



ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 69 за 2010 г.



**Кудрявцев М.Ю.**

Алгоритм прогноза сильных падений индекса Dow Jones Industrial Average

**Рекомендуемая форма библиографической ссылки:** Кудрявцев М.Ю. Алгоритм прогноза сильных падений индекса Dow Jones Industrial Average // Препринты ИПМ им. М.В.Келдыша. 2010. № 69. 22 с. URL: <http://library.keldysh.ru/preprint.asp?id=2010-69>

## ВВЕДЕНИЕ

Задача прогнозирования в финансовом мире вообще и на рынках акций в частности всегда была одной из самых острых. Существует целая индустрия профессионального управления финансовыми активами, например, инвестиционные банки. Но, несмотря на очевидную важность и практическую значимость этой задачи, практически каждый раз очередной финансовый кризис оказывается неожиданным, и банкротами становятся, в том числе, старейшие и наиболее уважаемые компании, банки и пр. (один из последних случаев – Lehman Brothers).

Вплоть до настоящего времени классические и широко используемые методы прогнозирования на финансовых рынках, фундаментального и технического анализа [1-2], не ориентированы на предсказание «сильных событий», и даже предсказание смены тренда является для них практически неразрешимой задачей. Существует множество индикаторов и моделей смены тренда (модель «голова–плечи» и ее варианты, например), и почти всегда часть из них противоречит другим. Большинство же игроков фондовой биржи эффективно зарабатывают на следовании рыночным тенденциям.

С точки зрения макроэкономики, предсказание финансовых кризисов тесно связано с существованием так называемых «пузырей». Интуитивно все ясно. Пузырь возникает из-за завышенных и необоснованных ожиданий, а затем по тем или иным причинам лопается, вызывая финансовый кризис. К сожалению, заранее предугадывать существование «пузырей» с достаточной степенью достоверности не получается ни у финансовых аналитиков, ни у экономистов.

В своей статье [3] Гуркайнак резюмирует: «Эконометрические методы определения ценовых пузырей на рынках активов не обладают удовлетворительной точностью. Для каждой статьи, которая обнаруживает свидетельства присутствия пузыря, находится другая, которая объясняет существующие данные ровно так же хорошо без предположений о несоответствии цены и реальной стоимости активов. Мы будем неспособны отличить пузыри от меняющихся со временем или вследствие переключения режима фундаментальных свойств рынков до тех пор, пока большинство модельных задач в области нахождения пузырей остаются нерешенными».

Если рассматривать задачу прогноза с точки зрения естественных наук, то здесь существует два подхода.

Первый основан на возможности предложить для описываемого процесса или явления достаточно подробную модель. Но на данной стадии развития науки фондовый рынок не может быть «жестко» смоделирован по следующим причинам:

- Большое число участников, которые сильно разнородны по своим свойствам и целям. Игроками фондовых рынков являются мелкие инвесторы, государство, пенсионные фонды, профессиональные игроки, инвестиционные бан-

ки, частные компании. Каждый имеет свои временные горизонты и решает свои собственные задачи.

- Открытость системы. Динамику фондовых рынков невозможно рассматривать в отрыве от экономики и информационного пространства, катастроф и стихийных бедствий.

- Невозможность на настоящий момент смоделировать человека. Исследования показывают, что даже на фондовом рынке, где, казалось бы, основной целью любого игрока есть получение прибыли, люди ведут себя нерационально (например, поддаются панике).

Второй подход основан на следующем наблюдении: сложные нелинейные системы, такие как экономика, общество, человек, демонстрируют большую часть времени достаточно регулярное поведение. И часто перед сильным событием возникают предкризисные сигналы, на основании которых можно построить прогноз. По сути, именно этим и занимаются врачи, устанавливая диагноз по косвенным признакам.

Одним из наиболее исследованных примеров большой нелинейной иерархической системы является литосфера Земли, представляющая собой систему блоков, разделенных разломами. Каждый из этих блоков делится на более мелкие – от тектонических плит протяженностью в тысячи километров до зерен горных пород миллиметрового размера и меньше. Тепловая конвекция Земли заставляет плиты двигаться и взаимодействовать между собой. Возникающие на разных масштабных уровнях напряжения могут приводить к разрушению соответствующих по размеру блоков. Разрушения мелких неоднородностей горной породы и микроблоков отвечают сейсмическому шуму, разрушение достаточно больших блоков фиксируются как землетрясения. Накапливающиеся в результате взаимодействия блоков напряжения имеют тенденцию распространяться от нижних уровней иерархии к верхним – обратный каскад, приводя к таким событиям, как сильные землетрясения. Разрядка напряжения землетрясениями может рассматриваться как перераспределение напряжений сверху вниз по иерархии – прямой каскад. Одно из характерных статистических свойств землетрясений – степенной закон распределения вероятностей, который является математическим выражением сложности системы и ее склонности к катастрофическим событиям.

В работах [4-7] подробно изучены статистические свойства распределения изменений цен акций на основных рынках ценных бумаг США (NYSE, AMEX, NASDAQ) и индекса S&P500. Установлено, что для редких событий (больших абсолютных значений доходности) имеет место степенное распределение. Поэтому интересен перенос методов прогноза сильных землетрясений на поведение рынка ценных бумаг [8-9].

Развитие сильных событий, таких как сильные землетрясения, в иерархических системах подготавливается предварительной активизацией на предшествующих иерархических уровнях. Соответствующее изменение формы распределения успешно используется не только при прогнозе сильных землетрясений [10], но и при прогнозе скачков тяжелой преступности [11].

В данной работе введено формальное определение «сильного» падения финансового индекса DJIA. При помощи методологии распознавания образов редких событий предложен алгоритм прогноза таких падений, основанный на изменении формы выборочной функции распределения детрендрованного ряда цен закрытий при приближении к объекту. Алгоритм способен предсказывать все падения индекса, которые являются предметом прогноза, при этом суммарное время тревоги (время, в течение которого ожидается наступление события), не превышает одной четвертой части рассматриваемого временного интервала, а доля ложных тревог не превышает одной трети. Для оценки практической значимости результатов предложена и оценена простейшая торговая стратегия. На данный момент алгоритм проходит верификацию в режиме работы в реальном времени, успешно предсказав три объекта за последний год.

## **1. ОБЩАЯ ПОСТАНОВКА ЗАДАЧИ**

### ***1.1 Описание данных и процессов***

Финансовые индексы являются основными показателями работы фондовых рынков. Индекс Dow Jones Industrial Average является самым старым и наиболее известным фондовым индексом. DJIA рассчитывается как среднее арифметическое (с поправкой на такие операции, как дробление акций, выплат дивидендов в форме акций и т.п.) курсов акций 30 крупнейших в своих отраслях промышленных корпораций. Список компаний пересматривается с развитием ситуации на фондовом рынке. Составлением этого списка занимается редакция газеты The Wall Street Journal.

Акции, используемые в данном индексе, котируются на Нью-Йоркской фондовой бирже (NYSE) и составляют порядка 15-20 процентов общей рыночной стоимости акций NYSE.

Широко известно, что финансовые индексы очень чутко реагируют на внешнюю информацию. Неожиданная экономическая или политическая (а иногда даже метеорологическая) новость может стать причиной очень сильного скачка на рынках. Например, 29 сентября 2009 года Палата представителей Конгресса отвергла проект закона о срочных мерах по преодолению финансового кризиса. Известие об этом привело к падению DJIA на 6,98% за один день. Понятно, что пытаться предсказывать такого рода события бессмысленно. С другой стороны, если рассматривать более длительные интервалы времени, то влияние отдельных новостей в значительной степени нивелируются. Для индекса DJIA, который в целом характеризует оценку фондовым рынком всей экономики США, в значительной мере можно считать сильные события порождением его собственной внутренней динамики.

Для примера кризисов, которые хотелось бы иметь возможность предсказывать, приводим выдержку из статьи «Уолл-стрит. Краткая история паник» международного информационно-аналитического агентства Washington Profile:

«1987 год. 19 октября 1987 года (в «черный понедельник») финансовый кризис накрыл весь мир: акции на мировых финансовых площадках обвалились, в среднем, на 22.6%, совокупные потери превысили \$500 млрд. В «черный понедельник» Dow Jones потерял 508 пунктов или 22.61%. Причины этого кризиса до сих пор дебатированы: неизвестно, почему многие инвесторы решили продавать свои ценные бумаги именно в этот день...

1998 год. На биржи США повлияли экономические кризисы в Азии и России, а также политический фактор: успешные испытания северокорейской баллистической ракеты. В конце августа Dow Jones упал почти на 7%...

2001 год. Этот крах — точнее, серия крахов — завершили восьмилетний период бурного экономического роста. В 2000 году в США началось обрушение «доткомов» — интернет-компаний, прибыльность которых оказалась намного ниже ожиданий инвесторов. В 2000 году примерно 1.8 тыс. интернет-компаний сменили владельцев или обанкротились, в 2001 году — их было около 1.7 тысяч. В 2000 году Нью-Йоркская Фондовая Биржа пережила несколько значительных падений индексов — к примеру, 14 апреля индекс Dow Jones упал более чем на 617 пунктов: эту неделю назвали худшей в истории биржевой неделей...

