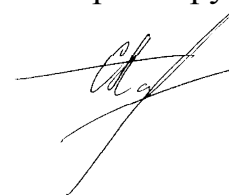


На правах рукописи



Павлов Сергей Владимирович

**ТЕХНОЛОГИЯ ПРОГНОЗИРОВАНИЯ  
СЛОЖНЫХ НАБЛЮДАЕМЫХ ВРЕМЕННЫХ РЯДОВ**

05.13.01 – Системный анализ, управление и обработка информации  
(по отраслям: информатика, вычислительная техника и управление)

Автореферат диссертации на соискание ученой степени  
кандидата технических наук

Красноярск – 2007



## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Обоснование актуальности работы.** В диссертации рассматривается проблематика сложных объектов в аспекте прогнозирования их наблюдаемой динамики. Под сложными объектами понимаются объекты, «погруженные» в иерархически структурированную окружающую среду, являющиеся открытыми и состоящими из большого количества компонент (включая слабо формализуемые, ненаблюдаемые и неуправляемые), отношения между которыми не могут быть строго установлены.

В диссертации в качестве представителей класса сложных объектов для проведения исследований выбраны: «Международный валютный рынок», «Международный фондовый рынок» и «Мировая экономика».

Выбранные для исследования сложные объекты финансово-экономической природы характеризуются динамической и организованной сложностью, наличием рефлексии, поведения и эффектов самоорганизации, мультифрактальностью структуры динамики и т.д.

Наблюдаемая динамика свойств выбранных сложных объектов является, с одной стороны, композицией детерминированных, стохастических и хаотических составляющих, а, с другой стороны, наложением движений разных масштабов, отличающихся по степени и частоте воздействия.

В диссертации решается задача прогнозирования наблюдаемых свойств сложных объектов финансово-экономического типа (котировки валют, курсы акций, индексы) на основе имеющейся информации о ретроспективных наблюдениях этих свойств – сложных наблюдаемых временных рядов.

**Объектом исследований** в диссертации является сложный наблюдаемый временной ряд – конечное и дискретное множество упорядоченных данных числовой природы единого информационного ранга.

Сложные наблюдаемые временные ряды характеризуются уникальной траекторией динамики, наличием сингулярностей, структурированностью значений по уровням, полимодальностью распределения значений и распределением с «тяжелыми хвостами» для приращений. Указанные характеристики для сложных наблюдаемых временных рядов являются уникальными на различных временных масштабах (минутные, часовые, дневные и другие значения), т.е. наблюдаемая динамика сложных объектов не является масштабно-инвариантной.

Актуальность решения задачи прогнозирования котировок валют, курсов акций и ключевых индексов определяется необходимостью в наличии информации прогностического характера при решении задач планирования и управления в области производства, инвестиций и др. для принятия упреждающего решения.

Использование мощной методологии теории случайных процессов и теории сигналов для прогнозирования сложных наблюдаемых временных рядов ограничено, поскольку эти ряды доступны в виде единственной траектории (ансамбль реализаций принципиально отсутствует), а их эргодическое «рассмотрение» неприемлемо, поскольку все характеристики (среднее,

дисперсия) сильно изменчивы, а корреляционные характеристики в принципе не могут быть корректно построены.

Использование методик кусочно-линейного моделирования, сингулярно-спектрального анализа, нейросетей и группового учета аргументов для прогнозирования ценовых последовательностей (на валютных и фондовых рынках) позволяет добиться удовлетворительного результата на несколько точек вперед (но не более 10).

Применение распространенных пакетов статистического анализа временных рядов (Эвриста, СТАТИСТИКА, PolyAnalyst, Vanguard и др.) не позволяет успешно решить задачу прогнозирования сложных наблюдаемых временных рядов (например, рядов котировочного типа), поскольку в большинстве таких пакетов применение методов прогнозирования (сводимых, в основном, к моделям авторегрессии, скользящего среднего и их обобщениям) осуществляется к не преобразованному временному ряду без специальной предварительной обработки.

В областях финансовой математики и финансовой инженерии существенный вклад внесли Акелис, Александер, Басс, Башелье, Блэк, Больерслев, Вильямс, Гирсанов, Кендалл, Крянев, Мандельброт, Марковиц, Медведев, Нейштадт, Петерс, Самуэльсон, Сорос, Терпугов, Фама, Фишер, Четыркин, Шарп, Ширяев, Шоулз, Элдер и др. Перечисленными учеными были заложены основы стохастической финансовой математики и математический аппарат ведения актуарных расчетов. Задача прогнозирования ценовых последовательностей была ими только сформулирована, но не решена.

Не существует моделей валютных, фондовых и других рынков, степень адекватности которых достаточна для прогнозирования ценовых последовательностей. Для некоторых сложных объектов физической природы (например, атмосфера) удалось построить модели, позволяющие осуществить краткосрочный прогноз погоды. В то же самое время, для сложных геофизических объектов не существует моделей, позволяющих построить удовлетворительный прогноз землетрясений.

**Основная идея** диссертации заключается в том, что прогнозирование сложных наблюдаемых временных рядов (на примере ценовых последовательностей) в условиях реального режима времени возможно на основе информационно-технологического подхода, базирующегося на системе наблюдаемых данных, а не на системе моделей объектов, порождающих эти данные. Назначение технологического подхода заключается в последовательном применении к сложным наблюдаемым временным рядам операционных преобразований, направленных на постепенное снижение сложности исследуемых временных рядов.

Исходным положением является то, что прогнозирование сложных наблюдаемых временных рядов доступными средствами может быть построено только с точки зрения определенной целевой функции, в качестве которой в диссертации выбрано обнаружение изменений в направлении движения траектории динамики наблюдаемого свойства сложного объекта.

В диссертации предлагается следующая структура технологии прогнозирования сложных наблюдаемых временных рядов:

1. повышение целостности сложных наблюдаемых временных рядов на основе процедуры сглаживания – цифровой фильтрации;
2. переход от сглаженной траектории к временному ряду-заместителю на основе приращений;
3. повышение целостности временного ряда-заместителя за счет процедуры сглаживания – цифровой фильтрации;
4. прогнозирование сглаженного временного ряда-заместителя;
5. интерпретация полученных прогнозных значений относительно исходной траектории сложного наблюдаемого временного ряда.

Разрабатываемая технология прогнозирования сложных наблюдаемых временных рядов применима не только к рассматриваемым в диссертации объектам финансово-экономического типа, а к более широкому классу сложных объектов, структура которых хорошо описывается иерархическими моделями и которые доступны в виде наблюдаемых траекторий с наличием определенных стереотипов поведения.

В диссертации вводится понятие «целостность временного ряда», означающее согласованность расположения соседних значений: чем сильнее «соседи» зависят друг от друга, тем целостность выше. Целостность временного ряда характеризует то, насколько отдельные значения единым образом формируют целое – временной ряд.

Процедура прогнозирования на четвертом этапе технологии осуществляется на основе рассмотрения работ таких ученых, как Бокс, Браун, Винер, Голяндина, Грешилов, Данилов, Дженкинс, Жиглявский, Зайченко, Ивахненко, Колмогоров, Котюков, Мюллер, Степашко, Хеннан, Хольт и Уинтерс. Для получения прогнозных значений были выбраны модели авторегрессионного типа.

**Предметом исследований** в диссертации являются целостные характеристики сложных наблюдаемых временных рядов, которые могут быть интерпретированы с точки зрения возможности прогнозирования этих рядов моделями в классе авторегрессионных.

