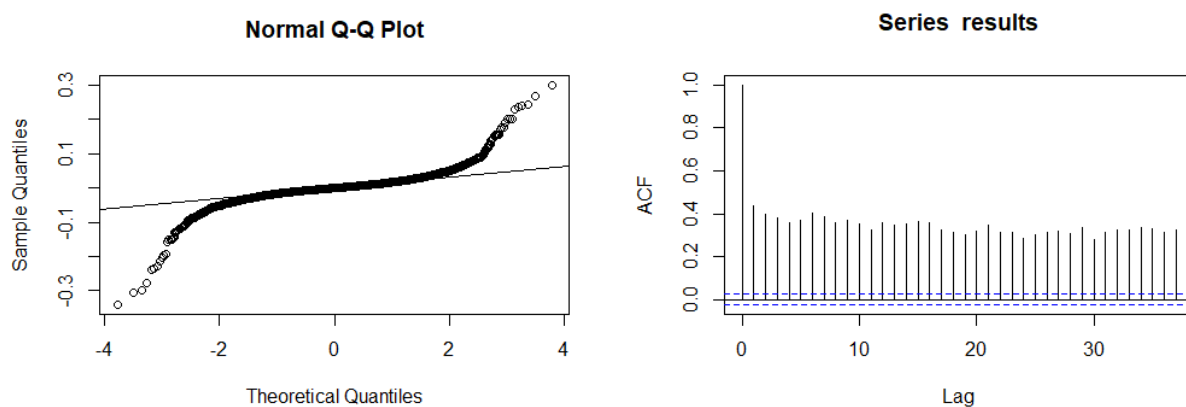


Рис. 2: – Пример наличия тяжелых хвостов и кластеризации волатильности в данных о доходности ВАС (Bank of America Corp), взятых из [1]

### 2.3. Гипотеза эффективного рынка

Популярной среди аналитиков является гипотеза об эффективном рынке.

**Эффективный рынок** – рынок, на котором стоимость ценных бумаг мгновенно реагирует на новую информацию. Специалисты обычно говорят о трех уровнях эффективности финансового рынка: слабой, эффективной в умеренной форме и эффективной в сильной форме. Исследования показывают, что большинство так называемых эффективных финансовых рынков характеризуются умеренной, а не сильной формой эффективности, т.е. текущие цены полностью отражают только доступную информацию.

## 3. Базовая система

Большинство статей, посвященных построению тех или иных систем, моделирующих фондовый рынок, не предоставляют исходный код программной реализации. Убедиться в этом можно как минимум по анализу, проведенному в [8]. Поэтому на начальном этапе было принято решение реализовать одну из уже описанных в литературе систем в программном виде и проверить её практическую обоснованность.

В качестве такой статьи была выбрана докторская работа [6], раздел Modeling. Помимо обширной исследовательской работы с литературой, обзора характеристик и построения прогнозирования и оценки рисков, в ней была описана модель реализации фондового рынка. Как утверждает автор, данная модель не претендует на высокую научность и реализм, но при этом является простой в реализации и отражает некоторые стилизованные факты при определенных значениях свободных параметров. Было решено взять данную модель как основу для программной реализации и дальнейшей доработки, предварительно протестировав её на предмет присутствия или же отсутствия выбранных в Разделе 2.2 стилизованных фактов.

### 3.1. Описание

В системе существует 2 типа агентов: фундаментальные аналитики и технические аналитики, переменные  $n_f$  и  $n_t$  определяют их соотношение в системе, соответственно.

$$n_t + n_f = 1 \tag{2}$$

#### 3.1.1. Фундаментальные аналитики

Фундаменталисты принимают решение о цене актива на основе новостей, связанных с организацией, которая эмитирует ценные бумаги для развития и финансирования своей деятельности. Новости по своей

сущности непредсказуемы, можно представить их движение как случайное геометрическое блуждание.

$\ln(V_t)$  – переменная, имитирующая процесс случайного блуждания:

$$\ln\left(\frac{V_t}{V_{t-1}}\right) \sim \mathcal{N}(0.0013, 0.023) \quad (3)$$

$$\ln\left(\frac{V_t}{V_{t-1}}\right) > 0 \text{ означает "хорошие" новости} \quad (4)$$

$$\ln\left(\frac{V_t}{V_{t-1}}\right) < 0 \text{ означает "плохие" новости} \quad (5)$$

В таком случае фундаментальная доходность актива в момент времени между  $t$  и  $t + 1$ :

$$F_t = r \ln\left(\frac{V_t}{P_t}\right) \quad (6)$$

где  $P_t$  – цена актива на данный момент времени  $t$ , а  $r$  принимает значение согласно следующему правилу:

$$\left\{ \begin{array}{ll} 1.1 & \text{если } V_t > V_{t-1} > V_{t-2} > V_{t-3} \text{ или} \\ & V_t < V_{t-1} < V_{t-2} < V_{t-3} ; \text{ иначе} \\ 1 & \text{если } V_t > V_{t-1} > V_{t-2} \text{ или} \\ & V_t < V_{t-1} < V_{t-2} ; \text{ иначе} \\ 0.9 & \end{array} \right. \quad (7)$$

### 3.1.2. Технические аналитики

Чартисты принимают решение на основе тренда, т.е. исключительно на основе наблюдаемых цен в прошлом. В данной системе чартисты смотрят на предыдущие три дня и вычисляют техническую доходность актива в момент времени между  $t$  и  $t + 1$  по следующей формуле, по сути своей представляющей собой скользящее окно:

$$T_t = c^3 \ln\left(\frac{P_{t-2}}{P_{t-3}}\right) + c^2 \ln\left(\frac{P_{t-1}}{P_{t-2}}\right) + c \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (8)$$

где  $c = 0.544$  – корень уравнения  $c^3 + c^2 + c = 1$ , цель которого – придать ценам различную значимость: чем цена новее, тем больше её вес.