11 сентября 2001 года произошли атаки террористов на Нью-Йорк и Вашингтон. После этого Нью-Йоркская Фондовая Биржа не работала четыре дня — самые длительные каникулы с 1933 года. А 17 сентября произошел очеред-

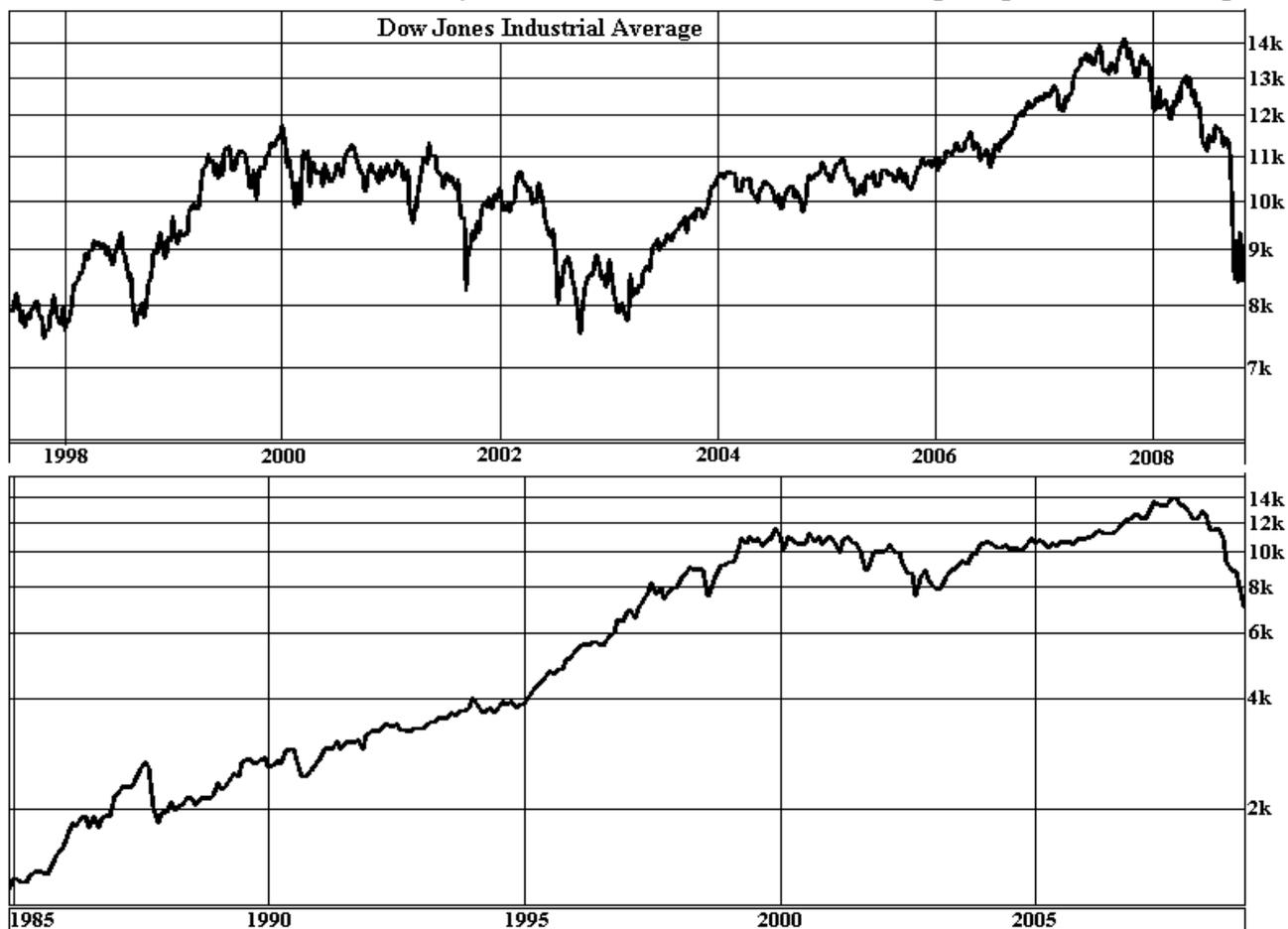


Рис.1. Цены закрытия индекса DJIA периода 01.07.1997 – 10.10.2008 (на верхнем графике) и 31.12.1984 – 06.02.2009 (на нижнем).

ной крупнейший биржевой крах — курсы акций упали почти на 685 пунктов. Ситуацию на рынке усугубила череда громких банкротств крупных компаний».

Предметом данного исследования являются дневные цены закрытия фондового индекса Dow Jones Industrial Average. Основной период рассмотрения — с 1-го июля 1997 года по 10-е октября 2008 — содержит 2838 отсчетов (торговых дней). При дополнительном анализе также исследуется временной промежуток с 31-го декабря 1984 года по 6-е февраля 2009, содержащий 6080 отсчетов (см. рис. 1).

## 1.2 Описание общей методологии предсказания

Задача прогноза формулируется следующим образом. Дан временной ряд, описывающий динамику системы вплоть до настоящего момента времени  $t$  и содержащий потенциальные предвестники сильного события. Требуется сказать: возникнет или не возникнет сильное событие в течение последующего периода времени  $(t, t+\Delta t)$ . Если ответ «да» (то есть мы ожидаем, что произойдет сильное событие), этот интервал назовем «время тревоги».

По мере прохождения по временной оси алгоритм прогноза формирует дискретную последовательность тревог. Возможные исходы такого предсказания изображены на рис. 2. Такой дискретный (в смысле ответов «да», «нет») алгоритм предназначен не для анализа динамики системы в целом, а для определения редких сильных событий.

Алгоритм прогноза содержит в себе много регулируемых составляющих, начиная от выбора данных и определения целей предсказания (формализация понятия объекта), и заканчивая численным выбором параметров, входящих в алгоритм. В силу невозможности существования теории, которая могла бы однозначно определить эти элементы, они должны быть найдены посредством ретроспективного прогноза. Такая ситуация создает потенциальную опасность



Рис. 2. Возможные исходы предсказания.

для исследователя увлечься подгонкой данных и, в конце концов, обмануть самого себя. Поэтому для проверки качества алгоритма прогноза необходимы следующие тесты:

- анализ чувствительности: являются ли результаты работы алгоритма устойчивыми по отношению к изменениям настраиваемых параметров;
- анализ данных, не входящих в материал обучения;
- предсказание будущих событий в режиме «реального времени».

Задача анализа эффективности алгоритма также включает оценку точности прогноза:

- процент пропуска целей,  $n$ ;
- отношение общей продолжительности времени тревоги к длительности всего рассматриваемого интервала,  $\tau$ ;
- относительное число (длительность) ложных тревог,  $\eta$ .

Этими характеристиками измеряется «вероятностное» качество алгоритма прогноза. Они обеспечивают статистическую оценку эффективности алгоритма и позволяют формализовать задачу управления при принятии решений относительно подготовки к возможным сильным событиям.

## 2. АЛГОРИТМ ПРЕДСКАЗАНИЯ

### 2.1 Определение объекта прогноза

Для того чтобы ограничить влияние внешнего информационного пространства, необходимо рассматривать изменения значений индекса на длительном интервале времени. С другой стороны, хочется иметь возможность прогнозировать события, которые являются неожиданными для самого фондового

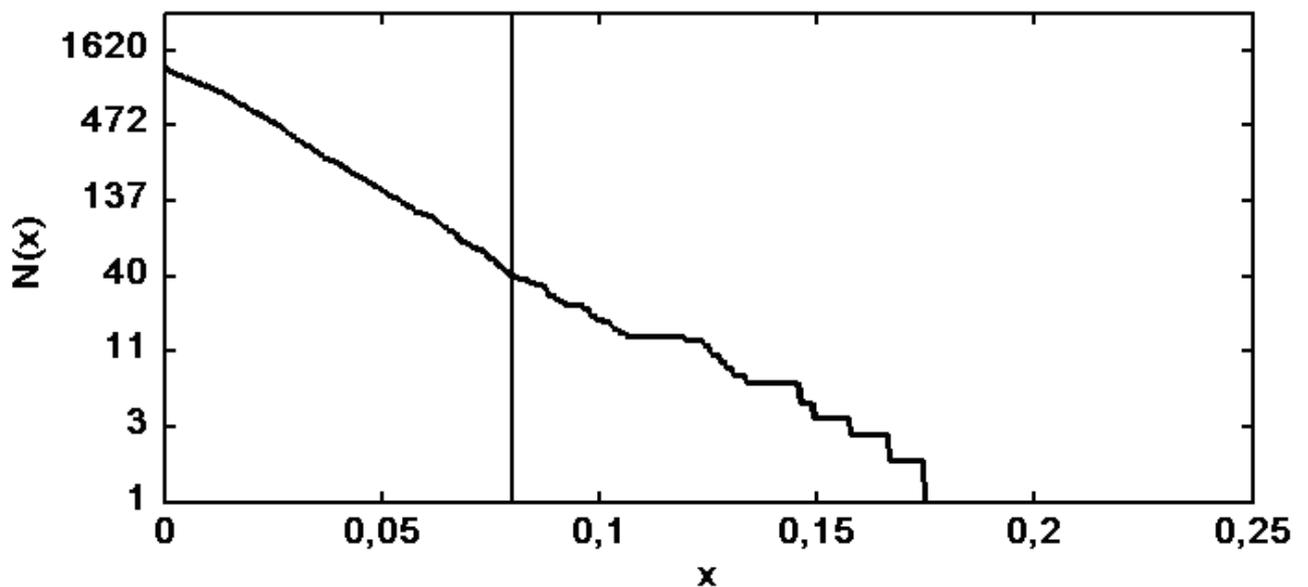


Рис. 3. Кумулятивное распределение  $N(x)$  модуля отрицательных значений  $v_i$ , по оси ординат выбран логарифмический масштаб, а вертикальная прямая ограничивает слева те дни, которые мы считаем сильными событиями.

рынка. Поэтому длительное медленное снижение индекса также не представляется интересным. Исходя из этих соображений, был сделан выбор измерять «силу» события (конкретной цены закрытия) относительно скользящего среднего, взятого за предшествующие 6 недель (30 рабочих дней). Таким образом, мы игнорируем и резкие однодневные колебания, и значительные, но очень медленные (а значит и сами по себе понятные участникам фондовой биржи) снижения значений индекса.

$$\text{Рассмотрим величину } v_i = \frac{p_i - \langle p_{i-1} \rangle_m}{\langle p_{i-1} \rangle_m}, \quad (1)$$

где  $p_i$  – значение цены закрытия, а  $\langle p_{i-1} \rangle_m$  – среднее величин  $p_{i-m+1}, \dots, p_i$ , а  $m$  выбирается таким образом, чтобы охватить временной период порядка полутора месяцев. Таким образом,  $v_i$  представляет собой некоторый аналог доходности.

