

**Список литературы:** 1. Сакович В. П., Колотвинов В. С., Шамов А. Ю. Хирургическое лечение интракраниальных аневризм из птерионального доступа с применением малых трепанационных отверстий // Вопросы нейрохирургии. 2000. №1. С. 3-7. 2. Олешкевич Ф.В., Олешкевич А.Ф. Нейрохирургия. Операции на головном мозге. Минск: Высшая школа, 1993. 294 с. 3. Хирургия опухолей основания черепа / Под ред. А.Н. Коновалова. М.: ОАО Можайский полигр. комбинат, 2004. 372 с. 4. Федоров В.Д., Кармазановский Г.Г., Цвиркун В.В. и др. Новые возможности спиральной компьютерной томографии – виртуальная хирургия // Медицинская визуализация. 2000. № 2. С. 15-18. 5. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2006. 1072 с. 6. Путятин Е.П., Аверин С.И. Обработка изображений в робототехнике. М.: Машиностроение, 1990. 330 с. 7. Осипов Л.В. Ультразвуковые диагностические приборы. М.: Видар, 1999. 256 с. 8. Gonzalez-Reimers E., Mas-Pascual M., Arnay-de-la-Rosa M. Noninvasive estimation of bone mass in ancient vertebrae // Am. J. Phys. Anthropol. 2004. Т. 125, № 2. Р. 121-131. 9. Справочник по конструкционным материалам / Под ред. Арзамасова Б.Н., Соловьевой Т.В. М.: МГТУ им. Н.Э. Баумана, 2005. 637 с. 10. Манжос Ф.М. Деревообрабатывающие станки (конструкция, эксплуатация и технологические расчеты). М.: Гослемуиздат, 1963. 674 с. 11. Станочный дереворежущий инструмент. Каталог-справочник, М. НИИмаш, 1976. 416с.

*Поступила в редколлегию 12.03.2007*

**Аврунин Олег Григорьевич**, канд. техн. наук, доцент кафедры биомедицинских электронных устройств и систем ХНУРЭ. Научные интересы: медицинское приборостроение, разработка новых методов и аппаратных средств. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 702-13-64, e-mail: bykh@kture.kharkov.ua.

**Носова Татьяна Витальевна**, канд. техн. наук, младший научный сотрудник кафедры биомедицинских электронных устройств и систем ХНУРЭ. Научные интересы: медицинское приборостроение, разработка новых методов и аппаратных средств. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 702-13-64, e-mail: bykh@kture.kharkov.ua.

---

УДК 519.246.8

*А.Н. ОДЕЙЧУК, Б.В. ШАМША, Е.Г. ФЕДОРОВ*

## **ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ВЫБОРА МЕТОДА ПРОГНОЗИРОВАНИЯ СТОХАСТИЧЕСКИХ РЯДОВ В УСЛОВИЯХ ГЕТЕРОСКЕДАСТИЧНОСТИ**

---

Описывается проблема выбора оптимального метода прогнозирования стохастических рядов в условиях гетероскедастичности. Для этого разрабатывается интеллектуальная система, приводится ее обобщенная структура. Основное внимание уделяется проблеме выбора метода оценивания моделей прогнозирования в условиях гетероскедастичности. Вычисляются параметры, влияющие на качество методов оценивания. Для определения оптимального метода оценивания впервые предлагается использовать дерево решений. Построенное дерево решений является основой интеллектуальной системы. В качестве апробации полученных результатов строится модель прогноза официального курса гривны по отношению к доллару США.

### **1. Постановка проблемы**

Прогнозирование финансовых временных рядов – необходимый элемент любой инвестиционной деятельности. Решение задач прогнозирования позволяет банкам, пенсионным и инвестиционным фондам эффективно осуществлять финансовые операции, рассчитывать эффективность принимаемых решений для повышения экономического потенциала. Таким образом, решение задач, связанных с прогнозированием финансовых показателей современных организаций с использованием современных информационных технологий, является актуальным.

*Цель данного исследования* состоит в разработке интеллектуальной системы выбора метода прогнозирования стохастических временных рядов в условиях гетероскедастичности.

Для этого необходимо решить следующие задачи: разработать схему интеллектуальной системы, определить показатели, влияющие на выбор метода прогнозирования, а также предложить механизм классификации временных рядов по методам их обработки.