Вместо цены, как было уже указано в Разделе 2.2, по ряду причин принято рассматривать доходность, т.е. логарифмическое приращение актива. Тогда новое значение доходности актива вычисляется маркет-мейкером по следующей формуле:

$$\ln \frac{P_{t+1}}{P_t} = n_f F_t + n_t T_t \quad (9)$$

## 3.2. Тестирование

Описанная в Разделе 3.1 система была реализована в виде программного модуля в среде RStudio (Приложение А).

RStudio – свободная среда разработки ПО с открытым исходным кодом на языке R, который предназначен для удобной и гибкой статистической обработки данных и работы с графикой. Базовых средств языка, а также дополнительных пакетов, вполне достаточно для аналитической работы в рамках поставленной задачи.

В процессе тестирования проводился многократный запуск симуляции 1000 дней торгов при различном соотношении фундаментальных и технических аналитиков. Чтобы наглядно оценить значения различных статистических показателей, производились визуальный анализ статистик и сопоставление стилизованным фактам.

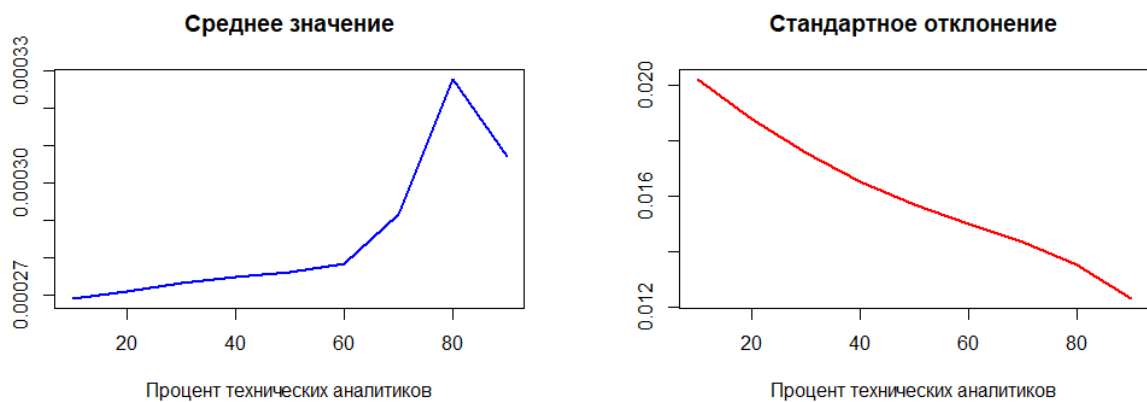


Рис. 3: – Среднее значение и стандартное отклонение доходности реалистичны при 30%-70% технических агентов в системе

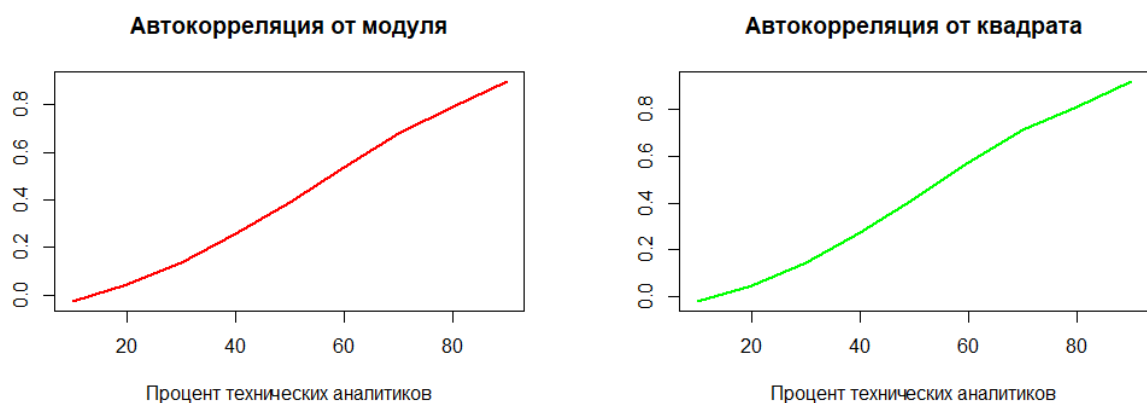


Рис. 4: – Автокорреляция квадратов и абсолютных значений доходности реалистичны при 30%-50% технических агентов в системе

