

МОДЕЛИРОВАНИЕ ПОВЕДЕНИЯ ТРЕЙДЕРОВ НА ФОНДОВОМ РЫНКЕ С ЦЕЛЬЮ ОБНАРУЖЕНИЯ АГРЕССИВНЫХ СТРАТЕГИЙ ТРЕЙДИНГА

М.Н. Каздагли

Рассмотрена задача построения модели, описывающей поведение совокупности трейдеров на фондовом рынке. Предполагается, что трейдеры могут вступать в сговор с целью искусственного манипулирования ценами на активы. Продемонстрирован алгоритм обнаружения целенаправленного манипулирования ценами. Дано сравнение предложенного алгоритма с другими способами детектирования согласованного воздействия брокеров на рынок. Получена качественная оценка рассмотренной модели на реальных биржевых данных.

Ключевые слова: рынок ценных бумаг, искусственное манипулирование ценами, скрытая марковская модель, анализ временных рядов.

ВВЕДЕНИЕ

Глобальный экономический кризис 2008 года и последовавшая за ним рецессия привели к повышенному вниманию со стороны государственных регулирующих структур к поведению игроков на фондовых рынках. Механизм обнаружения агрессивных стратегий трейдинга, реализуемых сразу несколькими участниками рынка, является первым шагом к глобальному анализу рынка и предотвращению финансовых кризисов.

Наряду с пассивными технологиями трейдинга, брокеры все чаще стали применять активные методы воздействия на рынок. Для этого участники рынка вступают в сговор с целью искусственного манипулирования ценами на активы. Подобное поведение создает ошибочные иллюзии у остальных трейдеров относительно ситуации на рынке и провоцирует их на принятие неверных решений. Все это приводит к изменению равновесия на рынке, т. е. цена на активы сдвигается в направлении, заданном агрессивными игроками, и как только она достигает благоприятного для них значения, они не замедляют воспользоваться сложившейся ситуацией. Довольно трудно обнаружить подобное поведение трейдеров, потому что сразу несколько брокеров вовлечены в агрессивную игру, и инициированные ими ордера распределены во времени.

На рынке ценных бумаг транзакции можно рассматривать как совокупность нескольких потоков данных — ордера на покупку/продажу и исполненные сделки. В соответствии с моделью рынка, эти временные последовательности проявляют высокую степень корреляции. Зачастую данные только одного типа используются в моделях детектирования агрессивных стратегий трейдинга [1–4], и при этом, как правило, теряется значительная часть информации, которая могла бы свидетельствовать о наличии сговора среди участников рынка. Только одновременное рассмотрение всех трех наборов данных позволяет узнать истинные намерения трейдеров [5].

Традиционный подход [1–4] к анализу временных рядов объединяет коррелированные последовательности в одну многомерную последовательность, что является оправданным, только если индивидуальные последовательности эволюционируют синхронно. В нашем случае мы имеем дело с тремя *асинхронно* взаимодействующими процессами, вследствие этого моделирование подобных процессов с помощью согласованных моделей значительно превосходит по точности традиционный подход. Кроме этого, (адаптивные) согласованные модели устойчивы к резким изменениям внутренней структуры потоков данных. Например, трейдер, изменяя свою стратегию, неявно оказывает влияние на структуру потока ордеров.

В данной статье рассматриваются два подхода к моделированию поведения инвесторов на фондовом рынке: на основе согласованных скрытых марковских моделей (ССММ) [6], и на основе адаптивных согласованных скрытых марковских моделей (АССММ). Второй из них, в отличие от первого, позволяет динамически изменять параметры модели вследствие изменения поведения игроков на рынке.

В основе этих моделей лежит классическая скрытая марковская модель (СММ) [7–9]. Стандартная СММ оперирует только одной переменной, которая используется для описания скрытых состояний. Вследствие этого стандартная СММ неточно описывает системы, включающие в себя несколько коррелированных процессов, таких, как упомянутые ранее три потока транзакций разного типа на фондовой бирже. В отличие от скрытой марковской модели, ССММ представляет собой более гибкий инструмент для моделирования временных последовательностей. Она состоит более чем из одной модели СММ-типа; каждая из входящих в ее состав СММ используется для описания отдельного процесса. Согласованная СММ добавляет дополнительные связи к составляющим ее СММ, так что скрытое состояние любой СММ, входящей в состав ССММ, в момент времени t зависит от скрытых состояний остальных СММ в момент времени $t - 1$.

Рассматриваемые модели апробированы на реальных биржевых данных. Результаты демонстрируют, что по качеству моделирования ССММ и АССММ превосходят стандартные СММ [7–10], не учитывающие корреляцию данных.

1. МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ ЗАДАЧИ ИССЛЕДОВАНИЯ

Действия брокеров можно представить в виде трех временных рядов: ордера на покупку/продажу и исполненные сделки. Таким образом, задача моделирования поведения трейдеров заключается в учете скрытых взаимосвязей между этими последовательностями.