## 2. Анализ последних исследований и публикаций, посвященных решению задач прогнозирования значений финансовых показателей

Финансовые временные ряды характеризуются неопределенностью и изменчивостью статистических характеристик с течением времени. Неопределенность приводит к увеличению доверительного интервала прогноза [1]. Для таких временных рядов характерна гетероскедастичность (в финансовой сфере распространены термины «волатильность», «кластеризация волатильности»). Гетероскедастичность характеризуется изменением дисперсии во времени в группах данных. Под волатильностью (volatility – изменчивость, непостоянство) понимают степень изменчивости во времени дисперсии временного ряда. Мерой волатильности служит дисперсия или среднее квадратическое отклонение. Для волатильности характерна так называемая “кластеризация”, т.е. в некоторые периоды волатильность принимает большие значения, а в другие – небольшие. Например, при рассмотрении курса валют, анализе инфляции и цен доходностей спекулятивных активов можно выделить периоды, когда колебания временного ряда были незначительными, и периоды, когда, среагировав на определённые события, ряд совершал значительные колебания.

Традиционные модели временных рядов, такие как модель ARMA, не могут адекватно учесть все характеристики, которыми обладают финансовые временные ряды, и требуют расширения.

Для эффективного принятия решений в условиях неопределенности необходимо повысить качество моделей прогнозирования с учетом изменения дисперсий и ковариаций во времени. Такие требования были учтены в моделях с условной авторегрессионной гетероскедастичностью (AutoRegressive Conditional Heteroskedasticity, ARCH), разработанных Энглем (Engle (1982)) [2,3].

Для получения оценок параметров моделей ARCH, GARCH, GARCH-M, EGARCH, HYGARCH и их модификаций может быть использован ряд методов: метод наименьших квадратов (МНК), метод взвешенных наименьших квадратов (ВМНК), обобщенный метод наименьших квадратов (ОМНК), метод максимального правдоподобия (ММП), метод моментов (ММ) [4,5].

## 3. Выделение нерешенных ранее частей общей проблемы в данной области

Основной проблемой, затрудняющей использование рассмотренной выше модели ARCH и ее модификаций в условиях гетероскедастичности, является то, что используемые методы для определения оценок параметров данной модели зависят от выполнения предположений и предположений, лежащих в основе данных методов. В каждом конкретном случае точность прогнозирования зависит от корректного применения одного из методов. Выбор метода и обоснование его применения является достаточно сложным и наукоемким процессом и требует от лица, принимающего решения, глубоких знаний в данной предметной области и методах прогнозирования.

С учетом развития информационных технологий и вычислительной техники актуальной является разработка интеллектуальной системы, позволяющей на основе статистических характеристик выборки определить наиболее эффективный метод ее обработки. Построение данной системы является достаточно сложной и наукоемкой задачей.

## 4. Постановка задачи исследования

Содержательную постановку задачи исследования сформулируем как разработку такой формализованной процедуры  $F$  выбора методов  $M$ , алгоритмов  $A$ , программ  $P$ , решения  $K$

задач, которая приводит к выбору системы  $S = \bigcup_{k=1}^K S_k$ , состоящей из подсистем  $S_k, k = \overline{1, K}$ .

Выбор каждой  $S_k$  сводится к выбору такого подмножества  $M^k, A^k, P^k$  решения задачи, которое удовлетворяло бы заданному критерию качества и статистическим характеристикам объекта. Представим синтез  $S_k$  следующим образом:

$$\{M, A, P, H, E\} \xrightarrow{F} S_k \left\{ M^k \in M; A^k \in A; P^k \in P \right\}$$
$$E^* = \text{extr}_{M^i \in M} \text{extr}_{A^{ij} \in A} \text{extr}_{P^{ijl} \in P} E(M^i, A^{ij}, P^{ijl}, H),$$

где  $M$  – класс методов;  $A$  – банк алгоритмов;  $P$  – библиотека программ решения отдельных задач;  $H$  – характеристики задачи;  $E$  – показатели качества подсистемы  $S_k$ .

В связи с этим выбор эффективного программного модуля может быть представлен как последовательный выбор метода, алгоритма и программы. Тогда для проектирования системы  $S_k$  необходимо решить следующие задачи: сформировать перечень  $K$  задач, решаемых в системе  $S$ , т.е. поставить задачу проектирования подсистем  $S_k$ ; сформировать класс методов  $M$ , банк алгоритмов  $A$ , библиотеку программ  $P$  решения каждой  $k$ -й задачи; выявить характеристики и показатели  $E$  качества отдельных модулей, по которым можно производить сравнение между собой альтернативных модулей; выявить специфические свойства объекта  $H$ , которые влияют на качество решения задач в условиях функционирования системы; установить математическую зависимость между характеристиками качества модулей  $E$  и свойствами объекта  $H$ ; разработать формализованную процедуру  $F$  автоматизированного проектирования прикладного программного обеспечения, т.е. разработать критерии и алгоритмы выбора наилучшего модуля.