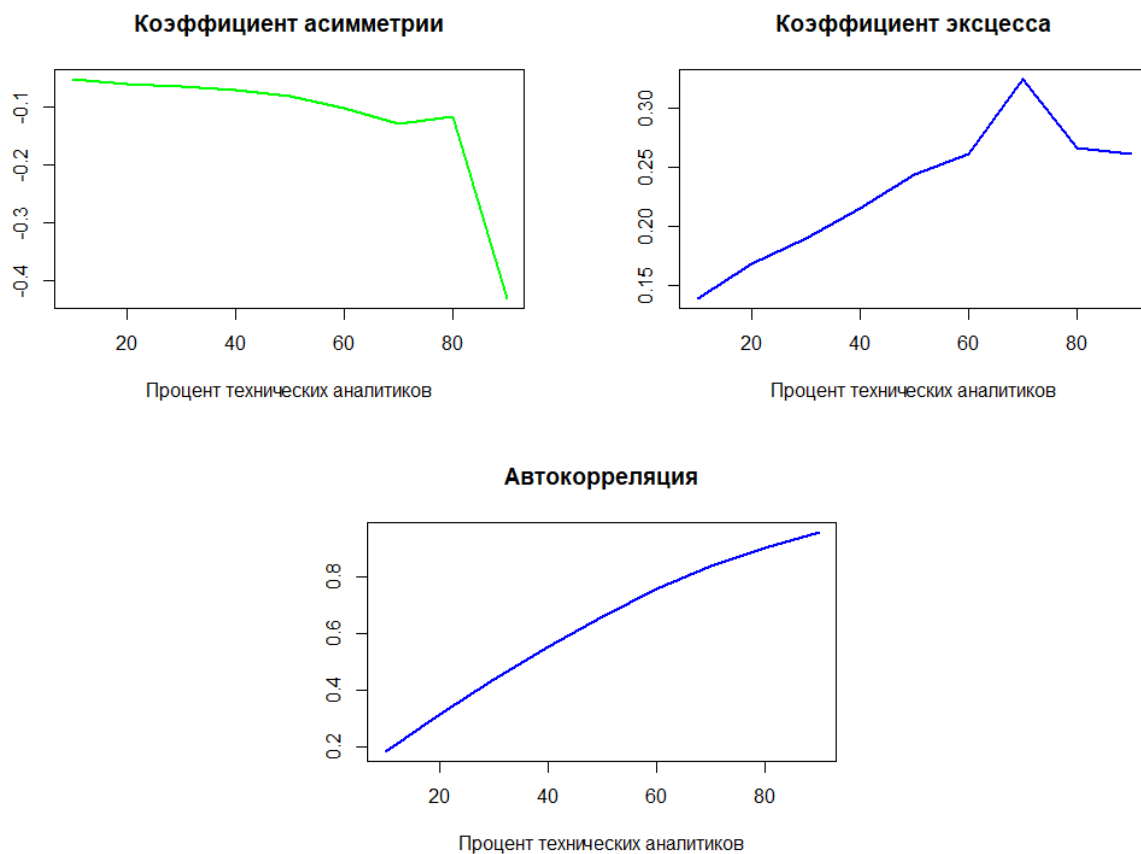


Рис. 5: – Статистические показатели для коэффициента асимметрии, коэффициента эксцесса и автокорреляции не отражают стилизованные факты ни при каком проценте технических аналитиков в системе

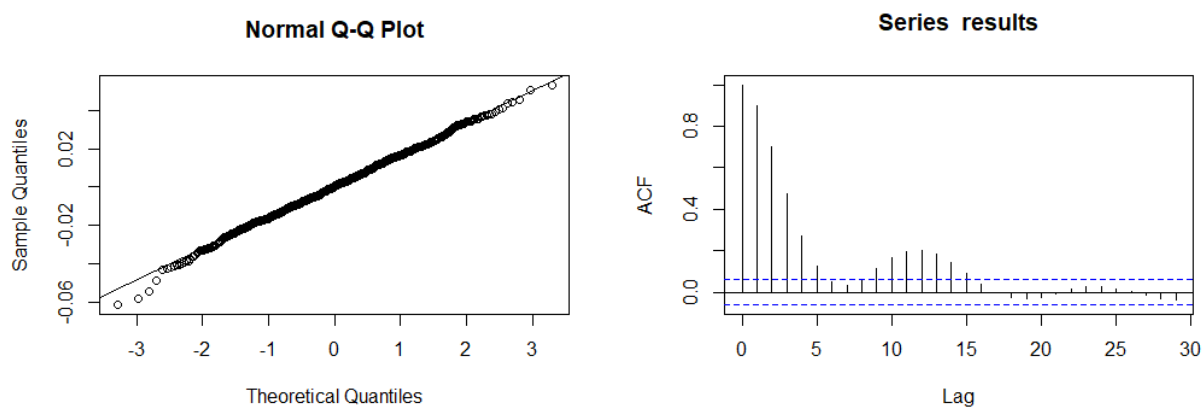


Рис. 6: – Тестирование распределения доходности на наличие ”тяжелых хвостов” и кластеризацию волатильности (Квантиль-Квантиль график и график автокорреляции с различными значениями лагов)



Таким образом, несмотря на то, что система смогла смоделировать временной ряд котировок, в котором присутствуют статистики, отражающие такие стилизованные факты, как среднее значение, стандартное отклонение, автокорреляция квадратов и абсолютных значений при 30%-50% технических аналитиков в модели, авторам не удалось добиться значений автокорреляции, асимметрии и эксцесса, которые можно было бы признать корректными с точки зрения эмпирических выкладок. Как результат, на Рисунке 6 не наблюдаются ни ”тяжелые хвосты”, ни кластеризации волатильности в распределении доходности.

## 4. Внедрение в базовую систему других подходов

Представленная модель в Разделе 3 имеет небольшое количество свободных параметров, поэтому недостаточно гибкая для аналитической работы с моделью.

В этом разделе представлены две доработки, которые в теории могли бы существенно улучшить показатели итоговой системы по сравнению с базовой, прежде всего, отразить все описанные в Разделе 2.2 стилизованные факты. Данные подходы были взяты из описанных ранее систем в [3] и [7].

Исходный код доработанной системы представлен в Приложении В.

### 4.1. Механизм ордеров

В базовой системе нет никакого механизма ордеров на куплю и продажу, как и понятия портфеля инвестора (количество акций и средств на счету). В доработанной версии системы можно попробовать реализовать простейшую версию этой структуры.

Добавим агентам возможность купить\продать одну единицу актива, причем это происходит, если текущая цена ниже\выше той, которую агент ожидает в будущем (поэтому также необходимо ввести дополнительную переменную, определяющую математическое ожидание цены на следующий день).

Поскольку фундаменталисты гораздо реже участвуют в торгах, вводится специальный параметр  $K$ . Фундаменталист становится активным, только если разница между текущей ценой и ценой, которую он ожидает, больше данного параметра.