**Цель исследований** заключается в построении информационной технологии прогнозирования сложных наблюдаемых временных рядов с точки зрения обнаружения изменений в направлении движения траектории этих рядов.

Достижение поставленной цели предполагается осуществить на основе решения следующих **задач**:

1. разработка алгоритмов оценивания целостности сложных наблюдаемых временных рядов, позволяющих оценивать меру «зазубренности» значений временных рядов на разных иерархических масштабах;
2. разработка алгоритмов повышения целостности сложных наблюдаемых временных рядов, позволяющих осуществить сглаживание временных рядов с учетом принципа рефлексии, выражающегося в

рекуррентности алгоритмов цифровой фильтрации, и с учетом принципа иерархии;

3. разработка методики иерархического прогнозирования сложных наблюдаемых временных рядов;

4. программная реализация разработанной информационной технологии прогнозирования сложных наблюдаемых временных рядов, включающей решения первых трех задач.

**Методологические основания исследований** базируются на теории вероятностей и математической статистике, теории случайных процессов, цифровой обработке сигналов, линейной алгебре, методах прогнозирования временных рядов и системологических представлениях Дж. Клира.

К основным **результатам** диссертационного исследования, выносимым на защиту, относятся:

1. алгоритмы оценивания целостности временных рядов – спектр показателей целостности;

2. алгоритмы цифровой фильтрации временных рядов, учитывающие принцип рефлексии;

3. методика иерархического прогнозирования временных рядов, основанная на формировании временных рядов разных масштабов по времени;

4. информационная технология прогнозирования сложных наблюдаемых временных рядов, включающая все полученные результаты.

**Научная новизна** полученных результатов заключается в построении обобщенной технологии прогнозирования наблюдаемых свойств сложных объектов.

**Теоретическая значимость** проведенных в диссертации исследований определяется обоснованностью применения информационно-технологического подхода к решению задачи прогнозирования наблюдаемой динамике сложных объектов. Значение для теории имеет предложенные спектр показателей целостности и алгоритмы цифровой рефлексивной фильтрации временных рядов.

Значение для **практики** определяется возможностью применения разработанной технологии прогнозирования сложных наблюдаемых временных рядов и алгоритмов оценивания целостности и цифровой рефлексивной фильтрации временных рядов не только для рассматриваемых в диссертации сложных объектов финансово-экономического типа, но и сложных объектов геофизической и другой природы, структура которых удовлетворительно описывается иерархически, а их наблюдаемая динамика представляется выраженными траекториями с наличием сингулярностей.

**Достоверность полученных результатов** (в части прогнозирования) определяется успешным сопоставлением прогнозных значений по отношению к наблюдаемым значениям, но не по критерию невязки, а согласно определенной целевой функции – обнаружение изменений в направлении движения.

**Использование результатов диссертации.** Материалы, положения и результаты диссертационной работы использованы в учебных программах дисциплин «Модели случайных процессов и сложных объектов» и «Методы и

средства экстраполяции и прогнозирования» кафедры «Системный анализ и управление» факультета информатики и процессов управления Политехнического института Сибирского федерального университета для преподавания студентам направления 221000.68 «Системный анализ данных и моделей принятия решений».

**Апробация результатов диссертации.** Основные положения и результаты диссертационных исследований докладывались на II всероссийской научно-технической конференции «Транспортные системы Сибири» (Красноярск, 2004), на IV и V всероссийской конференции по финансово-актуарной математике и смежным вопросам (Красноярск, 2005 и 2006), на VIII и IX всероссийском семинаре «Моделирование неравновесных систем» (Красноярск, 2005 и 2006), на VII всероссийской конференции молодых ученых по математическому моделированию и информационным технологиям (Красноярск, 2006), на X международной конференции «Решетневские чтения» (Красноярск, 2006), на семинарах «Системные задачи финансовой инженерии» Красноярского института экономики (Красноярск, 2006), а также на постоянно действующем семинаре кафедры «Системный анализ и управление» Политехнического института Сибирского федерального университета (Красноярск, 2005 – 2007).

**Публикации.** По теме диссертации опубликовано 15 научных работ, из которых: 2 статьи в периодических изданиях по списку ВАК, 6 статей в сборниках научных трудов, 8 работ в трудах всероссийских конференции, 1 работа в трудах международной конференции.

**Общая характеристика диссертации.** Диссертация состоит из 4 глав, содержит основной текст на 132 страницах, 63 иллюстрации (из них 20 находятся в приложениях), 9 таблиц, 2 приложения на 15 страницах, список использованных источников из 114 наименований.

## СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обоснована актуальность работы, сформулированы цели и задачи исследования, указаны методологические основания исследований, изложены основные положения, рассмотрены вопросы научной новизны, теоретической и практической значимости работы и изложены результаты, выносимые на защиту.

В **первой главе** («Сложные наблюдаемые временные ряды и характеристики их целостности») рассматривается системологическая модель сложных объектов и синтезируются алгоритмы оценивания целостности временных рядов, позволяющие, основываясь на числовом индикаторе, выносить суждения о внутренней целостности упорядоченных данных, представляющих временной ряд, и сравнивать различные временные ряды по критерию целостности.

В разделе 1.1 («Сложные объекты и сложные данные») проводится анализ основных свойств сложных объектов, являющихся продуцентами сложных наблюдаемых временных рядов.

К существенным свойствам сложных объектов относятся: наличие слабо формализуемых, ненаблюдаемых и неуправляемых компонент; многомерность; динамическая сложность; неравновесность; самоорганизация; глобальные масштабы функционирования (вплоть до планетарных масштабов).

Наличие слабо формализуемых компонент обусловлено социальным генезисом сложных объектов и вызваны организованной сложностью этих объектов, порождающей рефлексивность и наличие поведения, а также многофакторность и многоаспектность.

В этом же разделе определяется класс конкретных сложных объектов, исследуемых в рамках диссертационной работы, – сложные объекты финансово-экономического типа: «Международный валютный рынок» (свойства – котировки различных пар валют), «Международный фондовый рынок» (свойства – курсы акций различных компаний), «Мировая экономика» (свойства – различные индексы). Выбор сложных объектов финансовой и экономической среды для проведения исследований обосновывается их доступностью в среде Internet и наблюдаемостью с высокой разрешающей способностью (вплоть до минутных значений) на интервале жизни этих объектов.

Далее формулируется задача прогнозирования последовательностей наблюдаемых значений свойств сложных объектов – сложных наблюдаемых временных рядов, для которых рассматриваются основные характеристики с точки зрения визуального когнитивного анализа, статистического анализа и вейвлет-анализа на примере наблюдаемой последовательности котировки валют EURUSD (рис. 1).