В настоящей работе проводится анализ моделей для построения моделей прогноза и методов, позволяющих определить коэффициенты моделей в условиях гетероскедастичности, строится дерево решений для их классификации, которое может стать основой для разработки интеллектуальной системы выбора метода обработки данных в условиях гетероскедастичности.

## 5. Построение интеллектуальной системы

### 5.1. Обобщенное описание интеллектуальной системы

На основании анализа методов [4,5], позволяющих определить коэффициенты модели изучаемой выборки в условиях гетероскедастичности и подходов, предлагаемых в [4], представляем разрабатываемую интеллектуальную систему в виде следующей обобщенной схемы (рис. 1).

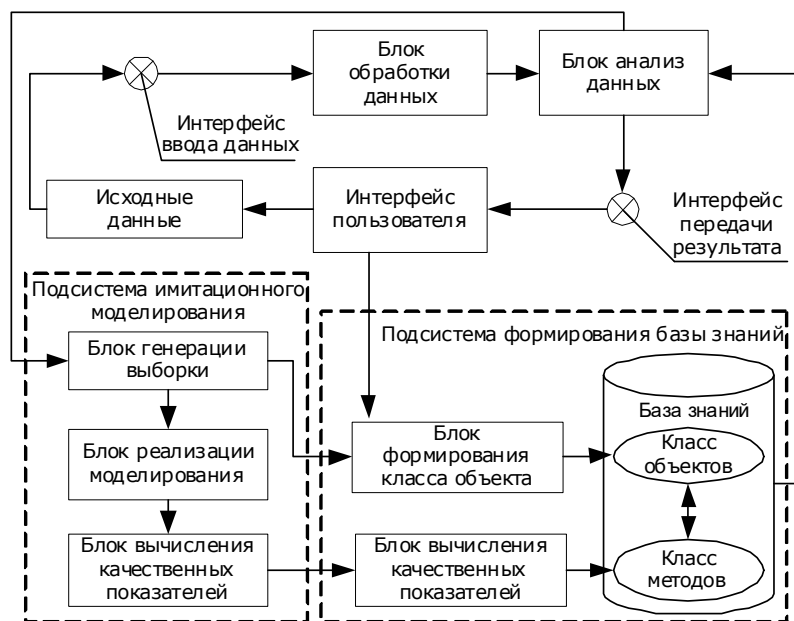


Рис. 1. Обобщенная структура интеллектуальной системы

Алгоритм работы системы представляется следующим образом.

1. Пользователь, используя интерфейс, подает на вход системы через интерфейс ввода данных исходные анализируемые данные, для которых необходимо определить метод их обработки.

2. В блоке обработки данных производится определение статистических характеристик, по которым выбирается метод. После этого они передаются в блок анализа данных.

Блок анализа данных соотносит статистические характеристики с определенным методом. Для этого используется аналитический аппарат, на основе которого принимается решение, а также данные, хранимые в базе знаний, касающиеся методов их обработки. После проведенного анализа, выполняемого данным блоком, получаем результат работы системы – наиболее эффективный метод обработки данной выборки.

Интерфейс пользователя содержит все необходимые для работы компоненты. Он позволяет производить ввод данных в систему, сохранять и загружать их из файла, выводит определяемые основные статистические характеристики выборки, а также предоставляет в удобном для пользователя виде модель прогноза в соответствии с решением, полученным на основе выбранного метода.

В блоке анализа данных решается классификационная задача. Для этой цели могут применяться нейронные сети, карты Кохонена, деревья решений, дискриминантный анализ, кластерный анализ, непараметрические статистики, нелинейное оценивание [4]. Применение одновременно нескольких из приведенных методов классификации позволит многократно повысить надежность получаемых решений, однако при этом возрастает время, затрачиваемое на проведение анализа.

В тех случаях, когда на выходе блока анализа данных не получен наиболее подходящий для этой выборки метод, с использованием подсистемы имитационного моделирования проводится ряд экспериментов, результатом которых будут выявлены статистические показатели, соотносящиеся с методами обработки данных, которые нужно проанализировать дополнительно для получения необходимого результата.