Впервые было обнаружено в [5], что доходность населения распре-

делена по степенному закону:

$$W(x) = x^{-(1+\alpha)} \quad (10)$$

Это наблюдение можно использовать при начальном распределении средств между агентами.

## 4.2. Гетерогенность внутри групп агентов

Несмотря на то, что есть две группы трейдеров и система позволяет управлять их соотношением, внутри групп трейдеры ведут себя абсолютно идентично. Предлагается увеличить гетерогенность внутри групп.

Чтобы добавить неоднородность среди технических аналитиков, предлагается в определение  $T_t$  ввести различные значения длины памяти  $M$  и равномерно распределить их между агентами:

$$T_t = c \sum_{i=1}^M c^{i-1} \ln\left(\frac{P_{t-i+1}}{P_{t-i}}\right) \quad (11)$$

где сглаживающий параметр  $c$  определяется для каждого чартиста в отдельности в соответствии со значением длины памяти.

Чтобы добавить неоднородность среди фундаментальных аналитиков, в определение  $F_t$  предлагается добавить шумовой параметр  $\eta$  – независимая случайная величина с нулевым средним значением.

$$F_t = r \ln\left(\frac{V_t}{T_t}\right) + \eta \quad (12)$$

Гетерогенность в предсказании цены актива на следующий день можно реализовать через параметры  $react_f$  и  $react_t$ , которые будут означать оценку скорости, с которой цена стремится к своему фундаментальному значению у фундаментальных аналитиков, и чувствительность технических аналитиков к изменениям цены, соответственно.

Дополнительную гетерогенность можно придать поведению агентов, если ввести шумовые параметры  $\sigma$ ,  $\tau$  и  $\beta$  в определении цены маркет-мейкером и математического ожидания цены у фундаментальных и технических аналитиков:

$$E_t^f(P_{t+1}) = P_t + react_f * (F_t - P_t) + \tau \quad (13)$$

$$E_t^t(P_{t+1}) = P_t + react_t * (P_t - T_t) + \beta \quad (14)$$

$$P_{t+1} = P_t * (1 + a * (D_t - S_t)) + \sigma \quad (15)$$

где  $a$  – коэффициент, регулирующий цену, а  $D_t$  и  $S_t$  – количество ордеров на покупку и продажу. Будем считать, что изначально ни один агент не обладает ни одним активом (поэтому может только приобрести его у эмитента через биржу) и количество выпускаемых акций не ограничено.

## 5. Анализ работы итоговой системы

После внедрения в систему, описанную в Разделе 3, структур из Раздела 4, в ней, помимо параметров  $n_f$  и  $n_t$ , появились другие свободные параметры, благодаря которым можно наделять систему новыми свойствами и корректировать поведение агентов в модели.

Комбинируя различные значения свободных параметров, помимо стилизованных фактов, можно также добиться численно близких статистических показателей, используя для анализа различные меры близости (метрики расстояния). В качестве примера в этом разделе параметры подбирались таким образом, чтобы статистические показатели были близки к показателям для распределения ежедневной доходности Banc of America Corp (Таблицы 1 и 2).

Таблица 1: Основные статистики ВАС

Среднее значение	6.79998e-05
Стандартное отклонение	0.02760361
Коэффициент асимметрии	-0.3184218
Коэффициент эксцесса	26.22164
Автокорреляция	0.4365003
Автокорреляция абсолютных значений	0.3910598
Автокорреляция квадратов значений	0.3265301

Как и в Разделе 3.2, для выявления присутствия или отсутствия стилизованных фактов были построены графики всех перечисленных ранее статистик при различном проценте технических аналитиков в модели и многократном запуске системы.

Таблица 2: Параметры модели

Параметр, регулирующий цену ( $a$ )	$0.2 * 10^{-4}$
Параметр активации фундаменталистов ( $K$ )	5%
Параметр в степенном законе ( $\alpha$ )	1.5
Сглаживающий параметр $c$	[0.3,0.99]
Коэффициент реакции фундаменталистов( $react_f$ )	$\mathcal{N}(0.04, 0.04)$
Коэффициент реакции чартистов( $react_t$ )	$\mathcal{N}(0.05, 0.04)$
Стандартное отклонение шумового параметра $\eta$	0.026
Стандартное отклонение шумового параметра $\sigma$	0.025
Стандартное отклонение шумового параметра $\tau$	0.01
Стандартное отклонение шумового параметра $\beta$	0.05
Параметр памяти ( $M$ )	$\mathcal{U}(1, 50)$

На Рисунке 7 изображены результаты тестирования итоговой системы. Можно заметить, что присутствует ярко выраженный положительный коэффициент эксцесса и небольшая отрицательная автокорреляция, что не наблюдалось в модели, построенной базовой системой. Среднее значение, стандартное отклонение, автокорреляция абсолютных значений и квадратов значений доходности сохранили свои реалистичные показатели. Коэффициент асимметрии по-прежнему сохраняет своё отрицательное значение, что с одной стороны не совпадает со стилизованным фактом, но с другой стороны наблюдается в статистике ВАС (Таблица 1). При необходимости можно скорректировать асимметрию так, чтобы она тоже стала отражать стилизованный фактор.

Положительный результат можно заметить также и на Рисунке 8, где отчетливо видны ”тяжелые хвосты” и кластеризация волатильности (Рисунок 2).