**Сильным событием** назовем день, когда соответствующее ему значение  $v_i < -0.08$ . Иными словами, в этот день цена закрытия индекса будет меньше среднего значения цен закрытия последних  $m = 30$  биржевых дней на 8%.

Рассмотрим интервал времени с 1-го июля 1997 года по 10-е октября 2008 года. Пусть  $N(x)$  – число дней, когда  $v_i < -x$  ( $x > 0$ ). Тогда  $N(x)$  – кумулятивное распределение абсолютных значений отрицательных  $v_i$  (см. рис. 3). Поскольку величина  $v_i$  аналогична логарифму цены, то график фактически изображен в двойных логарифмических координатах, и его форма означает, что для значения индекса имеет место степенное распределение.

Сильные события часто идут подряд одно за другим. Такую группировку сильных событий (расстояние между двумя соседними не должно превышать 8 торговых дней) назовем **объектом**. За последние 11 лет с 1-го июля 1997 года по 10-е октября 2008 года сильные события составляли 1.6% от числа всех отсчетов, наблюдалось 10 объектов, соответствовавшие таким кризисам американского фондового рынка, как «черный вторник» 1997 года, кризис августа 1998 года, серия кризисов 2000 года, падение после 11 сентября 2001 года и др.

### 2.3 Определение прогнозирующих функционалов

Для построения прогнозирующих функционалов используется выборочная функция распределения (ВФР) величины  $v_i$ , но уже для  $m = 7$  см. (1), посчитанная по предыдущим 2.5 месяцам наблюдений (что составляет 68 отсчетов). Это означает, что для вычисления значений функционалов на сегодняшний день используются последние 61 значений  $v_i$ , иными словами, скользящее окно алгоритма имеет ширину  $\text{win}=61$  торговых дней. Рассматривая величины  $v_i$ , мы уже имеем дело с детрендриванным рядом, трендом в данном случае является скользящее среднее, а ряд остатков преобразуется в  $v_i$  посредством формулы (1).

В ходе исследования было установлено, что при приближении к объекту ВФР  $v_i$  и соответствующая ей гистограмма меняют свою форму (см. рис. 4).

Подобные изменения формы гистограммы имели место и для других предкризисных периодов и характеризовались следующими особенностями:

- 1) Смещение гистограммы в сторону отрицательных значений  $v_i$ .

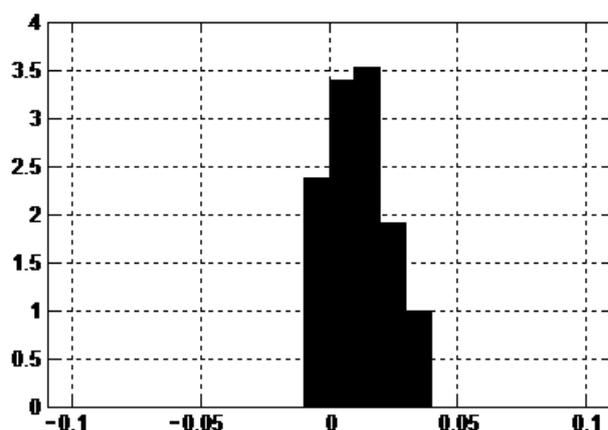
- 2) Увеличение доли относительно более сильных отрицательных  $v_i$ .
- 3) «Расползание» гистограммы.

На основании этих наблюдений составляются два функционала, F1 и F2, которые следят за изменением формы ВФР. F1 рассчитывается по относительному числу больших отрицательных событий с учетом дисперсии ВФР и медианы.

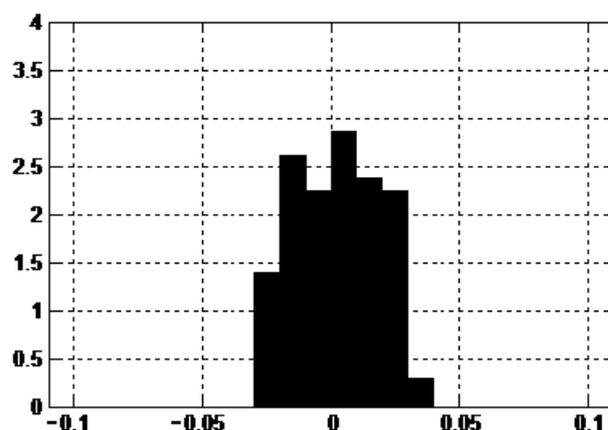
$$F1 = \mu \cdot \frac{S1}{S0} \cdot \delta - (2 - \mu) \cdot med,$$

где S1 – число отрицательных  $v_i$  в выборке, абсолютная величина которых больше среднеквадратичного отклонения  $\delta$  выборки, S0 – число отрицательных  $v_i$  (см. рис. 5), абсолютная величина которых меньше  $\delta$ ,  $med$  – медиана выборки,  $\mu$  – весовой коэффициент, принимающий значения от 0 до 2.

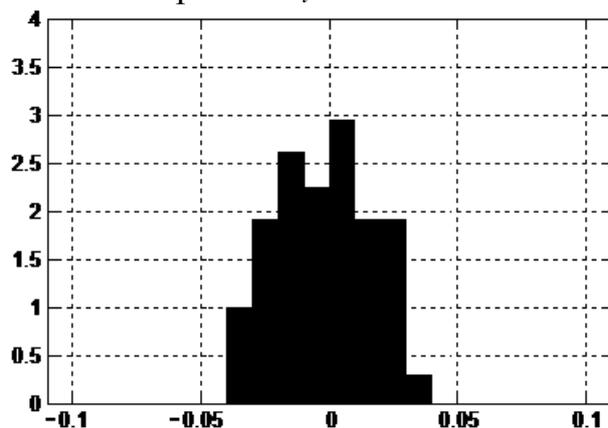
Отношение  $\frac{S1}{S0}$  характеризует относительное число больших отрицательных событий, учитывая некоторым образом ненормальность распределения (для нормального распределения значение этого отношения должно быть около 0.46, при приближении к объекту в большинстве случаев наблюдалось увеличение этой величины до 0.6). Умножение на  $\delta$  позволяет учесть «расползание»



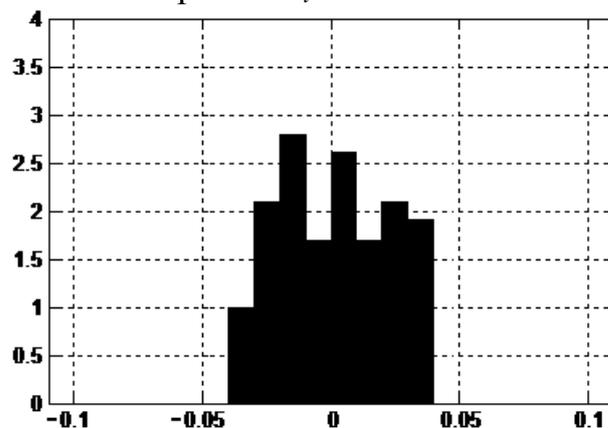
Гистограмма  $v_i$  на 18.04.1998 г.



Гистограмма  $v_i$  на 09.06.1998 г.



Гистограмма  $v_i$  на 07.07.1998 г.



Гистограмма  $v_i$  на 28.07.1998 г.

Рис. 4. Изменение формы гистограммы  $v_i$  (для  $m=20$ ,  $win=48$ , см. (1)) перед кризисом в начале августа 1998 года. При данных значениях  $m$  и  $win$  визуальные изменения в гистограмме более заметны, при прогнозе же используются  $m=7$  и  $win=61$ . Ось ординат выбрана в логарифмическом масштабе.

распределения и привести части функционала F1 к одной размерности. Медиана измеряет смещение гистограммы в область отрицательных значений  $v_i$ .

Эволюция составных частей функционала F1 показана на рис. 5. Как видно по рис. 6, при приближении к объекту увеличивается значение каждого функционала, но часть объектов лучше предсказывается величиной  $\frac{S1}{S0} \cdot \delta$ , а часть величиной

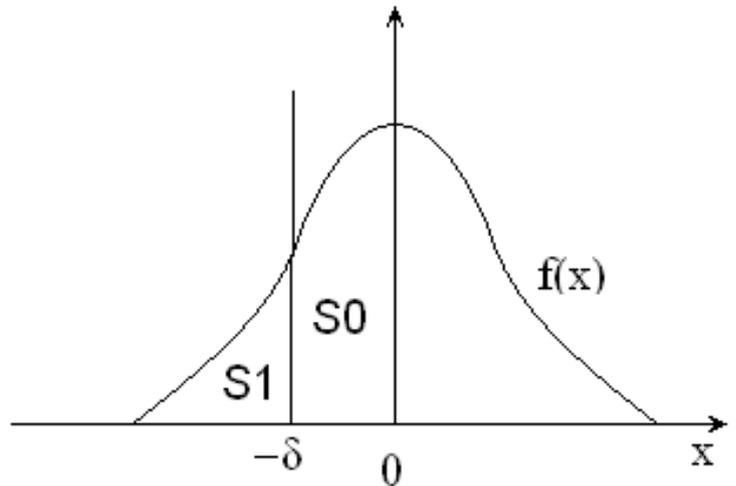


Рис. 5. Определение величин S1 и S0 на графике функции плотности ВФР  $f(x)$ .