Предположим, имеется S последовательностей $\Phi_1 = \{\phi_{11}, \phi_{12}, \dots, \phi_{1T}\}$, $\Phi_2 = \{\phi_{21}, \phi_{22}, \dots, \phi_{2F}\}$ и $\Phi_C = \{\phi_{C1}, \phi_{C2}, \dots, \phi_{CG}\}$, T , F и G — число элементов в каждой из них. В общем случае с каждым элементом последовательности связан L -мерный вектор атрибутов, описывающий его внутреннее состояние, $\phi_{ik}(p_{ik,1}, p_{ik,2}, \dots, p_{ik,L})$, $p_{ik,1}, p_{ik,2}, \dots, p_{ik,L}$ — атрибуты объекта ϕ_{ik} . Например, если мы имеем дело с ордером на продажу, то его атрибуты — это цена, объем, время подачи заявки.

Пусть атрибут $p_{ik,T}$ с индексом T содержит временную метку элемента ϕ_{ik} . Если ϕ_{ik} — ордер на

продажу, то $p_{ik,T}$ — время подачи заявки. Мы можем определить отношение согласованности между последовательностями Φ_i и Φ_j следующим образом: $\exists k, l \text{ s. t. } k \leq |\Phi_i|, l \leq |\Phi_j|, (p_{ik,T} < p_{jl,T} \wedge \phi_{ik} \rightarrow \phi_{jl}) \vee (p_{jl,T} < p_{ik,T} \wedge \phi_{jl} \rightarrow \phi_{ik})$. Данное определение означает, что одна из последовательностей содержит элемент, который влияет на некоторый элемент другой последовательности, обладающий более поздней временной меткой по сравнению с первым элементом. Если, например, две последовательности содержат ордера на покупку/продажу, то они согласованы, потому что размещение ордера на продажу ценных бумаг очевидно влияет сразу на один и более ордеров на покупку, размещаемых в последующие моменты времени.

Основное предположение, на которое опираются рассматриваемые в данной статье алгоритмы, — это трактовка согласованного поведения брокеров с целью агрессивного манипулирования ценами как *аномального*. Для аккуратного обнаружения аномалий требуется модель, способная точно описывать *нормальное* поведение трейдеров. Таким образом, *общую постановку задачи* моделирования согласованного поведения брокеров с целью обнаружения *агрессивных стратегий трейдинга* можно сформулировать следующим образом. Требуется разработать гибкую модель, способную эффективно учитывать отношение согласованности между коррелированными временными рядами, динамически обновлять свои параметры при изменении поведения трейдеров на рынке и точно обнаруживать аномалии.

Представленная модель, (А)ССММ, основана на теории согласованных марковских моделей [6]. Для качественного сравнения разрабатываемых моделей мы используем исторические биржевые данные, которые содержат подозрительные транзакции, обнаруженные биржевыми экспертами.

2. СКРЫТЫЕ МАРКОВСКИЕ МОДЕЛИ

Для моделирования взаимного влияния временных рядов мы воспользуемся согласованной СММ (ССММ) [6], которая состоит из трех СММ, описывающих потоки исполненных заявок и заявок на покупку/продажу соответственно.

Основные элементы СММ — это множество скрытых состояний $S = \{S_1, S_2, \dots, S_N\}$ и множество значений наблюдаемых переменных $O = \{O_1, O_2, \dots, O_T\}$. Для обозначения случайных и скрытых переменных мы используем прописные буквы, а для их реализаций — строчные. Таким образом, конкретное состояние в момент времени t обозначим s_t , а значение наблюдаемой переменной в момент времени t обозначим o_t . Вероятности пе-

переходов между скрытыми состояниями определяются матрицей $X = \{x_{ij}\}$, где $x_{ij} = \Pr(S_j = s_{t+1} | S_i = s_t)$. Распределение вероятности наблюдаемых переменных задается матрицей $Y = \{y_j(k)\}$, $y_j(k) = \Pr(O_k = o_k | S_j = s_j)$. И, наконец, распределение априорных вероятностей обозначается как $\pi = \{\pi_i\}$, где $\pi_i = \Pr(S_i = s_i)$. Таким образом, СММ, обозначаемая для краткости как $\lambda^{\text{СММ}}$, может быть полностью охарактеризована с помощью матриц X , Y , π , т. е. $\lambda^{\text{СММ}} = (X, Y, \pi)$. Оценка параметров СММ — элементов матриц X , Y , π — может быть эффективно выполнена с помощью стандартного алгоритма Баума — Велша [8]:

$$\alpha_t(i) = \Pr(O_1 = o_1, O_2 = o_2, \dots, O_t = o_t, S_t = s_i | \lambda),$$

$$\beta_t(i) = \Pr(O_{t+1} = o_{t+1}, O_{t+2} = o_{t+2}, \dots, O_T = o_T | S_t = s_i, \lambda),$$

$$\bar{x}_{iy} = \sum_{t=1}^{T-1} \alpha_t(i) x_{ij}(o_{t+1}) \beta_{t+1}(j) / \sum_{t=1}^{T-1} \alpha_t(i) \beta_{t+1}(j),$$

$$\bar{y}_j(k) = \sum_{t=1, o_t = O_k}^T \alpha_t(j) \beta_t(j) / \sum_{t=1}^T \alpha_t(j) \beta_t(j),$$

$$\bar{\pi}_i = \alpha_1(i) \beta_1(i) / \sum_{j=1}^N \alpha_1(j) \beta_1(j).$$

Первые два уравнения определяют вероятность получения заданной последовательности $o_1, o_2, \dots, o_t$ для состояния s_i в момент времени t и вероятность конечной заданной последовательности $o_{t+1}, o_{t+2}, \dots, o_T$ при условии, что СММ эволюционирует из исходного состояния s_i в момент времени t . Остальные уравнения позволяют оценить параметры СММ $\lambda^{\text{СММ}} = (X, Y, \pi)$.