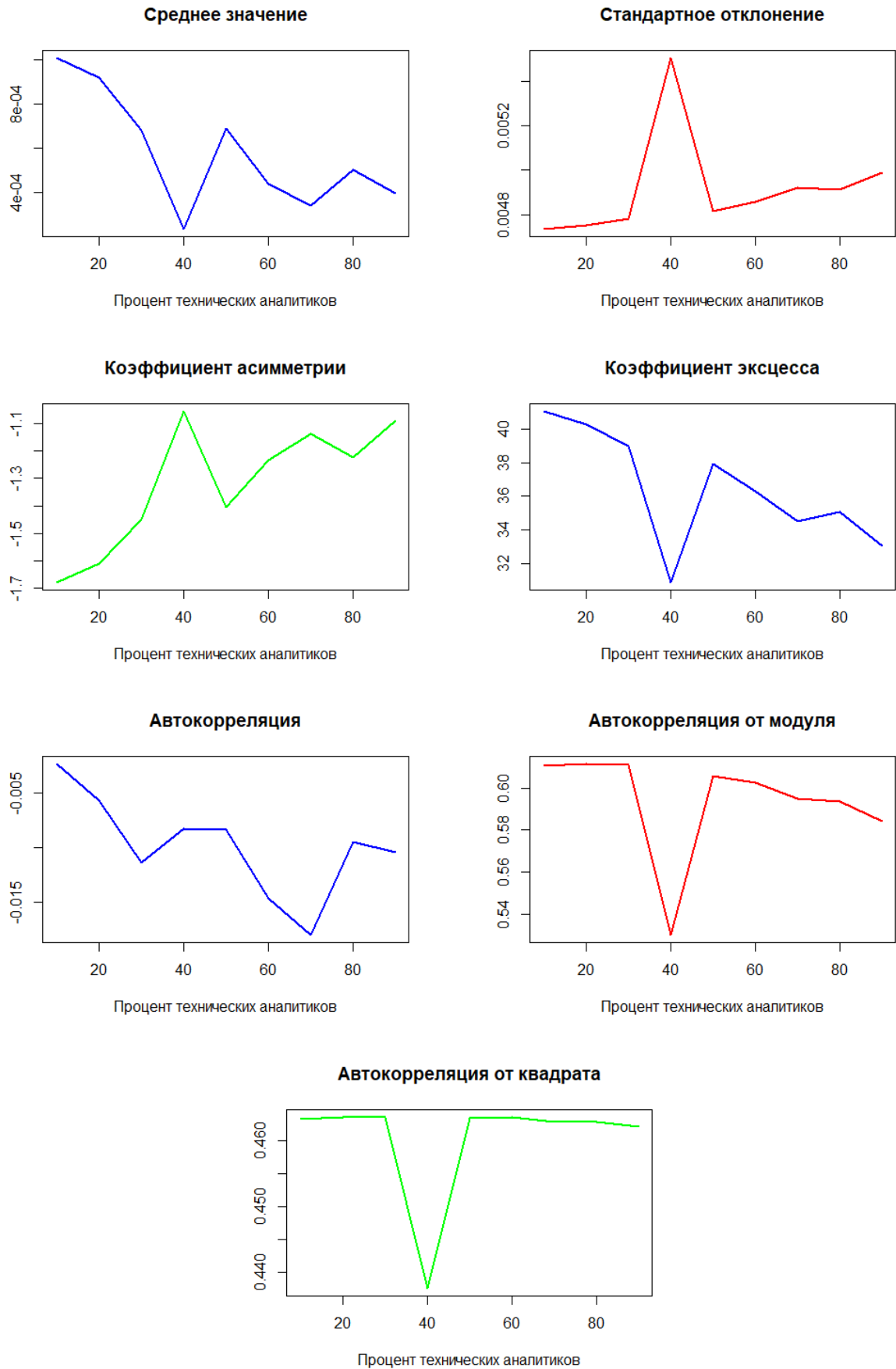


Рис. 7: – Статистические показатели итоговой системы

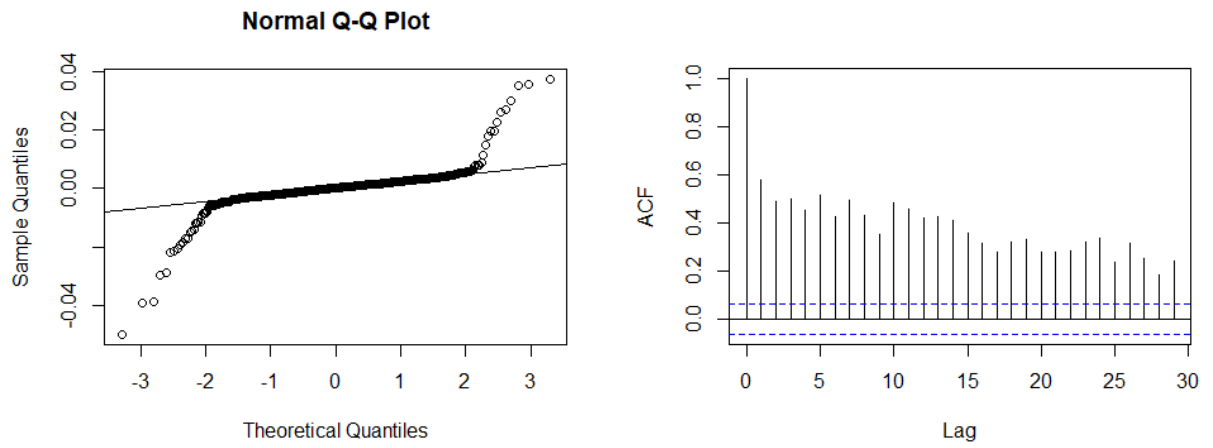


Рис. 8: – Тестирование распределения доходности на наличие ”тяжелых хвостов” и кластеризации волатильности в итоговой модели



## 6. Практическое применение

В этом разделе представлен результат, полученный при помощи практического применения построенной системы. В Листинге 1 описан алгоритм, при помощи которого, в отличие от визуального анализа графиков, можно с высокой точностью вычислять оптимальное соотношение технических аналитиков к фундаментальным аналитикам в модели.