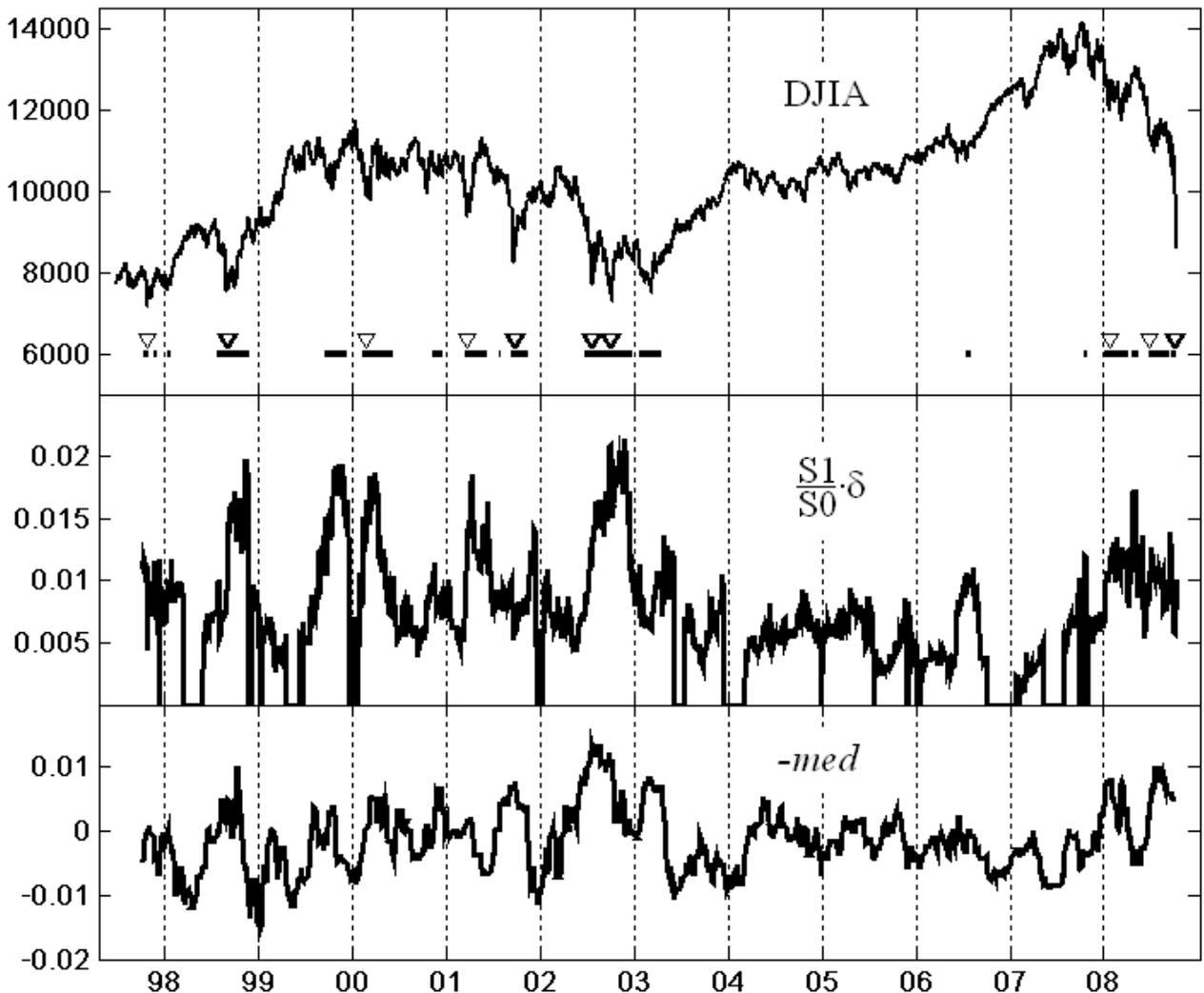


Рис.6. Результат вычисления составных частей F1. На рисунке показаны значения (вертикальная ось) индекса DJIA и величин  $\frac{S1}{S0} \cdot \delta$  и  $-med$ . По оси абсцисс отложено время. Под индексом сплошными линиями обозначено время тревоги, правила объявления которой будут представлены ниже, и треугольниками – объекты.

–*med*, поэтому в качестве прогнозного функционала F1 была предложена их взвешенная сумма.

При приближении к объекту гистограмма никогда не имеет достаточно компактной формы, на основании этого вводится дополнительно функционал F2, чтобы ликвидировать часть ложных тревог. F2 корректирует работу функционала F1, рассчитывая ширину ВФР по квартилям (считается расстояние между перцентилями 25 и 75%).

### 2.3 Объявление тревоги

По одновременному превышению заранее выбранных порогов  $b_1$  и  $b_2$  для F1 и F2, соответственно, объявляется тревога. Она длится до тех пор, пока функционалы не вернуться в «спокойное» состояние ( $F1 < b_1$  или  $F2 < b_2$ ) как минимум на 5 дней. Результат работы функционалов для  $b_1 = 0.011$  и  $b_2 = 0.014$ ,  $\mu = 1.2$  можно видеть на рис. 7.

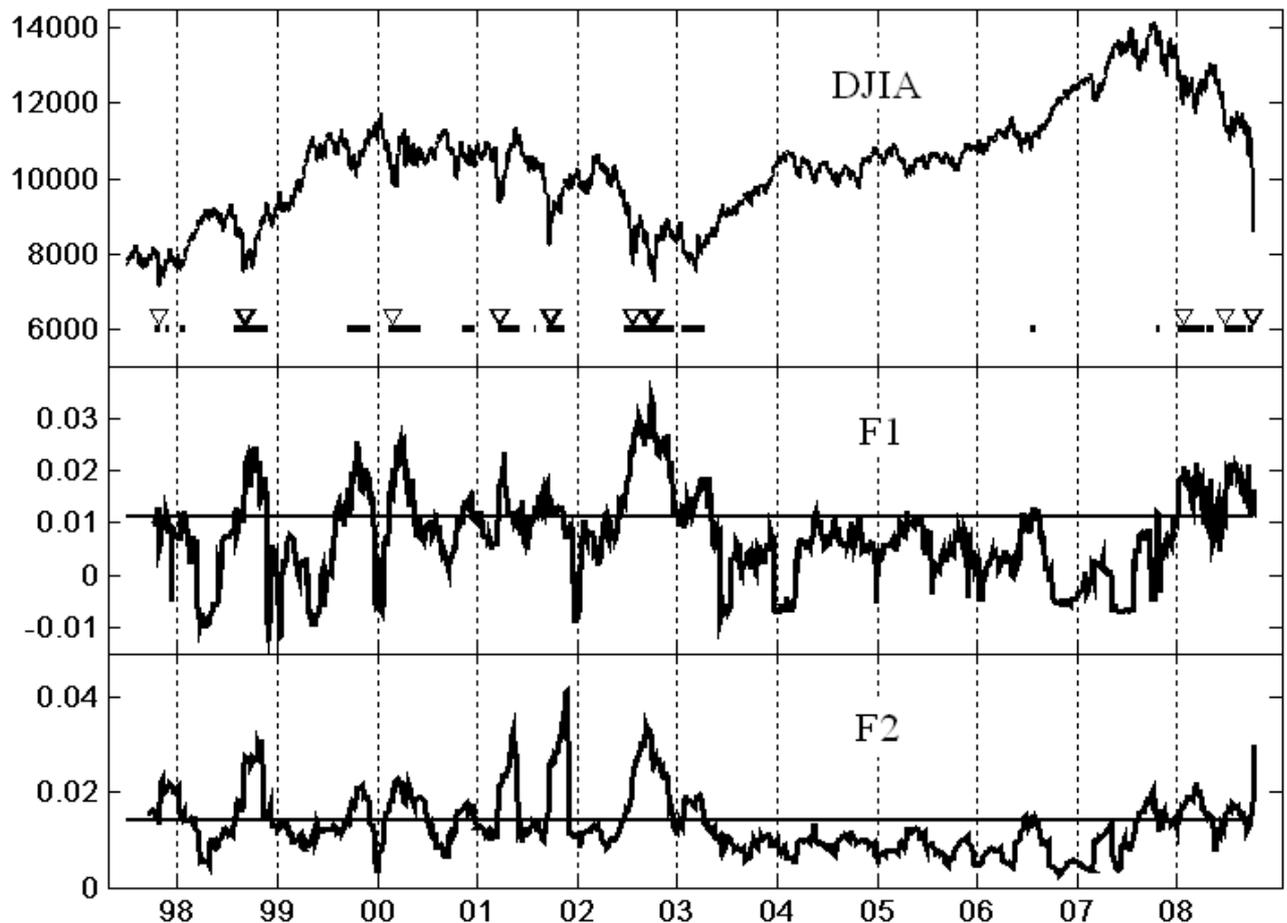


Рис.7. Результат вычисления функционалов F1 и F2. На рисунке показаны значения (вертикальная ось) индекса DJIA и функционалов F1 и F2 в течение всего времени (горизонтальная ось) прогноза. Под индексом сплошными линиями обозначено время тревоги и треугольниками – объекты. На графиках функционалов горизонтальными прямыми обозначены пороги  $b_1$  и  $b_2$ , по одновременному превышению которых объявляется тревога.

### 3. ОЦЕНКА ЭФФЕКТИВНОСТИ РАБОТЫ АЛГОРИТМА

Все тревоги, содержащие объекты прогноза, считаются **успешными**, а не содержащие – **ложными**.

Все объекты, попавшие в тревоги, считаются **предсказанными**, а не попавшие считаются **пропусками цели**.

Пусть  $N$  – это число объектов,  $N^-$  – число пропусков цели. Тогда определим одну из характеристик качества алгоритма, как  $n = \frac{N^-}{N}$ . Пусть  $T$  – общее время прогноза, а  $T_a$  – суммарное время тревоги, тогда  $\tau = \frac{T_a}{T}$ .

Эффективность прогноза удобно оценивать по  $n$ - $\tau$ -диаграмме [12]. При  $n = 1$ ,  $\tau = 0$  имеем предельный случай, когда пропускаются все события при времени тревоги, равном нулю («стратегия оптимиста»), а при  $n = 0$ ,  $\tau = 1$  предсказываются все события при постоянной тревоге («стратегия пессимиста»). Интегральное качество прогноза при изменениях свободных параметров алгоритма можно оценить величиной  $\varepsilon = \min(n+\tau)$  или, при разумных изменениях параметров, средним значением  $n+\tau$ . Условие  $n+\tau < 1$  ( $\varepsilon < 1$ ) соответствует нетривиальному прогнозу и чем меньше  $\varepsilon$ , тем лучше прогноз. Для прогноза землетрясений значения  $\varepsilon$  находятся в пределах 0.32-0.5.