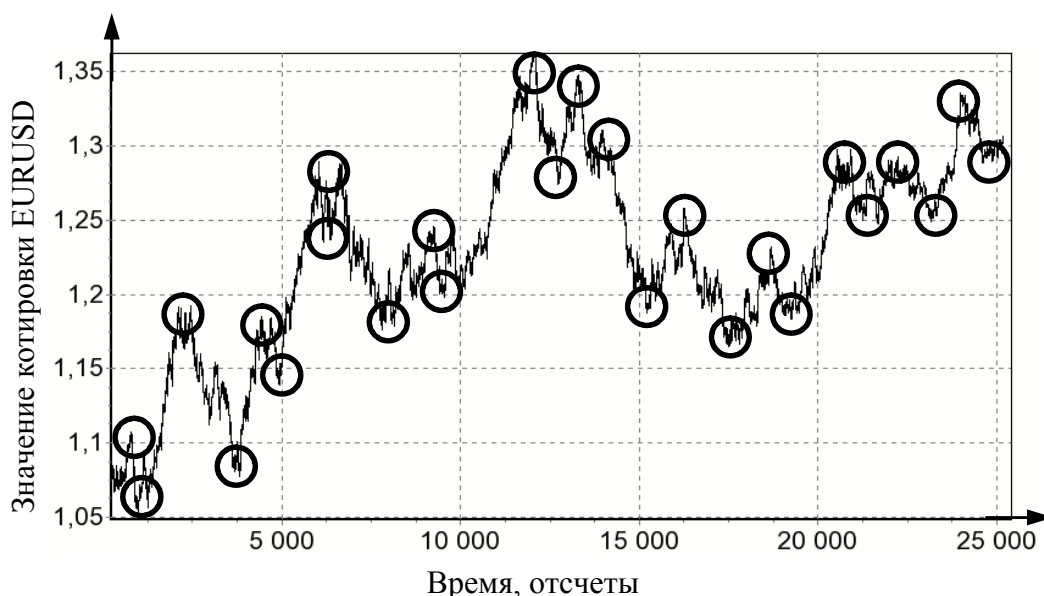


Рис. 1. Сложный наблюдаемый временной ряд на примере котировки EURUSD [данные предоставлены IDC Alpari], окружностями указаны целевые точки изменения направления движения

Значения сложных наблюдаемых временных рядов (рис. 1) структурированы по уровням, переход между которыми происходит в виде



сингулярностей, и имеют несколько мод в распределении значений. Распределение приращений такого временного ряда (рис. 1) симметрично относительно нуля и является унимодальным с «тяжелыми хвостами», эксцесс которого на порядки отличается от эксцесса нормального распределения.

Для того чтобы показать иерархическую структуру динамики траектории сложного наблюдаемого временного ряда (рис. 1), в диссертации осуществляется его вейвлет-преобразование, показывающее, что данный ряд представляется иерархической композицией составляющих, различающихся по силе воздействия и частоте изменчивости.

Далее указываются причины ограниченной применимости методов теории случайных процессов, сигнальной теории, теории мартингалов, регрессионной методологии для прогнозирования сложных наблюдаемых временных рядов (рис. 1) и в качестве решения предлагается построение технологии прогнозирования сложных наблюдаемых временных рядов, первый компонент которой (оценивание целостности) рассматривается далее в первой главе.

В разделе 1.2 («Проблема оценивания характеристик целостности сложных наблюдаемых временных рядов») рассматривается проблема оценивания характеристик целостности временных рядов, выражающейся в мере его «зазубренности» или степени гладкости, которая подчеркивается многими учеными (Климонтович, Кравцов, Мачарашвили, Петерс, Татарский, Челидзе) и решается, как правило, с использованием фрактальных или хаотических моделей либо с использованием метода нормированного размаха.

В разделе 1.3 («Анализ возможности использования методов оценивания «зазубренности» при исследовании наблюдаемых свойств сложных объектов») рассматриваются принципиальные и вычислительные трудности применение фрактальных размерностей, показателей Ляпунова и показателя Херста в качестве числовой меры «зазубренности» или индикатора степени гладкости временных рядов. С одной стороны, эти характеристики требуют грубых допущений о природе временного ряда. С другой стороны, их оценивание требует субъективной человеческой оценки «на глаз» и не может быть полностью автоматизировано. Вследствие указанных причин в диссертации предлагается новая методика оценивания целостности сложного наблюдаемого временного ряда.

Сложный наблюдаемый временной ряд рассматривается как абстрактный объект, для которого, как и для сложного объекта, прежде всего, важны целостные характеристики. Вероятно ожидать наличие у сложного наблюдаемого временного ряда множества целостных характеристик, соответствующим его разным целостным свойствам.

Ключевой идеей в понимании целостности временного ряда является идея соседства: чем сильнее зависимость между соседними значениями и чем шире диапазон зависимых значений, тем выше целостность временного ряда.

Для оценивания целостности временных рядов в разделе 1.4 («Синтез алгоритма оценивания целостности временных рядов в метрическом пространстве») предлагается рассматривать временной ряд как блуждание в

одномерном метрическом пространстве и изучать дисперсионные характеристики (средние пройденные расстояния) временного ряда на разных масштабах на основе различных локальных скользящих усреднений его значений.

Для исходного временного ряда  $\mathbf{X}_{(\Delta=1)} = \{x_1, \dots, x_n\}$  ( $\Delta$  – масштабный уровень) определим пройденное им среднее расстояние  $S_{(\Delta=1)}$  как

$$S_{(\Delta=1)} = \frac{1}{n-1} \sum_{i=1}^{n-1} |x_{i+1} - x_i|. \quad (1)$$

Далее из ряда  $\mathbf{X}_{(\Delta=1)}$  сформируем временной ряд  $\mathbf{X}_{(\Delta=2)}$ , каждый элемент которого равен среднему значению двух соседних значений ряда  $\mathbf{X}_{(\Delta=1)}$ :

$$\mathbf{X}_{(\Delta=2)} = \left\{ \frac{x_1 + x_2}{2}, \dots, \frac{x_{n-1} + x_n}{2} \right\}. \quad (2)$$

Для полученного ряда  $\mathbf{X}_{(\Delta=2)}$  также определим пройденное им среднее расстояние  $S_{(\Delta=2)}$  как

$$S_{(\Delta=2)} = \frac{1}{n-2} \sum_{i=1}^{n-2} \left| \frac{x_{i+1} + x_{i+2}}{2} - \frac{x_i + x_{i+1}}{2} \right| = \frac{1}{2(n-2)} \sum_{i=1}^{n-2} |x_{i+2} - x_i|. \quad (3)$$

Для старшего рассматриваемого уровня  $\Delta = k$  среднее расстояние  $S_{(\Delta=k)}$ , пройденное временным рядом  $\mathbf{X}_{(\Delta=k)}$  равно

$$S_{(\Delta=k)} = \frac{1}{n-k} \sum_{i=1}^{n-k} \left| \frac{x_{i+1} + \dots + x_{i+k}}{k} - \frac{x_i + x_{i+k-1}}{k} \right| = \frac{1}{k(n-k)} \sum_{i=1}^{n-k} |x_{i+k} - x_i|. \quad (4)$$

Из полученных средних расстояний  $S_{(\Delta=1)}, \dots, S_{(\Delta=k)}$  сформируем характеристику  $\Pi = \{\pi_1, \dots, \pi_k\}$  – спектр показателей целостности временного ряда, в котором каждое значение  $\pi_\Delta$  определяется как

$$\pi_\Delta = \frac{S_{(\Delta)}}{S_{(\Delta=1)}}. \quad (5)$$

Показатели целостности  $\pi_\Delta$  обладают свойством  $0 \leq \pi_\Delta \leq 1$ . Для выбора параметра  $k$ , характеризующего старший рассматриваемый масштаб, было установлено соотношение  $k \leq 10^{-3}n$ , исходя из статистической устойчивости оценивания показателей  $\pi_\Delta$ .