Листинг 1: Псевдокод системы поиска оптимального соотношения чартистов

```
vector <double> getStatistic (vector <double> ts)
{
    . . .
    return ({tech_proc, mean, sd, skewness,
            kurtosis, autocorrelation,
            sqr_autocorrelation,
            abs_autocorrelation });
}

vector <double> newModel (tech_proc)
{
    . . .
    return (ts);
}

vector <double> techProcVector ()
{
    vector <double> data_statistics =
        getStatistic (orig_data);
    vector <vector<double>> model_statistics;

    for (int i = 0; i <= 100; ++i){
```

```

        // заполнение model_statistics через
        // функцию getStatistic(newModel(i*0.01))
    }

    // количеством метрик, по которым i-ая статистика
    // находится ближе к data_statistics

    int succ_metrics_count[101];
    for (int i = 0; i <= 100; ++i){
        // вычисление succ_metrics_count[i]
    }

    int max_count = max(succ_metrics_count);
    vector<double> res;
    for (int i = 0; i <= 100; ++i){
        if (succ_metrics_count == max_count)
            res.append(i*0.01);
    }

    return res;
}

```

В случае ВАС, можно заметить, что наиболее удачные статистические показатели получаются, если 40% участников рынка – это технические аналитики. Этот процент попадает в диапазон [30%-50%], который был получен при анализе результатов базовой системы.

В качестве примера работы описанного алгоритма зафиксируем  $n_t = 0.4$ , оценим статистические показатели у обеих систем (Таблица 3) и сравним их с показателями для ВАС по различным метрикам расстояния при помощи пакета *philterropy* в RStudio (Таблица 4).

Таблица 3: Статистики при 40% технических специалистов в модели

Статистика	Базовая модель	Итоговая модель
Среднее значение	2.747364e-04	5.53097e-05
Стандартное отклонение	0.01653036	0.004983651
Коэффициент асимметрии	0.1573943	-0.9796748
Коэффициент эксцесса	0.2153644	32.03579
Автокорреляция	0.5522113	-0.01685278
Автокорреляция по модулю	0.26030875	0.5788605
Автокорреляция квадратов	0.27409132	0.4374577

Таблица 4: Сравнение результатов базовой системы и доработанной системы по различным мерам расстояния

№	Мера расстояния	Итоговая модель	Базовая модель
1	euclidean	<b>5.86921088651282</b>	26.0108877826056
2	manhattan	<b>6.9513887291</b>	26.6090826866
3	chebyshev	<b>5.81415</b>	26.0062756
4	sorensen	<b>0.12107967120687</b>	0.974364558040718
5	gower	<b>1.39027774582</b>	5.32181653732
6	soergel	<b>0.216005471005508</b>	0.987015851831983
7	kulczynski_d	<b>0.27551910506639</b>	76.0169892595044
8	canberra	1.46776535490475	<b>-1.00003996741987</b>
9	lorentzian	<b>2.82282743655568</b>	3.80599136905936
10	intersection	25.2301513807	<b>0.3500412598</b>
11	non-intersection	-24.2301513807	<b>0.6499587402</b>
12	wavehedges	<b>0.149516841482527</b>	5.37805435484683
13	czekanowski	<b>0.12107967120687</b>	0.974364558040718
14	motyka	<b>0.560539835603435</b>	0.987182279020359
15	kulczynski_s	3.62951236996446	<b>0.0131549540404216</b>
16	tanimoto	<b>0.216005471005508</b>	0.987015851831983
17	ruzicka	0.783994528994492	<b>0.0129841481680172</b>
18	inner_product	<b>840.335</b>	6836358245.838

№	Мера расстояния	Итоговая модель	Базовая модель
19	harmonic_mean	28.3314043211819	<b>1.55806639971235</b>
20	cosine	0.999682432115264	<b>0.362869611593447</b>
21	hassebrook	0.960621521192523	<b>0.00855589835429207</b>
22	jaccard	<b>0.0393784788074768</b>	0.991444101645708
23	dice	<b>0.0200846917071101</b>	0.983033368069627
24	squared_euclidean	<b>34.4476364303607</b>	676.566283239301
25	pearson	<b>-11.48400622297</b>	3141.85171902538
26	neyman	<b>0.405376198231118</b>	25.1174146143965
27	squared_chi	<b>0.748882848136174</b>	24.1930324067753
28	prob_symm	<b>1.49776569627235</b>	48.3860648135506
29	divergence	<b>3.85790727746621</b>	20.2789131023337
30	clark	<b>1.38886775422756</b>	3.18425133291436
31	additive_symm	<b>-11.078630024738</b>	3166.96913363977
32	avg	<b>6.38276936455</b>	26.3076791433

По 24 метрикам из 32, которые указаны в Таблице 4, показатели статистик итоговой модели ближе к показателям исходного индекса (Таблица 1). Таким образом, с 75% уверенностью можно утверждать, что доработанная система работает лучше, чем изначально взятая система.

## Заключение

В данной работе представлена система<sup>1</sup>, позволяющая моделировать фондовый рынок. При помощи настройки и подбора параметров, с помощью неё можно воссоздавать доходность реальных торговых индексов или отдельных компаний. Такую систему можно использовать для тестирования различных торговых стратегий, а также для проверки различных гипотез по отношению к акциям конкретной компании или же по отношению к конкретному фондовому индексу.

Практический результат работы – алгоритм вычисления процентного соотношения технических аналитиков к фундаментальным.

Данную систему можно внедрять в другие программные комплексы и улучшать за счет новых подходов в моделировании фондового рынка.