Данный алгоритм оказался способным предсказывать **все** объекты, согласно нашему определению (то есть  $n = 0$ ), величина  $\tau = 0.22$  (на основном интервале при  $\text{win}=60$ ,  $\mu=1.25$ ,  $b1=0.0131$ ,  $b2=0.0143$ ),  $n = 0$ , а значит  $\varepsilon = 0.22$ .

Однако точечная оценка мало говорит о качестве алгоритма. Один из необходимых тестов алгоритмов прогноза – это проверка на чувствительность результатов относительно малого изменения параметров. Другой тест – проверка результатов работы алгоритма на материале, не входящем в основной отрезок, на котором производилась настройка параметров. В качестве такого материала рассмотрим расширенный временной интервал: 31.12.1984–06.02.2009 (в отли-

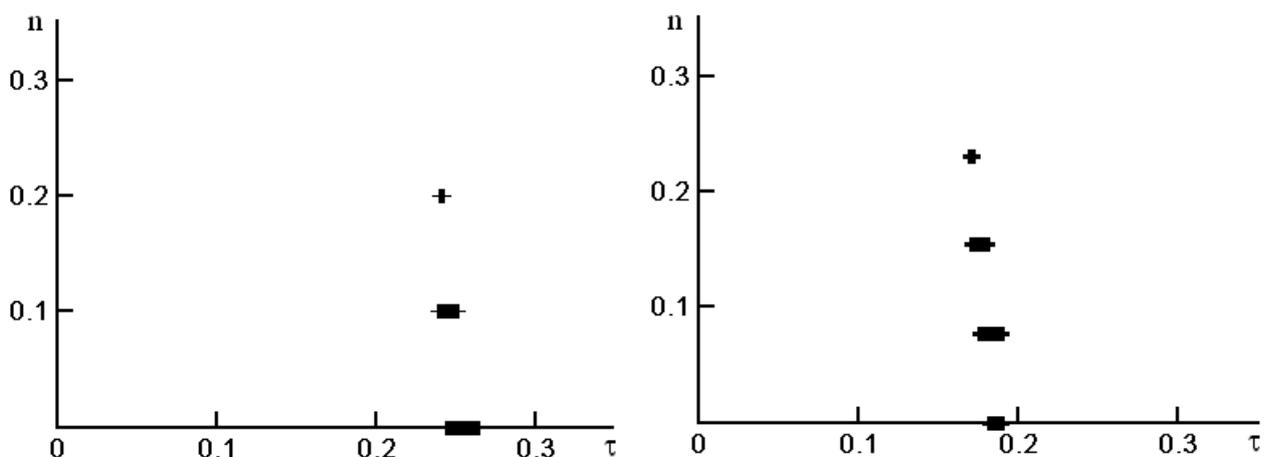


Рис. 8.  $n$ - $\tau$ -диаграмма на временном отрезке 31.12.1984–06.02.2009, справа – на отрезке 01.07.1997–10.10.2008 при изменении следующих параметров:  $\text{win} = 60$ -62,  $\mu = 1.07$ -1.08,  $b2 = 0.0138$ -0.0140,  $b1 = 0.0103$ -0.0113.

чие от основного 01.07.1997–10.10.2008). На нем появляются три новых объекта: кризис 1987 года, кризис 1991 года, связанный с войной в Персидском заливе, и последнее резкое падение индекса в 2009 году. На рис. 8 изображены  $n$ - $\tau$ -диаграммы прогнозов, выдаваемых алгоритмом, при небольшом изменении определяющих параметров.

Диаграммы демонстрируют устойчивость поведения и сохранение алгоритмом своей точности (в том числе способности определять все объекты) при тестировании на новом материале. Успешно сделано предсказание «назад» кризиса 1987 года.

Еще одним распространенным тестом для алгоритмов прогноза является эксперимент по изменению объекта. В данном случае мы исследуем зависимость  $\varepsilon = \min(n+\tau)$  и  $\eta$  (относительного числа ложных тревог) от порога при определении сильного события (переменная *barrier*). Из рис. 9 видно, что значение порога события  $-0.08$  является оптимальным для данного алгоритма в смысле баланса между значением  $\varepsilon$  и  $\eta$ . На рис. 10 построена  $n$ - $\tau$ -диаграмма.

Этот эксперимент также не выявил зависимости успешности прогноза при изменении совместно силы предсказываемого объекта и ширины окна *win*, по которому строится распределения. Лучшие результаты получались для  $\text{win}=61$ - $66$ . Это может говорить либо о характерном времени подготовки сильного события в такой системе, либо о том, что алгоритм лучше всего видит именно эту фазу подготовки, которая занимает около полутора месяцев, предшествующих объекту.

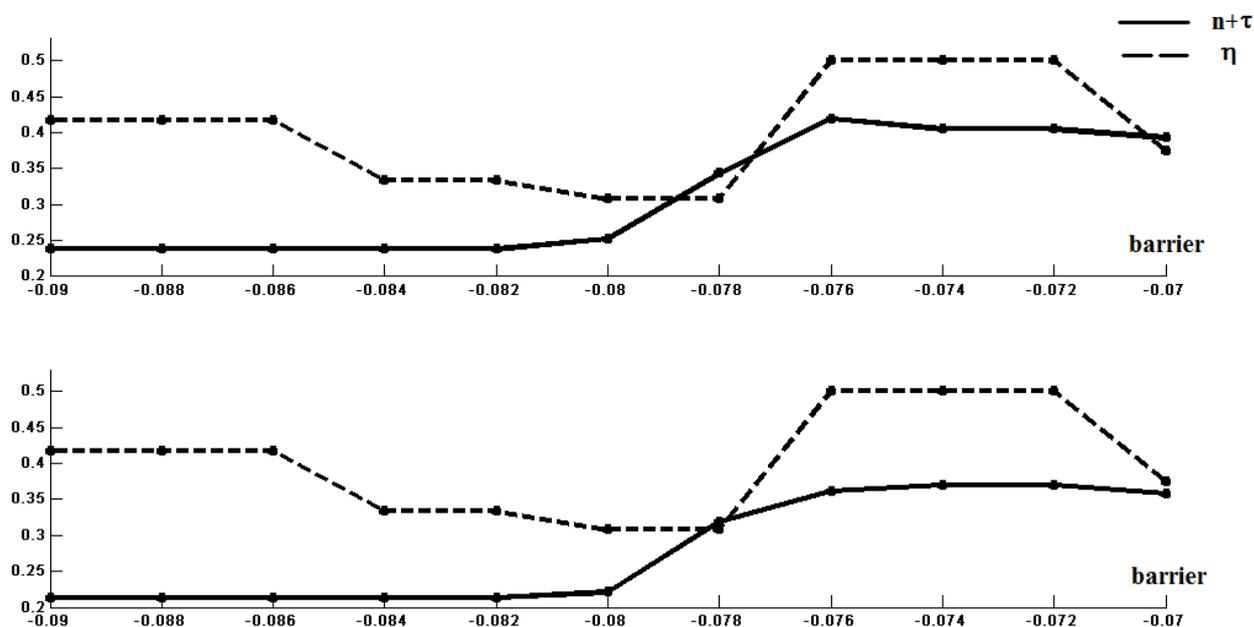


Рис. 9. Произведен расчет результатов работы алгоритма при следующей вариации параметров: порог определения объекта  $\text{barrier} = 0.07$ - $0.09$ ,  $\mu = 0.9$ - $1.8$ ,  $\text{win} = 49$ - $72$ ,  $b2 = 0.009$ - $0.0143$ ,  $b1 = 0.0075$ - $0.0150$ . На рисунках отмечены  $n+\tau$  и  $\eta$  при минимальном  $n+\tau$  и  $\eta$  для каждого значение *barrier*. На верхнем  $\min$  сначала берется по  $\eta$ , а затем по  $n+\tau$ , на нижнем рисунке – в обратной последовательности.