В ходе диссертационного исследования в разделе 1.5 было установлено, что для гауссовского блуждания (модели броуновского движения) показатели целостности хорошо аппроксимируются гиперболической зависимостью:

$$\pi_\Delta = \frac{1}{\Delta}. \quad (6)$$

На рис. 2 представлены спектры показателей целостности для различных временных рядов (рассматриваемые в разделах 1.5 и 1.6), демонстрирующие чувствительность предлагаемой методики оценивания целостности временных рядов к различной природе этих рядов (периодической, детерминированной, случайной).

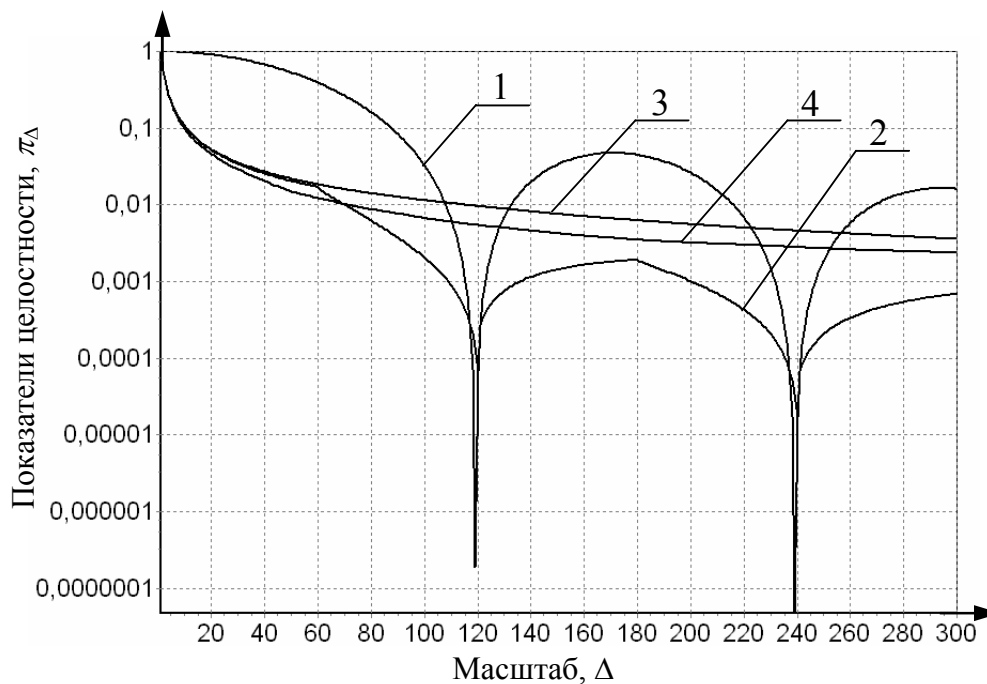


Рис. 2. Спектры целостности временных рядов: 1 – синусоида с периодом 120; 2 – последовательность прямоугольных импульсов с периодом 120; 3 – гауссовское блуждание; 4 – индекс Доу-Джонса DJI

Во **второй главе** («Конструирование цифровой рефлексивной фильтрации сложных наблюдаемых временных рядов») рассматривается второй компонент разрабатываемой технологии прогнозирования сложных наблюдаемых временных рядов – методика повышения целостности сложных наблюдаемых временных рядов, основанная на сглаживающих цифровых фильтрах.

В разделе 2.1 («Обоснование использования принципа рефлексии при решении задачи прогнозирования сложных наблюдаемых временных рядов») обосновывается применение принципа рефлексии в построении сглаживающей методики. Для этого анализируются работы таких ученых, как Бажанов, Воробьев, Лефевр, Разумовский, Сорос, и на основании того, что использование теории рефлексивности для описания динамики ценовых последовательностей конструктивно (наличие в ценовых последовательностях повторяющихся стереотипов поведения), предлагается внести принцип рефлексии на алгоритмическом уровне в процедуру цифровой фильтрации в виде рекуррентности фильтра.

В разделе 2.2 («Анализ цифровых фильтров») в качестве базового цифрового фильтра, способного учесть принцип рефлексии в рекуррентном виде и обладающего меньшим запаздыванием по сравнению с простыми скользящими средними, выбирается экспоненциальное скользящее среднее:

$$y_i = \mu x_i + (1 - \mu)y_{i-1}, \quad (7)$$

где  $\{x_i\}$  – значения исходного временного ряда;  
 $\{y_i\}$  – значения отфильтрованного временного ряда;  
 $\mu$  – степень фильтрации, удовлетворяющая условию  $0 \leq \mu \leq 1$ .

К разрабатываемым фильтрам кроме соответствия требованию рефлексии, предъявлялись следующие требования: управление запаздыванием/опережением отфильтрованного временного ряда относительно исходного, отсутствие принципиальных расхождений между исходным и фильтруемым временными рядами на концах, управление степенью фильтрации по иерархическому принципу – степень гладкости отфильтрованного временного ряда должна быть сильной в начале и слабой в конце (значения, уходящие в прошлое влияют на будущее сильнее в смысле основной тенденции, а значения, находящиеся близко к текущему моменту времени, влияют на будущее сильнее в смысле своей индивидуальности).

В разделе 2.3 («Базовая модель рефлексивной фильтрации временных рядов: прямой и обратный рефлексивные фильтры») на основании преобразования (7) вводится прямой рефлексивный фильтр **FRF** (forward reflexive filter):

$$\mathbf{Y}_{(F)} = \mathbf{FRF}(\mu_{(F)})[\mathbf{X}]: y_{(F),i} = \mu_{(F)}x_i + (1 - \mu_{(F)})y_{(F),i-1}, \text{ при } i = \overline{2, n}. \quad (8)$$

**FRF**-фильтр (8) вносит в отфильтрованный временной ряд  $\mathbf{Y}_{(F)}$  относительно исходного ряда  $\mathbf{X}$  запаздывание, величина которого определяется параметром  $\mu_{(F)}$ .

Далее на основе применения преобразования (7) к фильтруемому ряду  $\mathbf{X}$  с конца, вводится обратный рефлексивный фильтр **BRF** (back reflexive filter):

$$\mathbf{Y}_{(B)} = \mathbf{BRF}(\mu_{(B)})[\mathbf{X}]: y_{(B),i} = \mu_{(B)}x_i + (1 - \mu_{(B)})y_{(B),i+1}, \text{ при } i = \overline{(n-1), 1}. \quad (9)$$

**BRF**-фильтр (9) вносит в отфильтрованный временной ряд  $\mathbf{Y}_{(B)}$  относительно исходного ряда  $\mathbf{X}$  опережение, величина которого определяется параметром  $\mu_{(B)}$ .

В разделе 2.4 («Разработка двусторонней и многократной рефлексивной фильтрации временных рядов») на основе совместного использования фильтров **FRF** и **BRF** вводятся двусторонний (прямой→обратный) рефлексивный фильтр **FBRF** (forward→back reflexive filter):

$$\mathbf{Y}_{(FB)} = \mathbf{FBRF}(\mu_{(F)}, \mu_{(B)})[\mathbf{X}]: \mathbf{Y}_{(FB)} = \mathbf{BRF}(\mu_{(B)})[\mathbf{FRF}(\mu_{(F)})[\mathbf{X}]] \quad (10)$$

и двусторонний (обратный→прямой) рефлексивный фильтр **BFRF** (back→forward reflexive filter):

$$\mathbf{Y}_{(BF)} = \mathbf{BFRF}(\mu_{(B)}, \mu_{(F)})[\mathbf{X}]: \mathbf{Y}_{(BF)} = \mathbf{FRF}(\mu_{(F)})[\mathbf{BRF}(\mu_{(B)})[\mathbf{X}]] \quad (11)$$

В этом же разделе предложено многократное использование фильтров **FBRF** и **BFRF** ( $m$  раз) для придания отфильтрованному временному ряду гладких свойств (высокой целостности) без изменения основного темпа временного ряда при параметрах  $\mu_{(F)}$  и  $\mu_{(B)}$  близких к единице.