В данной работе впервые предложено выполнить блок анализа данных на основе дерева решений. Это обусловлено тем, что выборку могут описывать не только количественные показатели, но и качественные.

Дерево решений – это способ представления правил в иерархической, последовательной структуре, где каждому объекту соответствует единственный узел, дающий решение. Дерево решений позволяет использовать для ветвления как непрерывные, так и категориальные переменные, что значительно расширяет возможности проведения анализа данных.

## **5.2. Выделение параметров, на основе которых будет осуществляться выбор метода оценивания и построение имитационной модели**

Реализация процедуры автоматизированного выбора методов прогнозирования предполагает возможность их сравнительного анализа по некоторым критериям. Обязательным условием выбора показателя для сравнения методов является возможность количественной оценки критерия.

Для сравнения методов будем использовать следующие параметры [4]:  $h_1 = n$  – объем выборки;  $h_2 = p$  – количество независимых переменных;  $h_3 = \lambda_{\min}/\lambda_{\max}$  – показатель обусловленности матрицы плана (матрицы регрессоров);  $h_4 = \sigma_{\epsilon y}^2$  – дисперсия ошибки в зависимой переменной;  $h_5 = b^T b$  – длина вектора истинных коэффициентов в модели;  $h_6 = \sigma_{\epsilon y}^2 / \sigma_y^2$  – соотношение дисперсий шум/сигнал для зависимой переменной;  $h_7 = \eta$  – уровень засоренности распределения, если последнее является смесью  $f(\epsilon) = (1 - \eta)f_1(\epsilon) + \eta f_2(\epsilon)$ ;  $h_8 = \tau_3$  – время затухания автокорреляционной функции  $R(\tau_3)$ ;  $h_9 = \sigma_{\epsilon x}^2 / \sigma_x^2$  – соотношение дисперсии шум/сигнал для независимых переменных;  $h_{10} = \max r_{\epsilon x}$  – максимальная корреляция ошибок с переменными  $x$ .

Характеристики  $h_j$  могут изменяться в широких пределах, однако на основании проведенных исследований реальных процессов в [4] интервалы изменения по каждой  $h_j$  следующие:  $h_1 = 20 \div 200$ ;  $h_2 = 5 \div 50$ ;  $h_3 = 10^{-3} \div 1$ ;  $h_4 = 0 \div 10$ ;  $h_5 = 1 \div 100$ ;  $h_6 = 0 \div 60$ ;  $h_7 = 0 \div 20$ ;  $h_8 = 0 \div 50$ ;  $h_9 = 0 \div 60$ ;  $h_{10} = 0 \div 0.5$ .

Интервалы исследования необходимо принять более широкими, чем они обычно встречаются на практике. Это позволит изучить поведение алгоритма в экстремальных ситуациях на границах области изменения параметров объекта. Для того чтобы соотнести приведенные выше параметры и их значения с конкретным методом, была разработана про-

грамма на встроенном языке программирования математического пакета MATLAB, генерирующая выборки с варьируемыми статистическими характеристиками. Результатом выполнения программы в MATLAB является массив данных, содержащий значения параметров выборки и наиболее подходящий для ее анализа метод. В данной работе проведены исследования по методам МНК, ВМНК и ОМНК. Далее предполагается провести также исследования с учетом методов ММП и ММ.

Для того чтобы в дальнейшем соотносить подаваемую на вход интеллектуальной системы выборку с определенным методом ее обработки, необходимо построить дерево решений.

### 5.3. Построение дерева решений

Для построения дерева решений может быть использован один из множества существующих алгоритмов: CHAID, Exhaustive CHAID, C&RT, QUEST [6].

Дерево решений было построено на основе алгоритма C&RT, разработанного Л. Брейманом с соавторами (Breiman и др., 1984). Данный метод представляет собой алгоритм построения бинарного дерева решений – дихотомической классификационной модели; каждый узел дерева при разбиении имеет только двух потомков. Ошибка при выборе метода на основе дерева решений составляет 3,51% и является наименьшей среди предлагаемых алгоритмов.

Полученное дерево решений приведено на рис. 2, где V1 – количество измерений; V2 – количество переменных; V3 – дисперсия ошибки; V4 – процент ошибки в выборке; V5 – целевая переменная.

Из рис.2 видно, что в условиях гетероскедастичности применение МНК не дает эффективных оценок параметров модели. В большинстве случаев необходимо использовать метод ОМНК.