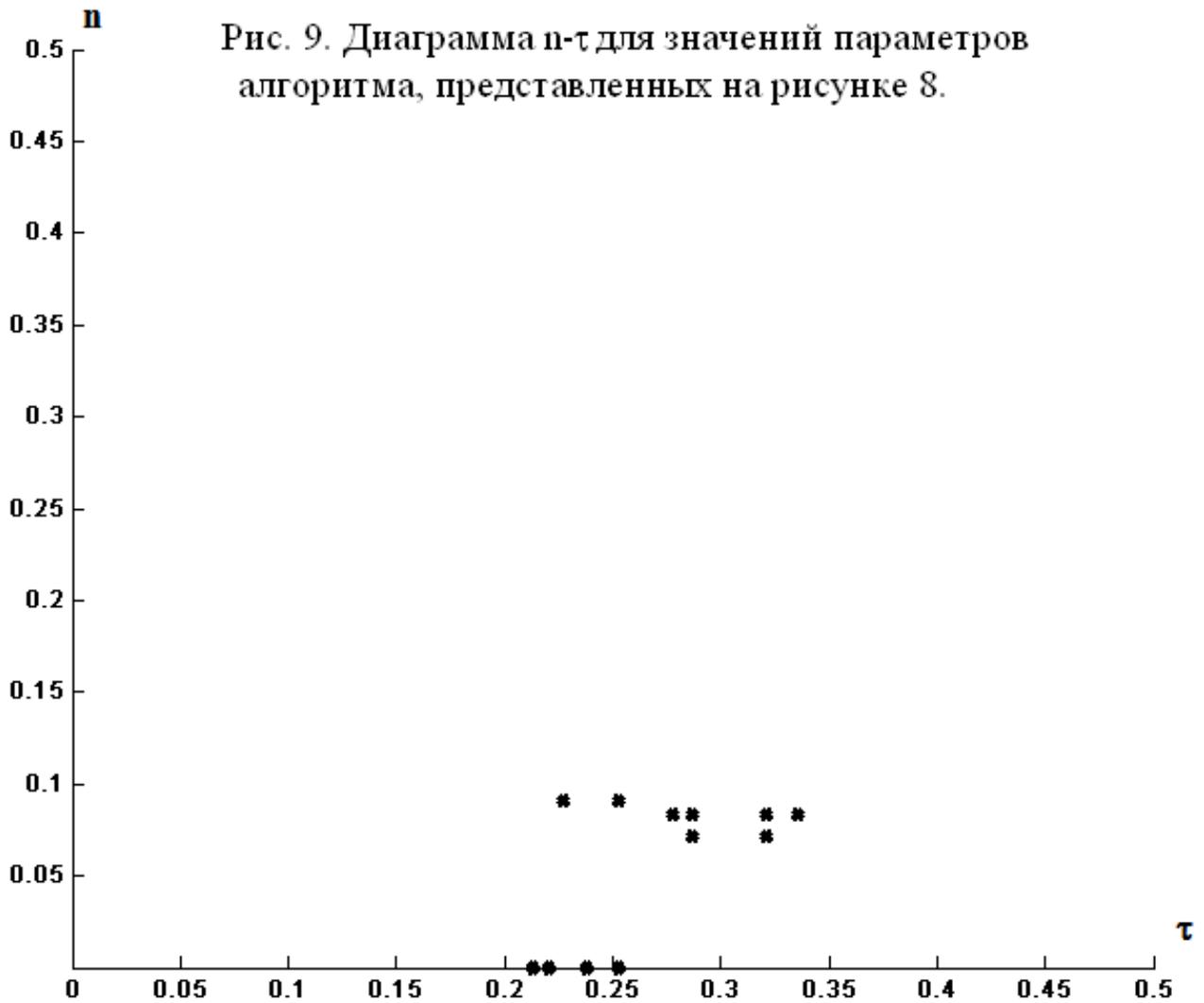


Рис. 10.

#### 4. ТЕСТИРОВАНИЕ ПРАКТИЧЕСКОЙ ПРИМЕНИМОСТИ АЛГОРИТМА

Несмотря на хорошие в математическом (вероятностном) смысле характеристики алгоритма, необходимо ответить на вопрос: является ли информация, которую дает алгоритм, ценной для участников фондового рынка. Для этого необходимо, в конечном итоге, предложить конкретную торговую стратегию и оценить ее успешность.

Главная проблема при этом – фондовый рынок имеет очень большое количество инструментов, и бесконечное многообразие различных типов стратегий, при помощи которых инвесторы решают свои задачи. К тому же зачастую сложно априори учесть все транзакционные издержки или небольшую дополнительную прибыль, возникающую при выплате дивидендов, что может оказаться критическим для оценки успешности торговли. Для того чтобы результаты тестирования предложенной в данной работе простейшей торговой стратегии не были искажены всеми этими факторами, объектом для применения ин-

формации о надвигающемся кризисе был выбран индексный фонд DIAMONDS Trust (DIA), очень точно следующий за индексом DJIA.

Один из главных критериев успешности для инвестора – умение «переиграть» рынок, то есть получить доходность выше той, которую дает рост всего фондового рынка в целом (или, что почти то же самое, рост индекса, который оценивает этот рынок). Исследования показывают, что на практике при долгосрочном инвестировании это удается лишь малой доле игроков (различных фондов), причем, как правило, в разные периоды времени это разные люди, поэтому здесь очень тяжело отличить везение от профессионализма.

Как сказал Уоррен Баффет: «Для абсолютного большинства инвесторов, и институциональных, и частных, самый лучший способ держать обычные акции – это вкладывать деньги в индексные фонды, которые имеют минимальные накладные издержки. Те, кто последуют этим путем, могут быть уверены, что в конечном результате (после выплат всех взносов) получают большую доходность, чем та, которую могут обеспечить большинство профессиональных управляющих».

Таким образом, с одной стороны цены акций DIA будут служить объектом торговли, а с другой стороны, результаты этой торговли можно эффективно сравнивать с доходностью, которую дают именно эти акции.

Как видно из рис. 11, алгоритм хорошо умеет предсказывать начало кризиса. Обладая этой информацией можно построить простейшую стратегию «покупай дешево, продавай дорого». К сожалению, данный алгоритм ничего не может сказать относительно начала выхода из кризиса, поэтому в данной работе применяется простейший прием, который используют трейдеры: если цена после падения при повышении превысит 4-хдневное скользящее среднее, то будем предполагать, что кризис окончен.

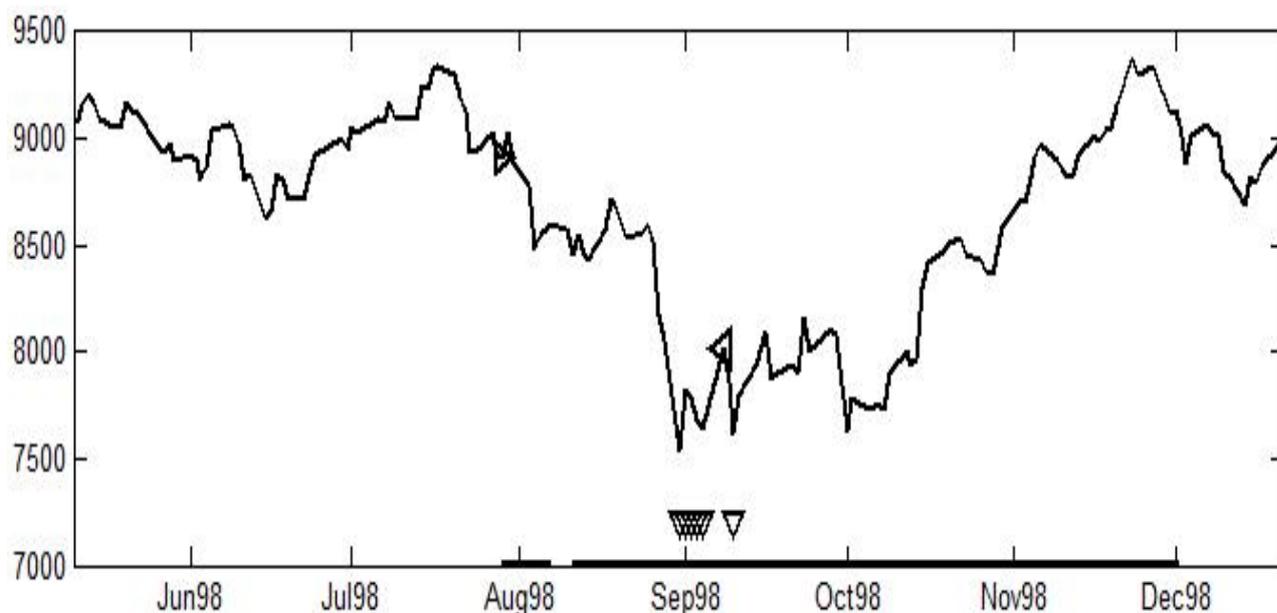


Рис. 11. Применение торговой стратегии. Треугольником, повернутым вправо, обозначена продажа активов, треугольником, повернутым влево, – покупка. На рисунке показан индекс DJIA, внизу жирной горизонтальной линией обозначено время тревоги, треугольники внизу – сильные события.

**Торговая стратегия:**

1) При объявлении тревоги продаем активы. В течение 14 торговых дней ничего не предпринимаем (режим «ожидания»), если не встретится объект прогноза. Если встречаем объект прогноза, то отменяем 14-тидневный режим «ожидания».

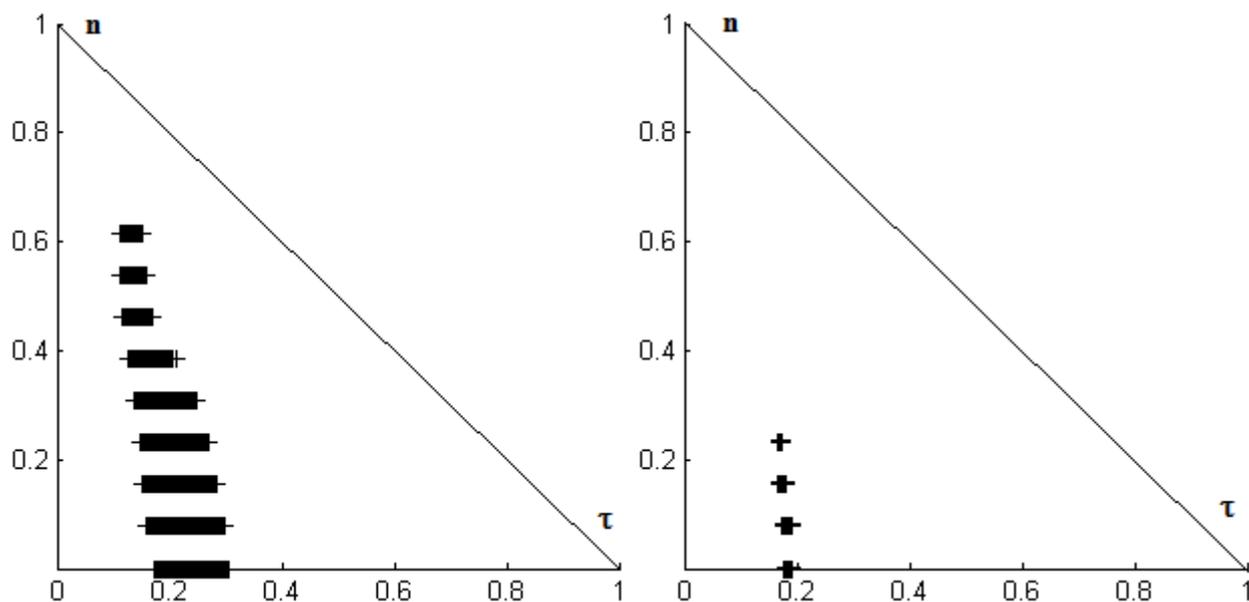


Рис. 12. Слева на рисунке  $n$ - $\tau$ -диаграмма для расширенного интервала времени (31.12.1984–06.02.2009) для следующей вариации параметров:  $\text{win}=59-69$ ,  $\mu=0.9-1.7$ ,  $b_2=0.0135-0.0145$ ,  $b_1=0.0100-0.0135$ . Справа на диаграмме те точки из этого диапазона параметров, для которых торговая стратегия на индексе DJIA давала суммарно более 1700% процентов доходности за все время (что составляет около 12% годовых).