В разделе 2.5 («Неравномерная и иерархическая рефлексивная фильтрация временных рядов») предлагается выбирать в (7) параметр  $\mu$  переменным, на основе чего вводится иерархический рефлексивный фильтр, позволяющий за счет выбора параметра  $\mu$  как

$$\mu_i = \mu_b + (\mu_e - \mu_b) \left( \frac{i-1}{n-1} \right)^p \quad (12)$$

добиться результата фильтрации, когда значения отфильтрованного временного ряда тем сильнее сглажены, чем сильнее они «уходят» в прошлое.

На рис. 3 представлено применение всех предложенных в третьей главе цифровых рефлексивных фильтров к котировочному ряду EURUSD в виде единого обобщенного рефлексивного фильтра **RF**.

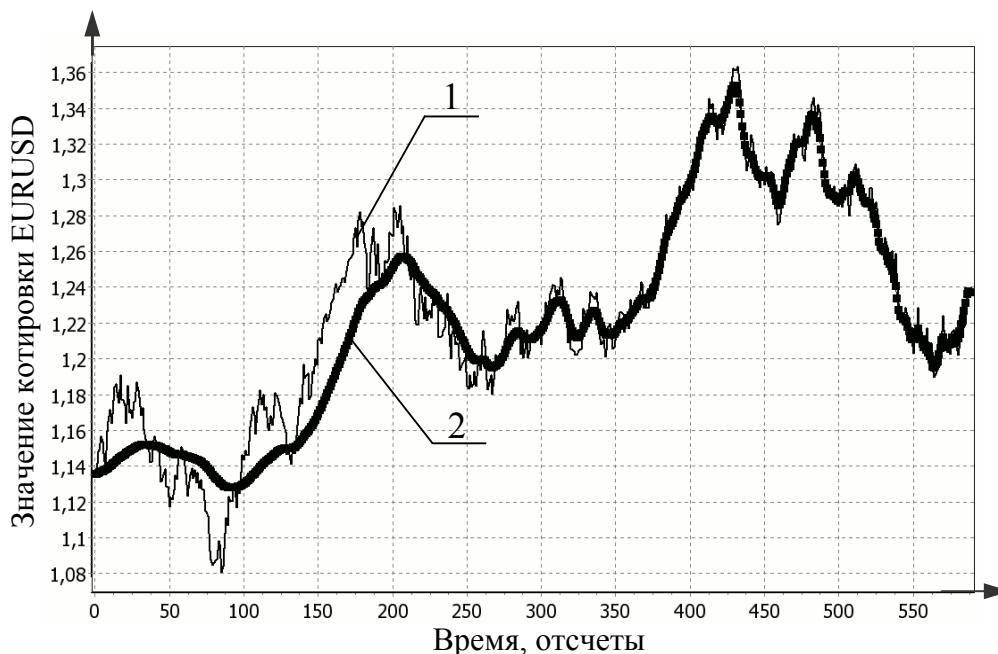


Рис. 3. Обобщенный вариант рефлексивной фильтрации, включающей в себя двустороннюю, многократную и иерархическую фильтрацию:  
1 – исходная котировка; 2 – отфильтрованная котировка

**Третья глава** («Синтез коллективов линейных рекуррентных моделей прогнозирования сложных наблюдаемых временных рядов») посвящена разработке компонента технологии, непосредственно предназначенного для прогнозирования. Разрабатываемые алгоритмы для прогнозирования временных рядов основываются на моделях авторегрессионного типа.

В разделе 3.1 («Задача прогнозирования состояния сложных объектов в контексте общих представлений о будущем») формулируются общие положения, исходя из которых для сложных объектов может быть построена процедура прогнозирования для формирования представлений о будущих состояниях таких объектов.

В разделе 3.2 («Анализ методов прогнозирования временных рядов – математические модели») рассматриваются основные модели (линейные и нелинейные, детерминированные и стохастические) и методики (сингулярно-спектральный анализ, кусочно-линейное моделирование, нейроинформатика, метод группового учета аргументов) прогнозирования временных рядов, среди которых для технологии выбираются модели авторегрессионного типа.

В качестве базовой модели для построения прогнозных значений в диссертации выбрана линейная рекуррентная формула (рассматриваемая далее как частная модель прогноза)

$$\tilde{x}_p = \sum_{j=1}^d a_j x_{p-j}, \quad (13)$$

характеристики которой, а также алгоритмы оценивания коэффициентов  $\{a_j\}$  рассмотрены в разделе 3.3. Там же описывается эксперимент по прогнозируемости тестового временного ряда частной моделью прогноза (13) при различных соотношениях сигнал/шум для определения уровня целостности, при котором временной ряд в принципе может прогнозироваться моделью вида (13).

В разделе 3.4 («Синтез коллективов частных моделей прогноза») модель (13) обобщается на случай коллектива таких моделей. Это необходимо для того чтобы довести модель (13) до технологического исполнения, наделив создаваемый коллектив регламентом. Кроме того, в силу высокой сложности исследуемых объектов и порождаемых ими наблюдаемых последовательностей частные модели прогноза могут «работать» не эффективно.

Построение коллективов частных моделей прогноза основано на выборе различных длин обучающих последовательностей для определения коэффициентов  $\{a_j\}$  в (13), а также на выборе различных порядков авторегрессии  $d$  в (13).

Для выбора «лучшей» частной модели прогноза из коллектива в пространстве параметров  $\mathbf{L} \times \mathbf{d}$  в диссертации использованы стандартные критерии оценивания качества: среднеквадратическая ошибка аппроксимации

$$\Sigma_A^2(L, d) = \sum_{p=n-L-v}^{n-v-1} \left( x_p - \sum_{j=1}^d a_j x_{p-j} \right)^2 \quad (14)$$

и среднеквадратическая ошибка прогнозирования

$$\Sigma_E^2(L, d) = \sum_{p=n-v}^n \left( x_p - \sum_{j=1}^d a_j x_{p-j} \right)^2. \quad (15)$$

В диссертации дополнительно к критериям  $\Sigma_A^2$  и  $\Sigma_E^2$  были предложены новые критерии оценивания качества частных моделей прогноза: точность решения СЛАУ

$$\Sigma_\varepsilon(L, d) = \sum_{k=1}^d \left| \sum_{s=1}^d a_s \sum_{p=n-L-v}^{n-v-1} x_{p-k} x_{p-s} - \sum_{p=n-L-v}^{n-v-1} x_{p-k} x_p \right|, \quad (16)$$

исходя из которой оцениваются коэффициенты  $\{a_j\}$  модели (13), баланс коэффициентов  $\{a_j\}$  в модели (13)

$$\mathbf{B}_{A_1}(L, d) = \sum_{j=1}^d a_j \quad \text{и} \quad \mathbf{B}_{A_2}(L, d) = \sum_{j=1}^d a_j^2 \quad (17)$$

и среднеквадратическая ошибка обратного прогнозирования

$$\Sigma_{E, \text{back}}^2(L, d) = \sum_{p=n-L-v-\gamma}^{n-L-v-1} \left( x_p - \sum_{j=1}^d a_j x_{p-j} \right)^2. \quad (18)$$

При использовании приведенных выше критериев, обучающая последовательность разбивалась на участки, показанные на рис. 4.