---

<sup>1</sup><https://github.com/mariapashkova/diploma>

## А. Исходный код базовой модели (язык R)

```
#parameters of the system

days_count <- 1000
c <- 0.544
mean_gaussian <- 0.0013
sd_gaussian <- 0.023
log_ratio <- rnorm(n = 3, mean = mean_gaussian, sd = sd_gaussian)
log_ratio_V <- rnorm(n = days_count, mean = mean_gaussian,
                    sd = sd_gaussian)

#modeling

basic_model <- function(x){

  nt <- x
  nf <- 1 - nt

  # experiment - data for teaching technical traders

  price_new <- experiment[4]
  price_0 <- experiment[3]
  price_1 <- experiment[2]
  price <- experiment[1]

  v_3 <- experiment[1]
  v_2 <- v_3 * exp(log_ratio[1])
  v_1 <- v_2 * exp(log_ratio[2])
  v_0 <- v_1 * exp(log_ratio[3])

  for (i in 1:days_count){

    p_3 <- p_2
    p_2 <- p_1
    p_1 <- p_0
    p_0 <- p_new

    v_3 <- v_2
```

```

v_2 <- v_1
v_1 <- v_0
v_0 <- v_1 * exp(log_ratio_V[i])

if (v_0 >= v_1 && v_1 >= v_2 && v_2 >= v_3 ||
    v_0 <= v_1 && v_1 <= v_2 && v_2 <= v_3) {

  r <- 1.1

} else if (v_0 >= v_1 && v_1 >= v_2 ||
           v_0 <= v_1 && v_1 <= v_2) {

  r <- 1

} else

  r <- 0.9

fundamental_value <- r * log(v_0 / p_0)
technical_value <- c^3*log(p_2/p_3) +
                  c^2*log(p_1/p_2) +
                  c*log(p_0/p_1)
p_new <- p_0 * exp(nf * fundamental_value +
                  nt * technical_value)
data <- rbind(data, c(log(v_0/v_1) > 0, p_new,
                    fundamental_value * nf, technical_value * nt))
}

colnames(df) <- c("day", "good_news", "price",
                 "fundamental", "technical")
df$price <- c(0, log(df$price[-1]) - log(df$price[-nrow(df)]))
return(df)
}

```

## В. Исходный код усовершенствованной модели (язык R)

```
library(actuar)

#parameters

days_count <- 1000
traders_count <- 1000
a <- 0.2 * 10(-4)
K <- 0.02
alpha <- 1.5
eta_sd <- 0.026
sigma_sd <- 0.025
tau_sd <- 0.01
beta_sd <- 0.05
mean_gaussian <- 0.0013
sd_gaussian <- 0.023

#inicialization

createTradersListMod <- function(nt){

  past_prices_vector <- head(experiment, 51)
  tech_traders_count <- round(traders_count * nt)
  fundamental_traders_count <- traders_count -
    round(traders_count * nt)
  strategy_type <- c(rep("technical", tech_traders_count),
    rep("fundamental", traders_count -
    tech_traders_count))

  while (max(cash) - min(cash) > 70 || max(cash) - min(cash) < 69){
    cash <- rpareto(traders_count, shape = alpha_const,
    scale = 0.5) + 30
  }
  f <- rnorm(fundamental_traders_count, 0.05, 0.04)
  c <- rnorm(tech_traders_count, 0.04, 0.04)
  type_coef <- c(c,f)
  technical_memory <- round(runif(traders_count, min = 1, max = 50))
}
```



```

smoothing_coef <- -(technical_memory - 50) * 0.69 / 49 + 0.3
technical_price <- vector(mode='numeric', length=0)
for (i in 1:tech_traders_count){
  sum <- 0
  for (j in 1:technical_memory[i]){
    sum <- sum + (smoothing_coef[i]^(j - 1)) *
      log(past_prices_vector[52 - j]/
          past_prices_vector[52 - j - 1])
  }
  technical_price <- c(technical_price, sum * smoothing_coef[i])
}
fundamental_price <- rep(experiment_data[50] +
  rnorm(1, 0, eta_sd),
  fundamental_traders_count)
type_price <- c(technical_price, fundamental_price)
type_expect <- rep(NaN, traders_count)
asset_amount <- rep(0, traders_count)
df_traders <- data.frame(strategy_type, cash, type_coef,
  technical_memory, smoothing_coef,
  type_price, type_expect, asset_amount)
return(df_traders)
}

#modeling

new_model <- function(nt){

  past_prices_vector <- head(experiment_data, 51)
  traders <- createTradersListMod(nt)
  iid_eta <- rnorm(days_count - 50, mean = 0, sd = eta_sd)
  iid_beta <- rnorm(days_count - 50, mean = 0, sd = beta_sd)
  iid_sigma <- rnorm(days_count - 50, mean = 0, sd = sigma_sd)
  iid_tau <- rnorm(days_count - 50, mean = 0, sd = tau_sd)
  tech_traders <- head(traders, round(nt*traders_count))
  tech_traders_count <- nrow(tech_traders)
  fund_traders <- tail(traders, traders_count -
    round(nt*traders_count))
  fund_traders_count <- nrow(fund_traders)
  log_ratio_V <- rnorm(n = days_count-50, mean = mean_gaussian,
    sd = sd_gaussian)
}

```

```

log_ratio <- rnorm(n = 3, mean = mean_gaussian, sd = sd_gaussian)
v_3 <- v_3 <- past_prices_vector[1]
v_2 <- v_3 * exp(log_ratio[1])
v_1 <- v_2 * exp(log_ratio[2])
v_0 <- v_1 * exp(log_ratio[3])

for (day in 51:days_count){

  buy_orders <- 0
  sell_orders <- 0

  for (i in 1:tech_traders_count){

    tech_traders$type_price[i] <-
      tech_traders$type_price[i] *
      tech_traders$smoothing_coef[i] -
      log(past_prices_vector[day-tech_traders$technical_memory[i]]/
          past_prices_vector[day-tech_traders$technical_memory[i]-1])*
          (tech_traders$smoothing_coef[i] ^
            (tech_traders$technical_memory[i] + 1)) +
      tech_traders$smoothing_coef[i] *
          log(past_prices_vector[day - 1] /
              past_prices_vector[day - 2])
  }

  v_3 <- v_2
  v_2 <- v_1
  v_1 <- v_0
  v_0 <- v_1 * exp(log_ratio_V[day-51])

  if (v_0 >= v_1 && v_1 >= v_2 && v_2 >= v_3 ||
      v_0 <= v_1 && v_1 <= v_2 && v_2 <= v_3) {

    r <- 1.1

  } else if (v_0 >= v_1 && v_1 >= v_2 ||
             v_0 <= v_1 && v_1 <= v_2) {

    r <- 1
  }
}