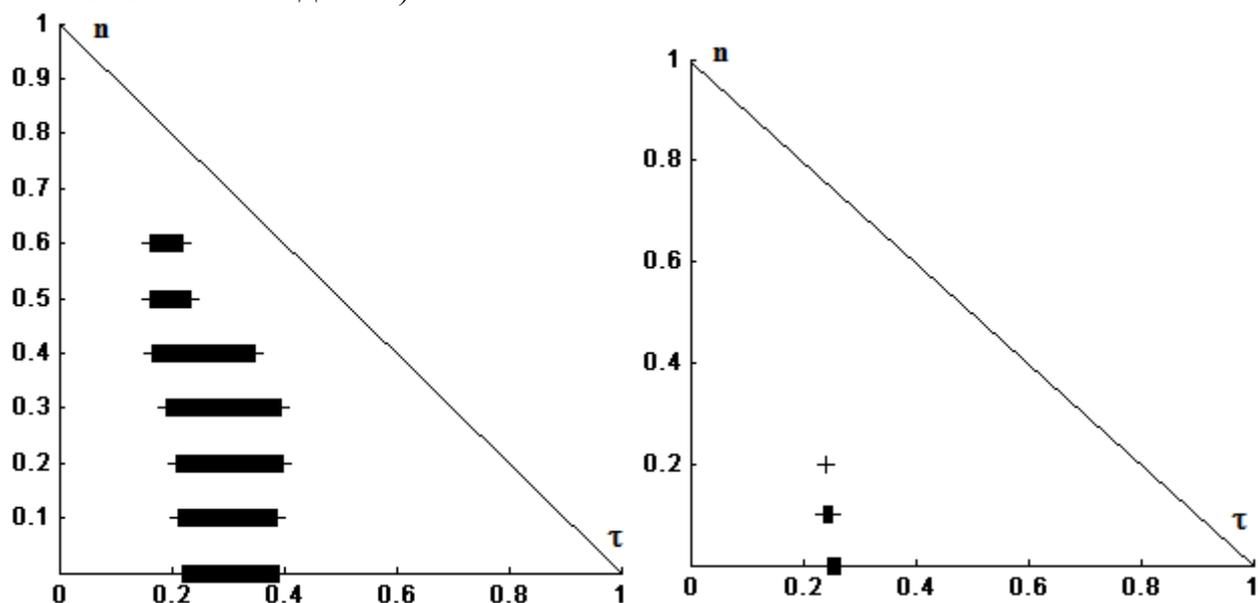


Рис. 13. Слева на рисунке  $n$ - $\tau$ -диаграмма для основного интервала времени для следующей вариации параметров:  $\text{win}=59-69$ ,  $\mu=0.9-1.7$ ,  $b_2=0.0135-0.0145$ ,  $b_1=0.0100-0.0135$ . Справа на диаграмме те точки из этого диапазона параметров, для которых торговая стратегия на индексе DJIA давала более 180% процентов доходности за все время (что составляет около 10% годовых).

2) Если мы не находимся в режиме «ожидания», мы продали свои активы, и цена превышает 4хдневное скользящее среднее на 0.12%, покупаем активы обратно.

Прежде чем обсуждать полученные результаты для конкретных значений параметров (параметров и торговой стратегии, и алгоритма), интересно ответить на следующий вопрос: как связана успешность работы торговой стратегии с эффективностью алгоритма.

Рис. 12 и 13 показывают, что для построения успешной торговой стратегии данного типа важна эффективность алгоритма предсказания. В силу того, что предложенная торговая стратегия не обладает большой гибкостью и в переборе не были использованы все возможности параметрической подстройки, не любое хорошее с точки зрения n-т сочетание параметров алгоритма обеспечивает высокую прибыль при торговле.

С точки зрения практического применения информации, полученной при помощи алгоритма, есть два основных параметра, определяющих возможный доход и риски, с ним связанные – относительное число ложных тревог  $\eta$ , от которого зависят возможные потери, и прибыль (profit) торговой стратегии.

Понятно, что для каждой конкретной торговой стратегии кривая риск-доходность будет разная, поэтому, чтобы примерно охарактеризовать потенциал, который имеет данный алгоритм, исследуем, как меняются значения пары

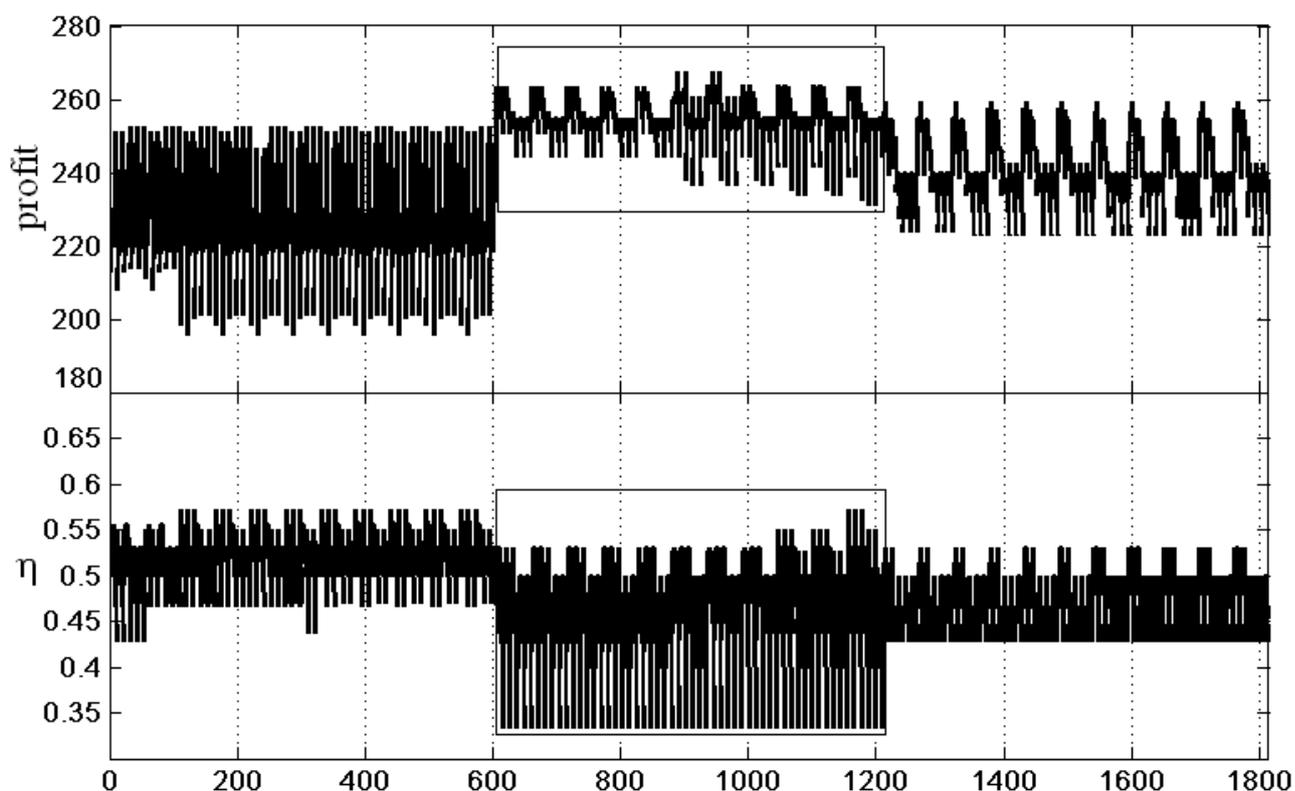


Рис. 14. Исследованы вариации значений ( $\eta$ , profit) при изменении параметров алгоритма  $win=60-62$ ,  $\mu=1.07-1.08$ ,  $b2=0.0138-0.0140$ ,  $b1=0.0103-0.0113$  для основного промежутка времени. Значения  $\eta$  и profit (в процентах) для каждого набора параметров отмечаются по вертикальной оси, по горизонтальной отмечается порядковый номер такого набора в цикле.

( $\eta$ , profit) при различных значениях параметров. На приведенных ниже рисунках 12 и 13 показаны значения ( $\eta$ , profit) при изменении win,  $\mu$ , b2, b1 в цикле (во внешнем цикле стоит win, в цикле, вложенном в него  $\mu$ , далее по порядку идут b2 и b1). Переменная win меняется с шагом 1, что определяет три последовательных участка графика (номера наборов параметров от 1 до 610, от 611 до 1220, от 1221 до 1830) на рис. 14 и десять «волн» на рис. 15. «Волны» на рисунке 12 соответствуют изменению параметра  $\mu$  в цикле.

На рис. 14 показан вариант с небольшой вариацией параметров, которая дает стабильные высокие значения прибыли и низкие значения ложных тревог одновременно.

На рис. 15 изображен вариант с очень большим диапазоном изменения параметров. Прямой вертикальной линией обозначено местоположение области, которая соответствует множеству параметров, отмеченных прямоугольной рамкой на рисунке 12. «Волны» на нижнем графике обусловлены изменением весового коэффициента  $\mu$ : первая часть первого функционала дает значительно больше ложных тревог, чем вторая, поэтому, когда  $\mu$  велико и больше вес первой части, возрастает относительное число ложных тревог.