После того, как оценка параметров модели выполнена, она может быть использована для вычисления вероятности того, что наблюдаемая последовательность $o_1, o_2, \dots, o_T$ сгенерирована именно этой моделью:

$$\Pr(O | \lambda^{\text{СММ}}) = \sum_{i=1}^N \alpha_T(i).$$

В отличие от стандартной СММ, согласованная СММ объединяет несколько ССМ, добавляя связи между скрытыми состояниями отдельных ССМ моделей. На рис. 1 показана ССММ, состоящая из трех связанных ССМ, каждая из которых моделирует поток заявок на покупку/продажу и исполненных ордеров. В виде овалов изображены скрытые состояния марковских моделей, прямоугольники символизируют наблюдаемые величины

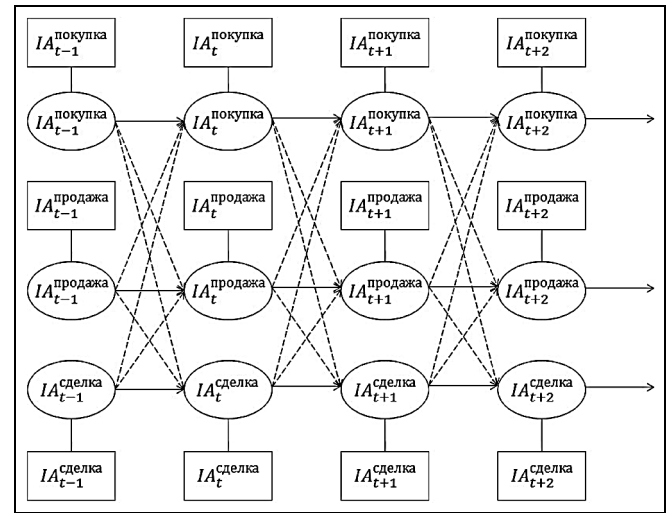


Рис. 1. Согласованная скрытая марковская модель

ны. Связи между скрытыми состояниями отдельных ССМ изображены штриховыми линиями.

С формальной точки зрения, дополнительные связи между парой стандартных СММ описываются матрицей, элементы которой характеризуют уровень взаимного влияния скрытых переменных друг на друга. Иными словами, если мы рассматриваем две согласованные СММ, λ и λ' , то матрица $Z = \{z_{ij}\}$ определяется как $z_{ij} = \Pr(s'_{t+1} = S'_j | s_t = S_i)$, где S'_j и S_i — скрытые состояния моделей λ' и λ в моменты времени $t+1$ и t соответственно. Ввиду сказанного, ССММ полностью описывается набором из четырех параметров $\lambda^{\text{СМММ}} = (X, Y, Z, \pi)$. По аналогии с алгоритмом Баума — Велша, для оценки элементов матрицы Z применяется итеративный алгоритм [6], основанный на методах динамического программирования.

Один из ключевых моментов в построении любой СММ состоит в выборе скрытых состояний. Поскольку результат исполнения ордера на покупку/продажу может быть либо *положительным/отрицательным*, либо *нулевым* в зависимости от того, насколько выгодна сделка для иницилирующей ее стороны, то потоки ордеров на покупку/продажу могут быть смоделированы с помощью СММ с данными тремя скрытыми состояниями:

$$S^{\text{продажа/покупка}} = \{\text{положительный, нулевой, отрицательный}\}.$$

Последовательность исполненных сделок представлена в виде СММ с двумя скрытыми состояниями, первое из которых соответствует восходящему тренду на рынке, а второе — нисходящему:

$$S^{\text{сделка}} = \{\text{восходящий, нисходящий}\}.$$

3. ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ВХОДНЫХ ДАННЫХ

Для построения эффективных моделей, способных в нужной степени обобщать экспериментальные данные, последние подвергаются предварительной обработке — усреднению и дискретизации. Все рассматриваемые СММ обладают дискретной природой, т. е. наблюдаемые переменные дискретны. Каждую транзакцию в потоке входных данных можно описать с помощью следующей контекстно-свободной грамматики:

$$A = \{a, p, v\},$$

$$a = \{\text{продажа}|\text{покупка}|\text{сделка}\},$$

$$p = \{\text{цена: численные данные}\},$$

$$v = \{\text{объем: численные данные}\},$$

где A (*activity*) — информация о конкретном ордере/сделке, переменные a , b и v определяют тип транзакции, цену и объем соответственно.

Входные данные необходимо подвергнуть интервальному усреднению внутри скользящего временного окна размером w , содержащем n элементов, чтобы уменьшить дисперсию и выявить тренд:

$$IA = \{AS, \bar{p}, \bar{v}, w\},$$

$$AS = \{A_1, A_2, \dots, A_n\},$$

$$\bar{p} = \frac{1}{n} \sum_{i=1}^n p(i),$$

$$\bar{v} = \frac{1}{n} \sum_{i=1}^n v(i),$$

где оценка IA (*interval activity*) описывает усредненные параметры транзакций на интервале фиксированного размера w , AS (*sequence of activities*) — совокупность всех транзакций, инициированных внутри временного интервала w . Параметры $\bar{p}$ и $\bar{v}$ представляют собой средние значения цен и объемов на интервале w ; n — число транзакций на интервале w .