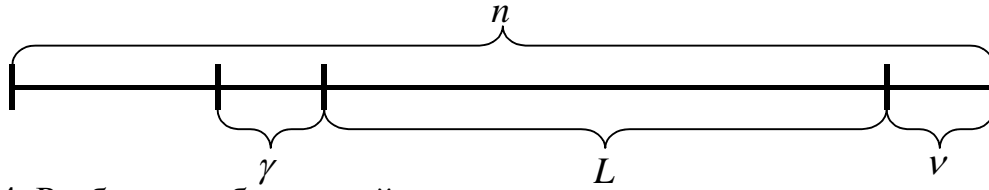


Рис. 4. Разбиение обучающей последовательности на отдельные участки

Регламент выбора «лучшей» частной модели прогноза из всего коллектива дедуктивный. Прежде всего, из коллектива исключаются модели по критерию  $\Sigma_{\varepsilon}$ . Затем исключаются модели, плохо аппроксимирующие обучающую последовательность, по критерию  $\Sigma_A^2$ . После этого исключаются модели по критерию  $\Sigma_E^2$ . Затем исключаются модели, не удовлетворяющие критериям баланса  $\mathbf{B}_{A_1}$  и  $\mathbf{B}_{A_2}$ . Из оставшихся моделей «лучшей» считается та, у которой лучшее значение критерия  $\Sigma_{E,back}^2$ .

Для того, чтобы учесть иерархическую структуру динамики сложных наблюдаемых временных рядов, в разделе 3.5 («Разработка иерархического обобщения частной модели прогноза») предлагается методика, позволяющая построить иерархическое прогнозирование, основанное на формировании временных рядов разных масштабов по времени.

Принцип формирования временных рядов разных масштабов по времени, используемый в диссертационном исследовании, представлен на рис. 5 и заключается в регулярном исключении  $(\Delta - 1)$  элементов из исходного временного ряда для масштабного уровня  $\Delta$ .

Далее временной ряд каждого масштабного уровня прогнозируется с использованием предложенных выше коллективов частных моделей прогноза, после чего из полученных прогнозов синтезируется результирующий прогноз с весами, пропорциональными масштабным уровням  $\Delta$  либо как простое среднее.

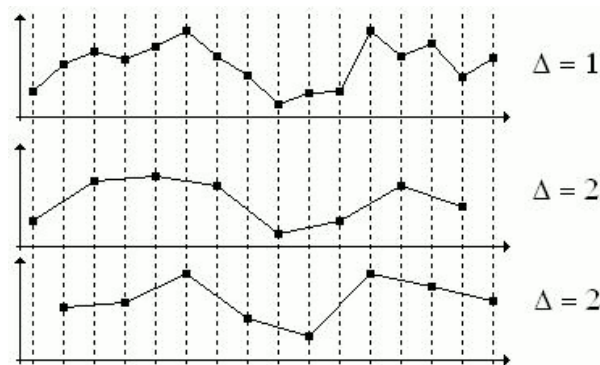


Рис. 5. Принцип формирования временных рядов различных масштабов по времени (схематично)

В четвертой главе из разработанных в первой, второй и третьей главах компонент синтезируется единая информационная технология прогнозирования сложных наблюдаемых временных рядов, структура которой представлена на рис. 6-9.

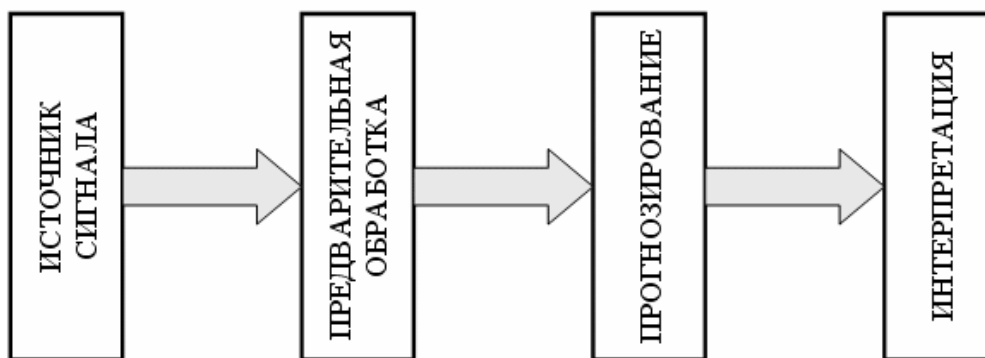


Рис. 6. Структура технологии прогнозирования сложных наблюдаемых временных рядов (концептуально)

Предварительная обработка сложного наблюдаемого временного ряда (рис. 7) заключается в повышении целостности исходного ряда на основе рефлексивной фильтрации и применения оператора  $\mathbf{RF}_X[\bullet]$ .

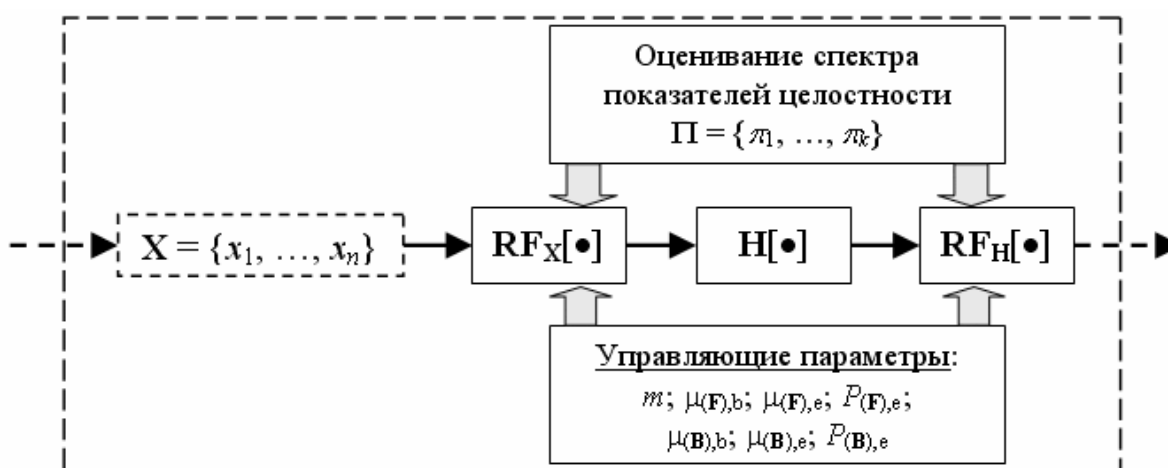


Рис. 7. Предварительная обработка сложных наблюдаемых временных рядов

Затем от сглаженной траектории осуществляется переход замещающему временному ряду, стационарному в смысле среднего и дисперсии, на основе применения оператора  $\mathbf{H}[\bullet]$ , в качестве которого в диссертации были выбраны простые приращения. Далее построенный заместитель дополнительно сглаживается оператором  $\mathbf{RF}_H[\bullet]$ .