Таким образом, удастся выделить области параметров с относительно небольшим значением  $\eta=0.35-0.50$ , и значениями profit=255-265%.

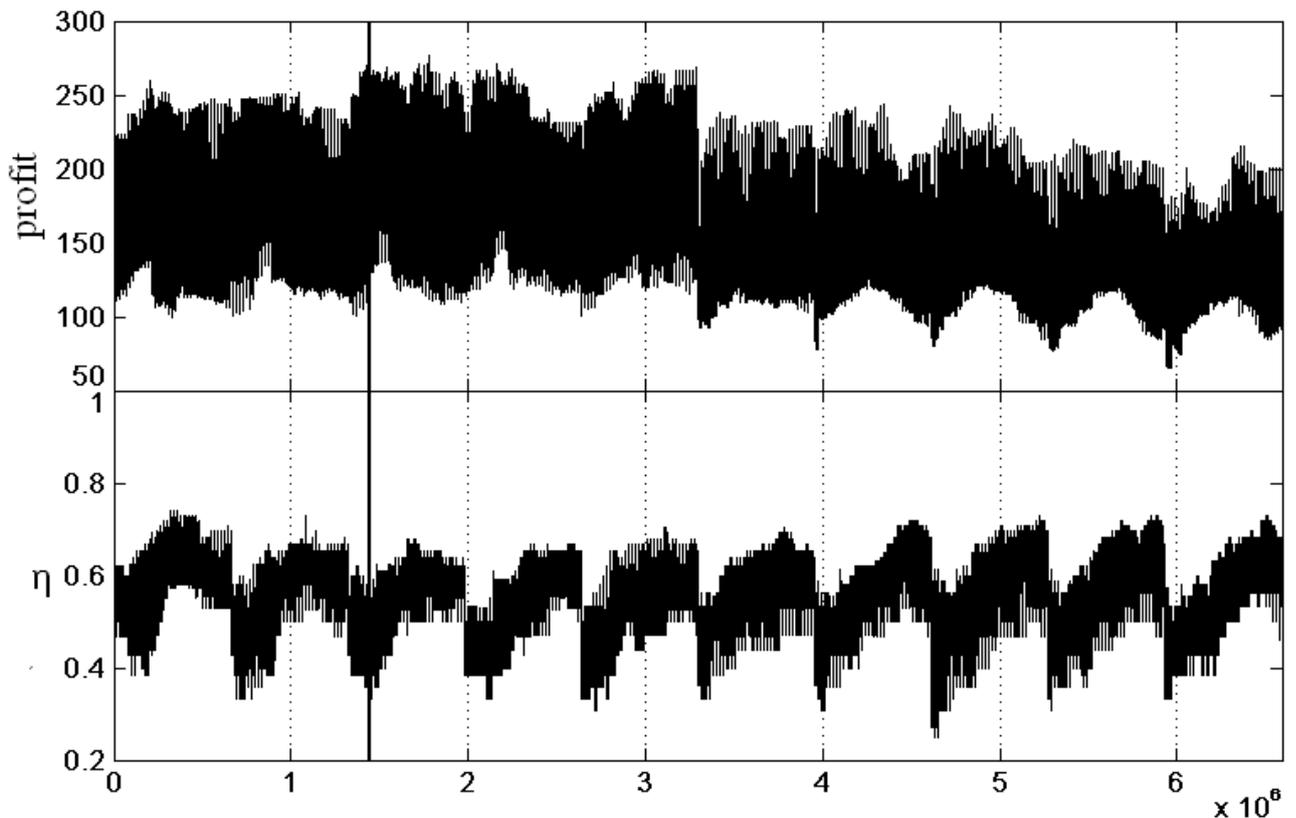


Рис. 15. Исследованы вариации значений ( $\eta$ , profit) при изменении параметров алгоритма win=59-68,  $\mu=0.9-1.7$ , b2=0.0135-0.0145, b1=0.0100-0.0135 для основного промежутка времени. Значения  $\eta$  и profit для каждого набора параметров отмечаются по вертикальной оси, по горизонтальной отмечается порядковый номер такого набора.

Для оценки успешности торговой стратегии будем сравнивать результаты с доходностью индексного фонда DIA. При этом здесь есть один важный момент. Идеальными условиями для применения стратегии «купил и держи» является период до начала финансового кризиса 2008 года. Максимум индекса был достигнут 9 октября 2007, эту дату и будем использовать для измерения доходности до кризиса.

Для значений параметров  $w_{in}=61$ ,  $\mu=1.07$ ,  $b_2=0.0138$ ,  $b_1=0.0101$  получаем следующие результаты, приведенные в таблице.

Стратегия	Доходность, измеренная до кризиса		Доходность, измеренная 05.11.2008 г.	
	Общая за 11 лет	В % годовых	Общая за 11 лет	В % годовых
«Купил и держи»	114%	7,1	68%	4,77
Торговая стратегия	257%	12,3	192%	10,4

Таким образом, использование торговой стратегии позволяет не только значительно переиграть рынок, но и смягчить последствия финансового кризиса: потери общей прибыли за 11 лет снизились с 40% до 25%.

## ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

В данной работе введено и обосновано формальное определение «сильного» падения финансового индекса DJIA. При помощи методологии распознавания образов редких событий предложен алгоритм прогноза таких падений, основанный на изменении формы выборочной функции распределения детрендированного ряда цен закрытий при приближении к объекту. Алгоритм способен предсказывать все падения индекса, которые являются предметом прогноза, при этом суммарное время тревоги (времени, в течение которого ожидается наступление события) не превышает одной четвертой части рассматриваемого временного интервала, а доля ложных тревог не превышает одной трети. Для оценки практической значимости результатов предложена и оценена простейшая торговая стратегия, которая способна значительно переигрывать рынок, смягчая последствия кризисов. Проведено исследование взаимосвязи основных характеристик качества работы алгоритма. Все выполненные численные эксперименты подтверждают нетривиальность полученных результатов.

Однако окончательную оценку качества прогноза может дать только проверка алгоритма в режиме работы в реальном времени. С момента выбора значений параметров произошло три сильных падения, классифицируемых согласно нашему определению как объекты, и все они были предсказаны.

Стоит также отметить недостатки, присущие данному алгоритму. Определение кризиса, рассматриваемое в данной статье, оставляет временной зазор между моментом начала падения рынка и собственно моментом, когда фиксируется появление объекта прогноза. В небольшом числе случаев алгоритм объявляет тревогу уже во время падения рынка, но до того, как это падение превы-

сит критический порог, что считается успешным предсказанием. Так же, исходя из физических соображений, было бы естественно ожидать, что редким и крупным событиям, являющиеся предметом исследования, должна предшествовать длительная подготовка. Однако данный алгоритм не способен делать прогноз, например, за месяц до наступления события.

Данная работа была построена на знании лишь одного временного ряда цен. В будущем перспективной выглядит тема исследования той информации, которая может быть извлечена из других существенных для кризисных явлений на финансовых рынках данных, таких как объем продаж на фондовых рынках, ставки по кредитам, государственным облигациям и другие макроэкономические показатели.

## СПИСОК ЛИТЕРАТУРЫ

1. Шарп У., Александер Г., Бэйли Дж. Инвестиции. М.: ИНФРА-М, 2001.
2. Мерфи Дж. Технический анализ фьючерсных рынков: теория и практика. М.: Диаграмма, 1998.
3. Gurkaynak, Refet S., Econometric Tests of Asset Price Bubbles: Taking Stock, Journal of Economic Surveys 22 (1), 166-186, 2008.
4. B. Mandelbrodt, The variation of certain speculative prices. J. Bus. 67 (1963), 394-419.
5. Gopikrishnan P., Meyer M., Amaral L.A.N., Stanley H.E. Inverse cubic law for the distribution of stock price variation // Eur. Phys. J.B. 1998. V. 3. 139-140.
6. Mantegna R. N., Stanley H.E. Scaling behavior in the dynamics of an economic index // Nature. 1995. V. 376. 46-49
7. Mantegna R.N., Stanley H.E. An introduction to econophysics. Correlation and Complexity in Finance. Cambridge: Cambridge Univ. Press, 2000.
8. A. Johansen and D. Sornette, Evaluation of the quantitative prediction of a trend reversal on the Japanese stock market in 1999, Int. J. Mod. Phys. C. 2000. V. 11(2), 359-364
9. Zhou, W.-X. and D. Sornette, Testing the stability of the 2000 US stock market "antibubble", Physica A. 2005. 348, 428-452
10. Keilis-Borok VI, Soloviev AA (eds) Nonlinear Dynamics of the Lithosphere and Earthquake Prediction, 2003. Springer-Verlag, Berlin-Heidelberg
11. Кузнецов И.В., Родкин М.В., Серебряков Д.В., Урядов О.Б. Иерархический подход к динамике преступности // Новое в синергетике. Новая реальность, новые проблемы, новое поколение. Часть 1. М.: Радиотехника, 2006. С.381-396 (Сер. Фракталы. Хаос. Вероятность).
12. Молчан Г.М. Оптимальные стратегии в прогнозе землетрясений. Современные методы интерпретации сейсмологических данных // Выч. сейсмология. №24. С.3-18. – М: Наука, 1991.

**Содержание.**

<b>Введение .....</b>	<b>3</b>
<b>1. Общая постановка задачи .....</b>	<b>5</b>
1.1 Описание данных и процессов .....	5
1.2 Описание общей методологии предсказания.....	7
<b>2. Алгоритм предсказания .....</b>	<b>8</b>
2.1 Определение объекта прогноза.....	8
2.3 Определение прогнозирующих функционалов .....	9
2.3 Объявление тревоги .....	12
<b>3. Оценка эффективности работы алгоритма .....</b>	<b>13</b>
<b>4. Тестирование практической применимости алгоритма.....</b>	<b>15</b>
<b>Основные результаты работы .....</b>	<b>20</b>
<b>СПИСОК ЛИТЕРАТУРЫ .....</b>	<b>21</b>