Поскольку наши модели работают с дискретными значениями величин, мы дискретизируем интервальные оценки IA наблюдаемых величин с помощью k -means алгоритма [11], заменяя непрерывные величины центрами полученных кластеров. Например, если $\alpha_i^{\text{цена}}$ — центр i -го кластера, построенного для параметра «цена», то все величины p , принадлежащие этому кластеру, можно заменить на $\alpha_i^{\text{цена}}$. Математически, переход от усредненных непрерывных параметров транзакций

IA к их дискретным значениям $IA^{\text{дискр}}$ описывается формулами:

$$p^{\text{дискр}} = \operatorname{argmin}_i |\bar{p} - \alpha_i^{\text{цена}}|,$$

$$v^{\text{дискр}} = \operatorname{argmin}_i |\bar{v} - \alpha_i^{\text{объем}}|,$$

$$w^{\text{дискр}} = \operatorname{argmin}_i |\bar{w} - \alpha_i^{\text{интервал}}|,$$

$$IA = \{AS, \bar{p}, \bar{v}, w\} \rightarrow IA^{\text{дискр}} = \{AS, p^{\text{дискр}}, v^{\text{дискр}}, w^{\text{дискр}}\}.$$

4. АДАПТИВНАЯ СОГЛАСОВАННАЯ МАРКОВСКАЯ МОДЕЛЬ

Поведение трейдеров на рынке ценных бумаг может меняться достаточно резко по различным причинам, например, вследствие изменения внешних условий или поступления дополнительной информации. Поэтому ССММ должна своевременно обнаруживать изменившееся поведение участников рынка и мгновенно подстраиваться к новым условиям. Вариант ССММ, обладающий подобными характеристиками, именуется адаптивной ССММ или АССММ.

Для того чтобы определить, нужно ли обновлять значения параметров АССММ, специальный компонент в нашей системе в режиме реального времени оценивает, насколько качественно ССММ обобщает поступающие данные. Подобный анализ выполняется с помощью статистического t -теста, который способен отвергнуть гипотезу о том, что тестовая последовательность и текущая последовательность сгенерированы одной и той же моделью. В качестве тестовой последовательности рассматривается последовательность данных, проанализированная посредством АСММ сразу после последнего обновления параметров. Обнаруженная существенная разница свидетельствует о необходимости модификации параметров АССММ, потому что АССММ в данном виде не способна качественно описывать изменившееся поведение участников рынка.

Данные DS_1 из интервала $[t_1, t_2]$ (рис. 2) используются как тестовые, они описываются моделью типа ССММ (обозначим ее АССММ-1). Вероятность того, что исходные данные сгенерированы моделью АССММ-1, используется в качестве опорной величины для оценки способности модели АССММ-1 описывать наблюдаемые переменные на следующем интервале, на интервале $[t_3, t_4]$. Если t -тест свидетельствует о малой вероятности, что данные из текущего интервала описываются текущей АССММ, то параметры модели обновляются.

В целях уменьшения времени и вычислительных ресурсов, необходимых для обновления пара-

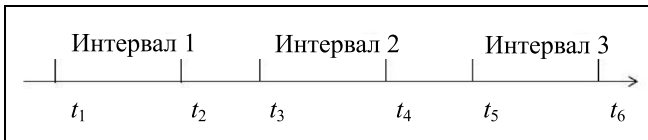


Рис. 2. Пример интервального обновления параметров модели типа АССММ

метров, применяется метод скользящего окна. Допустим, что модель АССММ-1 следует обновить, чтобы она смогла лучше описывать данные из интервала $[t_3, t_4]$. При вычислении новых значений параметров текущие значения параметров модели АССММ-1 используются в качестве исходных оценок, а сама модель обучается заново на данных из интервала $[t_3, t_4]$, таким образом мы получаем модель АССММ-2'. Окончательные оценки параметров модели АССММ-2 вычисляются как линейная комбинация параметров моделей АССММ-1 и АССММ-2'. Мы переобучаем модель только на последнем интервале, на котором она плохо работает, а не на всем наборе данных в отличие от классического подхода. Процесс обновления параметров представляется формулами

$$x_{ij} = (1 - w)x'_{ij} + wx''_{ij},$$

$$y_{ij} = (1 - w)y'_{ij} + wy''_{ij},$$

$$z_{ij} = (1 - w)z'_{ij} + wz''_{ij},$$

$$\pi_{ij} = (1 - w)\pi'_{ij} + w\pi''_{ij}.$$

Переменные, отмеченные одним штрихом, соответствуют старым значениям параметров (в нашем примере, параметрам модели АССММ-1), двумя штрихами — новым интервальным оценкам параметров (параметрам модели АССММ-2'), без штрихов — значениям параметров в обновленной модели (параметрам модели АССММ-2); w — весовой коэффициент.