Следующим этапом является непосредственное построение прогнозных значений (рис. 8). Для этого из сглаженного временного ряда-заместителя  $\{\chi_1, \dots, \chi_n\}$  формируются временные ряды различных масштабов по времени, для каждого из которых строится коллектив частных моделей прогноза и получаются прогнозные значения каждого иерархического уровня. Затем из полученных прогнозов для разных уровней с определенной весовой функцией синтезируется искомый прогноз.

Заключительным этапом технологии является интерпретация полученных прогнозных значений сглаженного заместителя относительно исходного сложного наблюдаемого временного ряда (рис. 9).



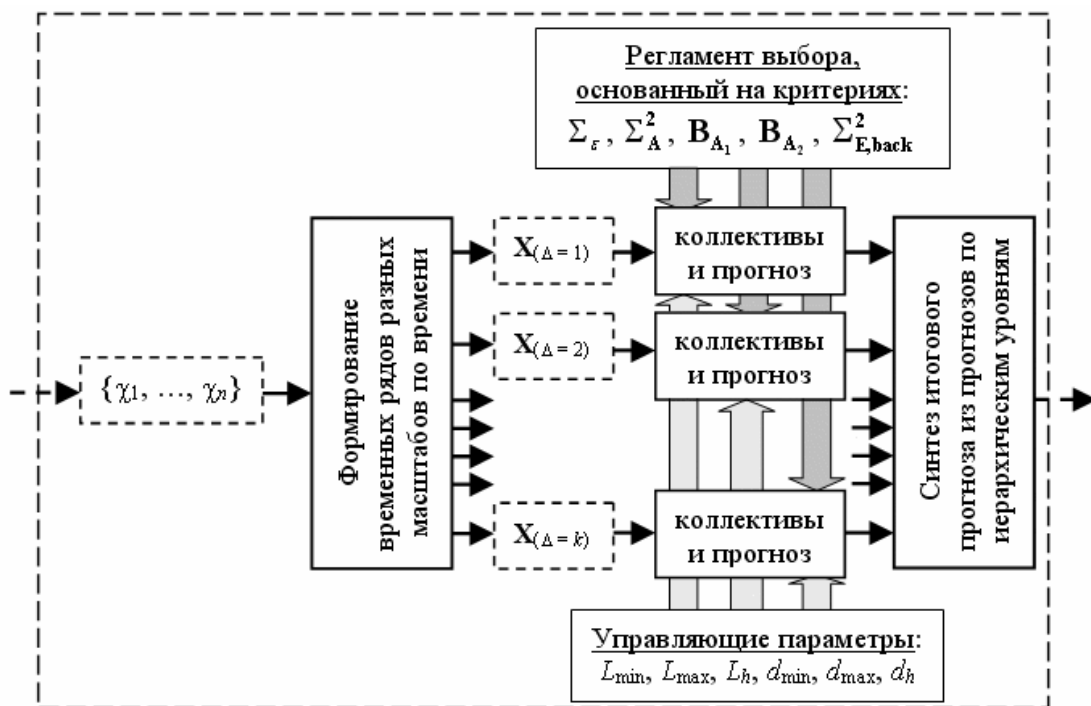


Рис. 8. Прогнозирование обработанных сложных наблюдаемых временных рядов

Интерпретация основана на применении обратных рефлексивных фильтров  $\mathbf{RF}_H^{-1}[\bullet]$  и  $\mathbf{RF}_X^{-1}[\bullet]$ , а также оператора восстановления  $\mathbf{H}^{-1}[\bullet]$ , позволяющего вернуться от временного ряда-заместителя (приращений) к траектории сложного наблюдаемого временного ряда.

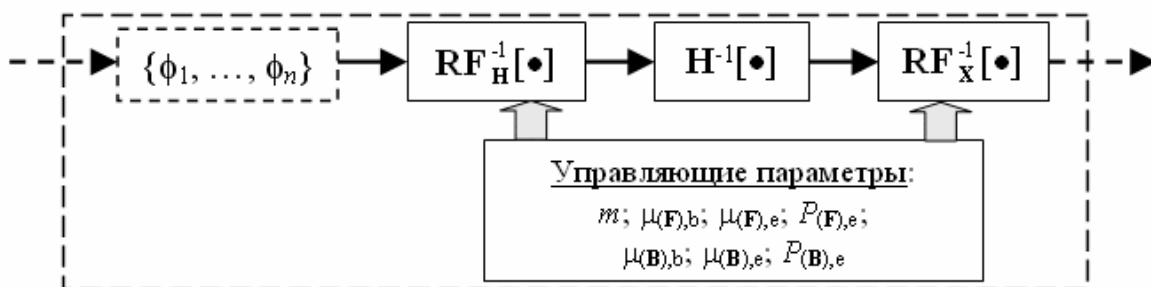


Рис. 9. Интерпретация прогнозных значений обработанного сложного наблюдаемого временного ряда

На основе разработанной информационной технологии прогнозирования сложных наблюдаемых временных рядов (рис. 6-9) в среде «Borland Delphi 7» был создан программный инструмент «**ForIT**», с использованием которого в диссертации были проведены эксперименты с натурными данными выбранных сложных объектов финансово-экономического типа. Полученные результаты (рис. 10) подтверждают состоятельность разработанной технологии и предположения о сложных объектах, связанные с иерархической структурой, наличием рефлексии и возможности перехода к ряду-заместителю.

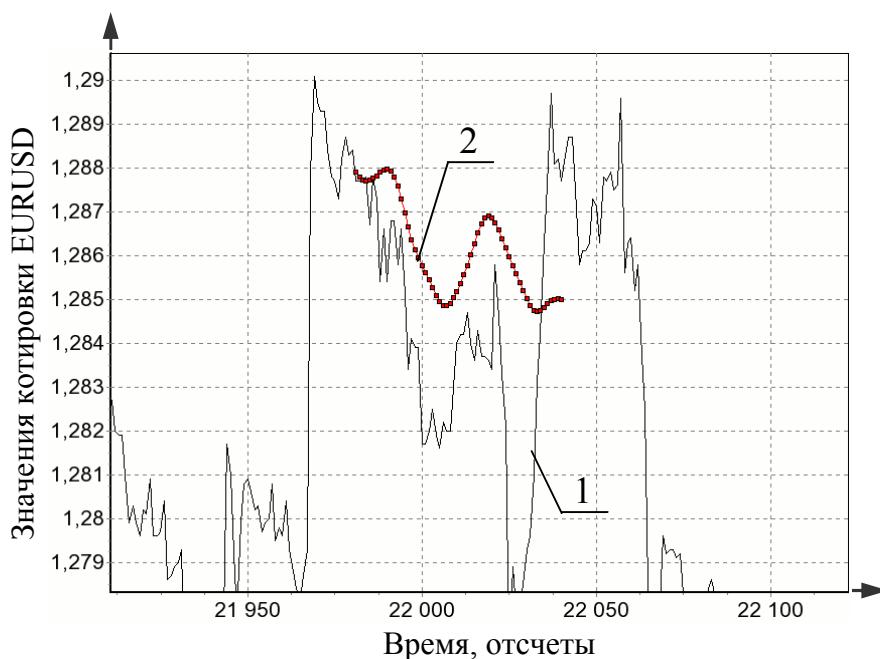


Рис. 10. Прогнозирование котировки EURUSD (увеличенный фрагмент) на 60 точек вперед по 2 иерархическим уровням при отступе назад 10 точек:  
1 – исходный ряд EURUSD; 2 – прогнозные значения

В **заключении** приведены основные результаты, полученные в ходе выполнения диссертационного исследования, сформулированы выводы по работе в целом и определены перспективы дальнейших исследований.