На основе построенного дерева решения было разработано ядро интеллектуальной системы выбора метода прогнозирования финансовых временных рядов в условиях гетероскедастичности.

### 5.4. Апробация полученных решений

Поиск метода построения модели прогноза был проведен для финансового временного ряда официального курса гривни по отношению к доллару США, установленного Национальным банком Украины за период с января 1996 г. по декабрь 2001 г. (среднее за период). Данные получены из официального сайта Национального банка Украины [7]. В качестве фундаментальных факторов, влияющих на курс гривни, были выбраны: валовой внутренний про-



Рис. 2. Дерево решений

дукт в фактических ценах; денежные доходы населения, млн. грн; денежные затраты населения, млн. грн; уровень безработицы (на конец периода), в процентах; индекс цен потребительского рынка; индекс цен производителей промышленной продукции; размер дефицита бюджета; ставка рефинансирования; межбанковская ставка; депозитная ставка; ставка кредитования.

Анализ построенного прогноза показал, что реальные данные курса валют практически полностью попадают в доверительный интервал. Некоторые незначительные отклонения реального курса от прогнозных значений обусловлены тем, что при проведении расчета не учитывались такие качественные факторы, как административное регулирование валютного рынка, тип проводимой в стране фискально-денежной политики, а также интервенции центрального банка на межбанковском валютном рынке.

Расчеты были проведены с использованием математического пакета MATLAB.

## 6. Выводы

Научная новизна: разработана структурная схема интеллектуальной системы выбора метода прогнозирования финансовых временных рядов; впервые предложено использовать дерево решений для определения метода обработки данных в зависимости от статистических характеристик временного ряда. Построено дерево решений на основе алгоритма C&RT, которое является основой интеллектуальной системы выбора метода обработки данных.

Практическая ценность: в процессе проведения исследований была разработана программа, которая позволила провести имитационное моделирование временных рядов с различными статистическими характеристиками в условиях гетероскедастичности, что дало возможность оценить эффективность различных методов прогнозирования в условиях гетероскедастичности. Предложенные решения прошли апробацию при составлении научного прогноза финансового временного ряда официального курса гривны по отношению к доллару США. В результате апробации (составлении прогноза) показана высокая эффективность использования предложенного подхода. Все процедуры и модели прогноза были рассчитаны с помощью математического пакета MATLAB.

Созданное ядро интеллектуальной системы позволяет получать эффективный прогноз в различных отраслях хозяйственной деятельности и науки. Наиболее эффективно применение данной разработки в задачах составления прогнозов курсов валют, оценки валютного риска, котировок акций, уровня цен на нефть, газ, уголь, электроэнергию и др.

**Список литературы:** 1. *М. Кендэл*. Временные ряды. М.: Финансы и статистика, 1981. 199 с. 2. *Грін, Вільям Г.* Економетричний аналіз: Пер. з англ. А.Олійник, Р.Ткачук.; Наук. ред. пер. О.Комашко; Передм. О.І.Черняка, О.В.Комашка. К.: Видавництво Соломії Павличко „Основи”, 2005. 1197 с. 3. *Бородич С.А.* Эконометрика / Пособие для ВУЗов. Мн.: Новое знание, 2004, 596 с. 4. *Шамша Б.В., Гурій А.М., Дудар З.В., Левикін В.М.* Математичне забезпечення інформаційно-управляючих систем: Підручник для студентів ВНЗ. Харків: ТОВ „Компанія СМІТ”, 2005. 448 с. 5. *Титов О.А.* Математические методы обработки наблюдений. СПб.: Санкт-Петербургский Государственный Университет, 2001. 70 с. 6. *Breiman L., Friedman J.H., Olshen R.A., and Stone C.T.* Classification and Regression Trees. Wadsworth, Belmont, California, 1984. 7. Сайт Национального Банка Украины. – <http://www.bank.gov.ua>.

Поступила в редколлегию 17.03.2007

**Одейчук Андрей Николаевич**, магистрант кафедры ИУС ХНУРЭ. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 8(057) 335-32-32.

**Шамша Борис Владимирович**, канд. техн. наук, проф., доц. кафедры ИУС ХНУРЭ. Адрес: Украина, 61166, Харьков, пр. Ленина 14, тел. 8(057) 343-27-78.

**Федоров Евгений Григорьевич**, аспирант кафедры ИУС ХНУРЭ. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 8(057) 716-82-85.