5. АЛГОРИТМ ПОСТРОЕНИЯ МОДЕЛЕЙ

Алгоритм состоит из двух логических частей: первоначальная обработка данных и применение АССММ для обнаружения сговора среди участников рынка.

На фондовом рынке ордера, размещаемые трейдерами, автоматически аннулируются в конце торгового дня, если они не исполнены. Сделки совершаются только с ордерами, полученными в течение текущего рабочего дня. Поэтому причинно-следственная связь между ордерами и сделками существует только в рамках одного рабочего дня. В связи с чем мы разбиваем каждый торговый день на несколько интервалов, и для каждого интервала

вычисляем интервальные активности в соответствии со следующим алгоритмом.

Шаг 1. Разбить торговый день на L интервалов равной длительности.

Шаг 2. Для каждого временного интервала вычислить активности $IA^{\text{продажа}}$, $IA^{\text{покупка}}$, $IA^{\text{сделка}}$.

Шаг 3. Дискретизировать интервальные активности: $IA^{\text{продажа}}$, $IA^{\text{покупка}}$, $IA^{\text{сделка}} \rightarrow IA^{\text{дискр-продажа}}$, $IA^{\text{дискр-покупка}}$, $IA^{\text{дискр-сделка}}$.

Шаг 4. Построить последовательность интервальных активностей, описывающих транзакции в течение одного торгового дня, т. е. $IA^{\text{тип}} = IA_1^{\text{тип}}$, $IA_2^{\text{тип}}$, ..., $IA_L^{\text{тип}}$.

Шаг 5. Повторить шаги 1—4 для всех дней.

После первоначальной обработки данные используются для построения АССММ. Стратегии, вызывающие подозрение, детектируются с помощью оценки вероятности того, что проверяемые последовательности ордеров сгенерированы имеющейся в нашем распоряжении моделью. Если эта вероятность ниже порогового значения, то это служит признаком того, что данные могут быть отправлены на экспертизу для проведения более тщательного анализа. Приведенный ниже алгоритм детектирует подозрительные действия трейдеров.

Шаг 1. Разбить выходные данные приведенного выше алгоритма на два множества: данные, необходимые для обучения АССММ, и данные, необходимые для проверки качества построенной АССММ.

Шаг 2. Обучить АССММ на первом множестве.

Шаг 3. Вычислить статистические параметры полученной АССММ: ее среднее значение

$$\mu = \frac{1}{K} \sum_{i=1}^K \Pr(Seq_i | \text{АССММ})$$

и дисперсию

$$\sigma = \left(\frac{1}{K} \sum_{i=1}^K \Pr(Seq_i | \text{АССММ} - \mu) \right)^{1/2},$$

где Seq_i — i -й набор входных данных, K — число наборов входных данных, используемых для обучения модели.

Шаг 4. Для каждой тестовой последовательности вычислить, насколько вероятно то, что она сгенерирована построенной АССММ, по формуле

$$D = \frac{1}{\sigma} (\mu - \Pr(Seq' | \text{АССММ})),$$

где Seq' — тестовая последовательность.

Шаг 5. Если $D > \psi_0$, где ψ_0 — пороговое значение, то совокупность Seq' содержит подозрительную последовательность ордеров.

6. ЭКСПЕРИМЕНТАЛЬНЫЕ ДАННЫЕ

Описанный подход к моделированию биржевых данных был апробирован на реальных данных, собранных более чем за год. Они включают в себя информацию об обнаружении агрессивных стратегий, собранную системой биржевого мониторинга. Хотя эти данные не являются абсолютно точными, они используются в качестве референсных в наших экспериментах. Набор данных для обучения АССММ построен так, что он не включает в себя подозрительных операций, т. е. АССММ обучается только на нормальных данных. Одна треть данных использовалась для обучения, а остальные — для тестирования моделей типа (А)СММ.

Для сравнения различных вариантов скрытых марковских моделей мы воспользовались шестью моделями:

- СММ-ПР/СММ-ПО/СММ-СД — данные СММ моделируют последовательности заявок на продажу/покупку ценных бумаг и последовательность исполненных сделок соответственно; в этих моделях не используются понятия согласованности и адаптивности;
- СММ-ИН — интегрированная СММ; вероятность входной последовательности определяется как среднее значение вероятностей отдельных СММ-ПР/СММ-ПО/СММ-СД-моделей;
- ССММ;
- АССММ.

Для сравнения описанных моделей в качестве критериев мы воспользовались стандартными па-

раметрами [11], принятыми в теории машинного обучения:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN};$$

$$Precision = \frac{TP}{TP + FP};$$

$$Recall = \frac{TP}{TP + FN};$$

$$Specificity = \frac{TN}{FP + TN},$$

где TP означает true positive, FP — false positive (ошибка первого рода), FN — false negative (ошибка второго рода), TN — true negative.

Кроме приведенных критериев модели сравнивались по показателю прибыльности ($Return$), широко применяемому на фондовых рынках:

$$R = \ln \frac{p_t}{p_{t-1}},$$

где p_t и p_{t-1} — уровни цен, на которых были совершены сделки в моменты времени t и $t-1$ соответственно. Дни, в течение которых были подозрительные сделки, имеют более высокое значение показателя прибыльности по сравнению с обычными днями. В соответствии с этим модель, которая лучше всего описывает поведение рынка, должна в первую очередь выявлять дни с наибольшим значением прибыльности.