В диссертационной работе решена задача прогнозирования наблюдаемых свойств сложных объектов в контексте определенной целевой функции прогнозирования, определенной как обнаружение изменений в направлении движения траектории, представляющей наблюдаемые свойства.

Полученным решением задачи прогнозирования сложных наблюдаемых временных рядов, с одной стороны, является обобщенная информационная технология, основанная на предложенных в диссертации методики оценивания целостности временных рядов, рефлексивной фильтрации временных рядов. А также методики построения иерархических коллективов линейных рекуррентных моделей прогнозирования временных рядов.

С другой стороны, полученным решением является реализация синтезированной информационной технологии прогнозирования сложных наблюдаемых временных рядов в виде программного инструмента «**ForIT**».

Оба решения являются завершенными и готовыми к использованию.

В основе полученных в диссертации решений лежат следующие результаты, выносимы на защиту:

1. алгоритмы оценивания целостности временных рядов – спектр показателей целостности;
2. алгоритмы цифровой фильтрации временных рядов, учитывающие принцип рефлексии;
3. методика иерархического прогнозирования временных рядов, основанная на формировании временных рядов разных масштабов по времени;

4. информационная технология прогнозирования сложных наблюдаемых временных рядов, включающая все полученные результаты.

Задачи диссертационного исследования, поставленные во введении, решены успешно и цель диссертационной работы достигнута.

Синтезированная информационная технология прогнозирования наблюдаемых свойств сложных объектов может быть применима не только к рассматриваемым в диссертации объектам финансово-экономического типа, а к более широкому классу сложных объектов, структура которых хорошо описывается иерархическими моделями и которые доступны в виде наблюдаемых траекторий с наличием определенных стереотипов поведения.

#### **Публикации автора по теме диссертации**

1. Павлов, С. В. Иерархические коллективы линейных моделей прогноза сложных наблюдаемых временных рядов / С. В. Павлов // Системы управления и информационные технологии. – 2007. – № 1 (27). – С. 18-22.
2. Павлов, С. В. Иерархические линейные регрессионные модели для прогнозирования динамики аграрного сектора / С. В. Павлов, В. В. Сафронов // Вестник Красноярского государственного аграрного университета. – 2006. – № 15. – С. 177-183.
3. Павлов, С. В. Рефлексивная фильтрация временных рядов / С. В. Павлов // Вестник университетского комплекса: сб. науч. тр. / Красноярск: ВСФ РГУИТП, НИИ СУВПТ, 2005. – Вып. 5(19). – С. 56-64.
4. Павлов, С. В. Эпистемологические Хааровские модели наблюдений свойств сложных объектов / С. В. Павлов, В. Ф. Слюсарчук // Вестник университетского комплекса: сб. науч. тр. / Красноярск: ВСФ РГУИТП, НИИ СУВПТ, 2005. – Вып. 5 (19). – С. 65-80.
5. Павлов, С. В. Коллективы линейных прогнозных моделей финансовых временных рядов / С. В. Павлов // Управление и экономика. Теория и практика: сб. науч. тр. / Красноярск: ООО «Издательский центр «Платина», 2006. – Вып. 2. – С. 286-292.
6. Павлов, С. В. Вейвлет-представления финансовых процессов / С. В. Павлов, А. И. Максимов // Управление и экономика. Теория и практика: сб. науч. тр. / Красноярск: ООО «Издательский центр «Платина», 2006. – Вып. 2. – С. 302-305.
7. Максимов, А. И. Использование вейвлет-представлений для выделения иерархической структуры наблюдаемой динамики сложных неравновесных объектов / А. И. Максимов, **С. В. Павлов** // Моделирование неравновесных систем. Материалы IX Всероссийского семинара, 13-15 октября 2006 г. / Красноярск: ИВМ СО РАН, 2006. – С. 117-118.
8. Павлов, С. В. Выделение тенденций в нестационарном потоке данных / С. В. Павлов // Транспортные системы Сибири. Материалы II Всероссийской научно-технической конференции, 25-26 ноября 2004 г. / Красноярск: ИПЦ КГТУ, 2004. – С. 135-137.
9. Павлов, С. В. Оценивание спектра траекторной целостности наблюдений свойств сложных объектов / С. В. Павлов // Моделирование неравновесных

- систем. Материалы VIII Всероссийского семинара, 14-16 октября 2005 г. / Красноярск: ИВМ СО РАН, 2005. – С. 133-134.
10. Павлов, С. В. Целостные представления сложных объектов в фазовых пространствах / С. В. Павлов, В. В. Сафронов, В. Ф. Слюсарчук // V Всероссийская ФАМ конференция: тезисы докладов, 3-5 марта 2006 г. / Красноярск: Красноярский гос. ун-т, 2006. – С. 85-86.
  11. Павлов, С. В. Иерархическое прогнозирование наблюдаемых свойств сложных неравновесных объектов и систем / С. В. Павлов, В. Ф. Слюсарчук // Моделирование неравновесных систем. Материалы IX Всероссийского семинара, 13-15 октября 2006 г. / Красноярск: ИВМ СО РАН, 2006. – С. 133-134.
  12. Павлов, С. В. Синтез алгоритмов цифровой рефлексивной фильтрации временных рядов / С. В. Павлов // VII Всероссийская конференция молодых ученых по математическому моделированию и информационным технологиям (с участием иностранных ученых). Программа выступлений и тезисы докладов, 13 ноября 2006 г. / Новосибирск: ИВТ СО РАН, Красноярск: ИВМ СО РАН, 2006. – С. 24.
  13. Павлов, С. В. Синтез коллективов линейных рекуррентных моделей прогнозирования временных рядов / С. В. Павлов // VII Всероссийская конференция молодых ученых по математическому моделированию и информационным технологиям (с участием иностранных ученых). Программа выступлений и тезисы докладов, 13 ноября 2006 г. / Новосибирск: ИВТ СО РАН, Красноярск: ИВМ СО РАН, 2006. – С. 24-25.
  14. Павлов, С. В. Алгоритмы оценивания целостности наблюдаемых временных рядов / С. В. Павлов // Решетневские чтения: материалы X международной научной конференции, посвященной памяти генерального конструктора ракетно-космических систем академика М.Ф. Решетнева, 8-10 ноября 2006 г. / Красноярск: Сиб. гос. аэрокосмич. ун-т., 2006. – С. 256-257.
  15. Сафронов, В. В. Алгоритм оценивания целостности наблюдаемых траекторий динамики сложных многомерных объектов / В. В. Сафронов, **С. В. Павлов** // VII Всероссийская конференция молодых ученых по математическому моделированию и информационным технологиям (с участием иностранных ученых). Программа выступлений и тезисы докладов, 13 ноября 2006 г. / Новосибирск: ИВТ СО РАН, Красноярск: ИВМ СО РАН, 2006. – С. 27-28.

Павлов Сергей Владимирович

**Технология прогнозирования сложных наблюдаемых временных рядов**

Автореф. дисс. на соискание ученой степени кандидата технических наук

Подписано в печать 3.04.2007. Заказ № 633  
Формат 60×90/16. Усл. печ. л. 1. Тираж 100 экз.  
Типография Политехнического института  
Сибирского федерального университета