```

```

} else

  r <- 0.9

fund_traders$type_price <- r *
  log(v_0 / past_prices_vector[day-1]) +
  iid_eta[day - 51]

fundamental_new_comp <- fund_traders$type_coef *
  (fund_traders$type_price - past_prices_vector[day-1]) +
  iid_tau[day - 51]
fund_traders$type_expect <- past_prices_vector[day-1] +
  fundamental_new_comp

technical_new_comp <- tech_traders$type_coef *
  (past_prices_vector[day-1] - tech_traders$type_price)
+ iid_beta[day - 51]
tech_traders$type_expect <- past_prices_vector[day-1] +
  technical_new_comp

#making of the decicion: "sell", "buy" or "no action"

tech_decicions <- rep("no_action", tech_traders_count)
fund_decicions <- rep("no_action", fund_traders_count)
for (i in 1:tech_traders_count){

  if (tech_traders$type_expect[i] < past_prices_vector[day-1] &&
      tech_traders$asset_amount > 0) {

    tech_decicions[i] <- "sell"
    sell_orders <- sell_orders + 1
  }

  if ( tech_traders$type_expect[i] > past_prices_vector[day-1] &&
      past_prices_vector[day-1] < tech_traders$cash[i]){

    tech_decicions[i] <- "buy"
    buy_orders <- buy_orders + 1
  }
}
}

```

```

for (i in 1:fund_traders_count){

  if (fund_traders$type_expect[i] < past_prices_vector[day-1] &&
      ((past_prices_vector[day-1] - fund_traders$type_expect[i])/
       past_prices_vector[day-1]) > K &&
       fund_traders$asset_amount > 0) {

    fund_decicions[i] <- "sell"
    sell_orders <- sell_orders + 1
  }

  if ( fund_traders$type_expect[i] > past_prices_vector[day-1] &&
      ((fund_traders$type_expect[i] - past_prices_vector[day-1])/
       past_prices_vector[day-1]) > K &&
       past_prices_vector[day-1] < fund_traders$cash[i]){

    fund_decicions[i] <- "buy"
    buy_orders <- buy_orders + 1
  }
}

#succeeding of the orders

for (i in 1:tech_traders_count){

  if (tech_decicions[i] == "buy"){

    tech_traders$asset_amount[i] <- tech_traders$asset_amount[i]+1
    tech_traders$cash[i] <- tech_traders$cash[i] -
                                past_prices_vector[day-1]
  }

  if (tech_decicions[i] == "sell"){

    tech_traders$asset_amount[i] <- tech_traders$asset_amount[i]-1
    tech_traders$cash[i] <- tech_traders$cash[i] +
                                past_prices_vector[day-1]
  }
}
}

```

```

for (i in 1:fund_traders_count){

  if (fund_decisions[i] == "buy"){

    fund_traders$asset_amount[i] <-
      fund_traders$asset_amount[i] + 1
    fund_traders$cash[i] <- fund_traders$cash[i] -
      past_prices_vector[day-1]
  }

  if (fund_decisions[i] == "sell"){

    fund_traders$asset_amount[i] <-
      fund_traders$asset_amount[i] - 1
    fund_traders$cash[i] <- fund_traders$cash[i] +
      past_prices_vector[day-1]
  }
}

#market-maker making new price

new_price <- past_prices_vector[day-1] * ((1 + a *
      (buy_orders - sell_orders)) +
      iid_sigma[day - 51])
past_prices_vector <- c(past_prices_vector, new_price)
}
return(past_prices_vector)
}

```

## Список литературы

- [1] Yahoo finance.
- [2] Frankel J., Froot. K.A. Understanding the U.S. Dollar in the Eighties: The Expectations of Chartists and Fundamentalists // Economic Record. — 1986. — P. 24–38.
- [3] Hessary Yasaman Kamyab, Hadzikadic Mirsad. Role of Behavioral Heterogeneity in Aggregate Financial Market Behavior: An Agent-Based Approach // International Conference on Computational Science. — 2017. — P. 12–14.
- [4] McDonnell Philip J. Optimal Portfolio Modeling: Models to Maximize Return and Control Risk in Excel and R. — Wiley, 2015. — ISBN: 9780470117668.
- [5] Pareto Vilfredo. Cours d'`economie politique // Librairie Droz. — 1964. — Vol. 1.
- [6] Sewell Martin Victor. Application of Machine Learning to Financial Time Series Analysis. — 2017. — University College London.
- [7] W.B. Arthur J.H. Holland Blake Lebaron R. Palmer, Tayler P. Asset Pricing Under Endogenous Expectations in an Artificial Stock Market // The economy as an evolving complex system. — 1997. — P. 15–44.
- [8] А.С. Федеряков. Разработка моделей, алгоритмов и программных средств для повышения качества прогнозов биржевых показателей с применением мультиагентного подхода. — 2010. — Российский новый университет.
- [9] Берзон Н.И. Аршавский А.Ю. Буянова Е.А. Красильников А.С. Фондовый рынок. — Вита-Пресс, 2009. — ISBN: 9785775516215.