Результаты тестирования шести моделей на реальных данных представлены на рис. 3–7. По горизонтальной оси отложено число P временных окон, в течение которых зарегистрированы подозрительные сделки, а по вертикальной — соответствующий показатель.

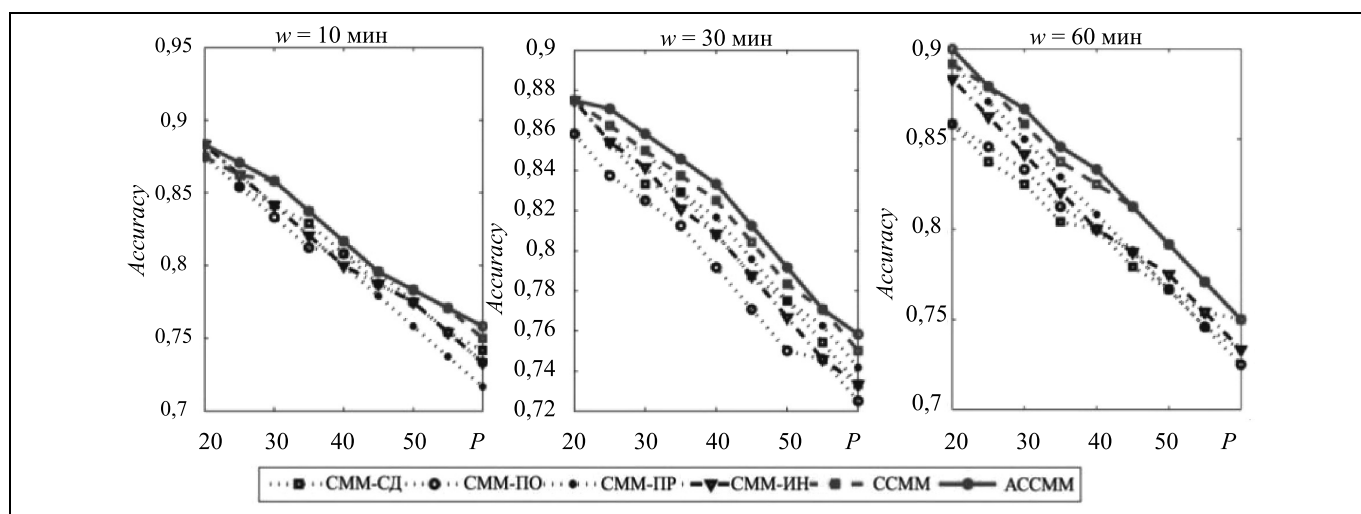
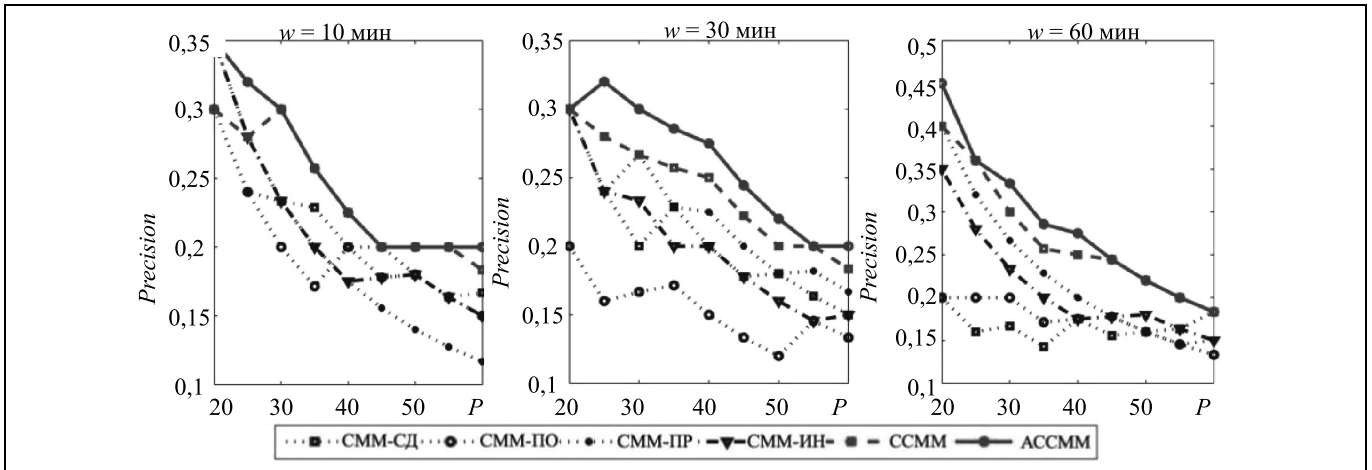
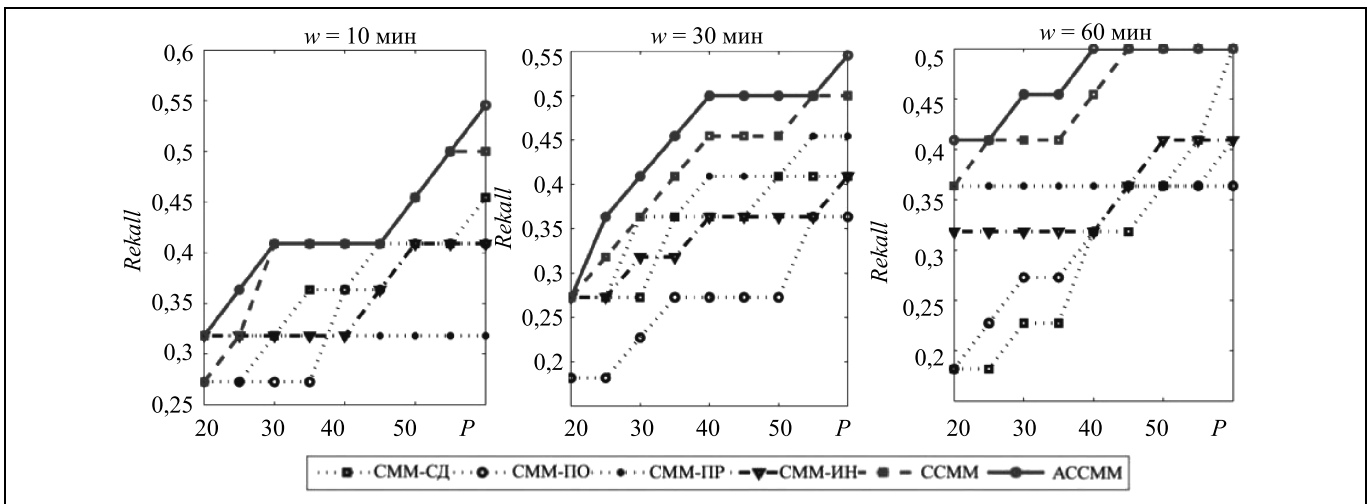
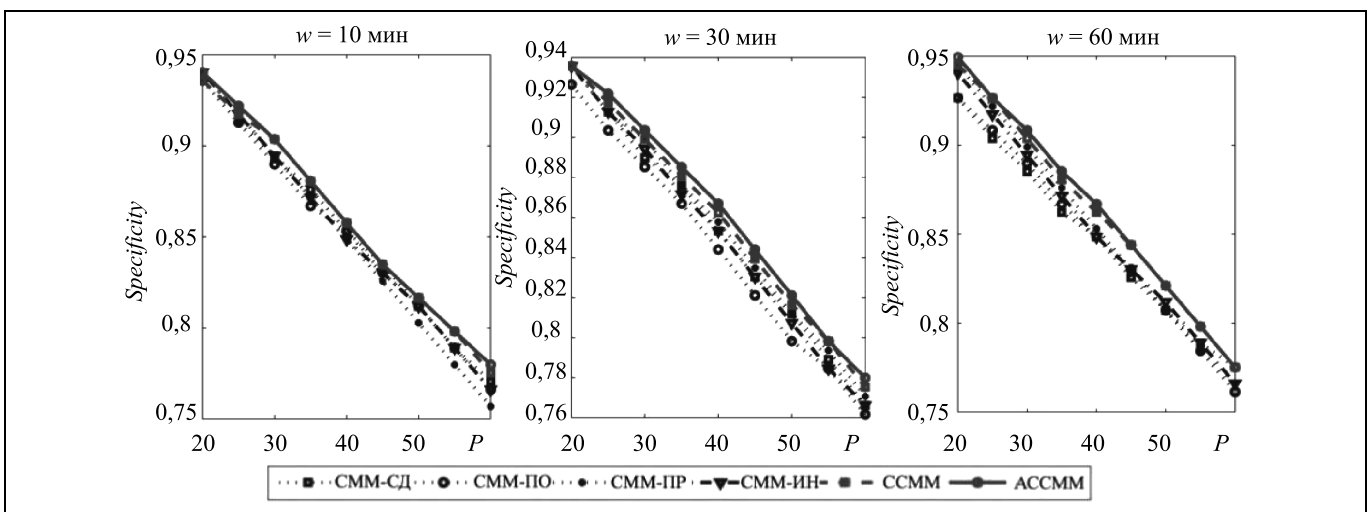
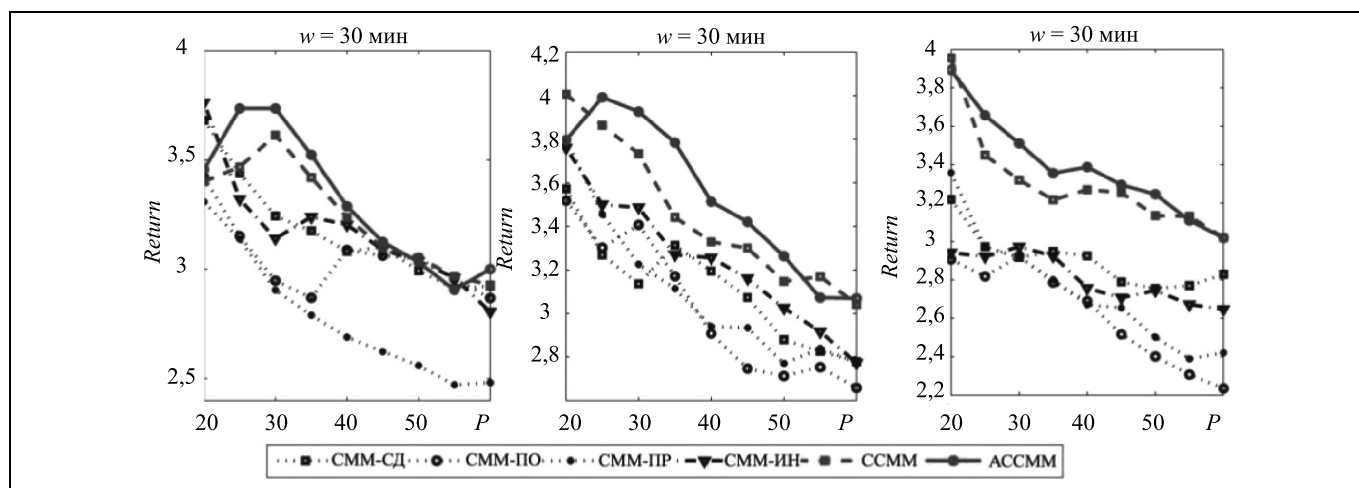


Рис. 3. Динамика параметра $Accuracy$

Рис. 4. Динамика параметра *Precision*Рис. 5. Динамика параметра *Recall*Рис. 6. Динамика параметра *Specificity*


Рис. 7. Динамика параметра *Return*

Эксперименты выполнены при значениях длительности интервала: 10, 30 и 60 мин.

Наиболее точно входные данные моделирует АССММ. Однако отметим, что, хотя ССММ уступает АССММ, она все же превосходит по качеству моделирования стандартные СММ. Из этого следует, что учет взаимосвязей между отдельными СММ необходим для более точного описания поведения трейдеров.

ЗАКЛЮЧЕНИЕ

Глобальный финансовый кризис 2008 года послужил стимулом к разработке автоматических методов обнаружения агрессивных стратегий на фондовом рынке, которые реализуются группой брокеров, преднамеренно вступивших в сговор. Подобные случаи привлекают внимание надзорных органов и могут послужить поводом к усилению контроля на рынке ценных бумаг [12, 13].

В данной статье продемонстрировано, что согласованное поведение брокеров с целью агрессивного манипулирования ценами аномально и точность его обнаружения зависит от точности моделирования нормального поведения трейдеров. По сравнению с существующими системами достигнуты лучшие показатели благодаря учету взаимосвязей между поведением отдельных трейдеров с помощью (адаптивных) согласованных скрытых марковских моделей. Рассмотренные методы применимы не только к фондовому рынку, но и к любым другим распределенным системам, отдельные части которых способны вступать в сложные взаимодействия друг с другом.

ЛИТЕРАТУРА

1. Zaki M. SPADE: An Efficient Algorithm for Mining Frequent Sequences // Journal Machine Learning. — 2001. — Vol. 42, iss. 1–2. — P. 31–60.
2. Srikant R., Agrawal R. Mining sequential patterns: Generalizations and performance improvements // The 5th International Conference on Extending Database Technology Proceedings. — London, 1996. — P. 3–17.
3. PrefixSpan: Mining sequential patterns efficiently by prefix-projected pattern growth / J. Pei, et al. // The 17th International Conference on Data Engineering Proceedings. — Heidelberg, Germany, 2001. — P. 215–226.
4. Sequential pattern mining using a bitmap representation / J. Ayres, et al. // The Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining Proceedings. — Edmonton, Canada, 2002. — P. 429–435.
5. Gwadera Z., Crestani F. Discovering significant patterns in multi-attribute sequences // The Eight IEEE International Conference on Data Mining Proceedings. — Pisa, 2008. — P. 827–832.
6. Brand M. Coupled hidden markov models for modeling interacting processes // Techn. Report. — URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.50.8099&rep=rep1&type=pdf> (дата обращения: 31.07.2014).
7. Gales M., Young S. The Application of Hidden Markov Models in Speech Recognition // Foundations and Trends in Signal Processing. — 2007. — Vol. 1, N 3. — P. 195–304.
8. Rabiner L. A tutorial on hidden Markov models and selected applications in speech recognition // Proceedings of the IEEE. — Piscataway, 1989. — Vol. 77. — P. 257–286.
9. Dymarski P. Hidden Markov models, theory and applications. — URL: <http://www.intechopen.com/books/hidden-markov-models-theory-and-applications> (дата обращения: 20.04.2016).
10. Russell S., Norvig P. Artificial Intelligence: A Modern Approach. — London: Pearson Education Ltd., 2014. — 1132 p.
11. Bishop C. Pattern recognition and machine learning. — N.-Y.: Springer, 2009. — 738 p.
12. Каздагли М.Н. Выбор подходящих для высокочастотного трейдинга (ВЧТ) активов. — URL: http://www.mlsl.info/confmlsd/MLSD_2012/Section2.php (дата обращения: 20.04.2016).
13. Каздагли М.Н. Перспективы использования методов оперативного принятия решений на торговых площадках. — URL: http://www.mlsl.info/confmlsd/MLSD_2011/Section3.php (дата обращения: 20.04.2016).

Статья представлена к публикации членом редколлегии А.С. Манделем.

Каздагли Михаил Нижмеддинович — науч. сотрудник, Техасский университет, г. Остин, mkazdagli@gmail.com.