



УДК 62-504.6

© 2006 г. **Н.С. Безруков,**
Е.Л. Еремин, д-р техн. наук
(Амурский государственный университет, Благовещенск),
Е.В. Ермакова,
В.П. Колосов, д-р мед. наук,
Ю.М. Перельман, д-р мед. наук
(Дальневосточный научный центр физиологии и патологии дыхания
СО РАМН, Благовещенск)

АВТОМАТИЗИРОВАННАЯ СИСТЕМА «MEDICAL TOOLBOX» ДЛЯ ДИАГНОСТИКИ БРОНХИАЛЬНОЙ АСТМЫ ПО ПОКАЗАТЕЛЯМ РЕОЭНЦЕФАЛОГРАФИИ¹

В работе рассматриваются общие вопросы создания автоматизированной медицинской диагностической системы на основе Matlab-приложения с использованием гибридной сети.

Введение

В настоящее время необходимость применения современных информационных технологий в медицине и здравоохранении не вызывает сомнений, особенно в задачах принятия врачом «управленческих решений» по диагностике, стратегии лечения и т.п. В медицине резко увеличился поток и объем информации, в то время как способ ее анализа, как правило, остался прежним, иначе говоря, разрешение проблем обработки и анализа медицинской информации – это узкое место, сдерживающее дальнейшее развитие медицинской науки и повышение эффективности лечебно-диагностической помощи.

Результаты медицинских исследований как аналитических, так и описательных, в значительной мере зависят от правильной организации и планирования научных работ, надлежащего сбора, анализа и интерпретации полученных данных. Состоятельность и результативность подобных исследований приобретают особую актуальность в связи с разработкой

¹ Работа выполнена в рамках комплексной программы фундаментальных исследований ДВО РАН «Системный анализ и моделирование влияния экологических и социальных факторов на здоровье населения Дальневосточного региона на основе применения современных информационных технологий с целью прогнозирования и разработки систем поддержки принятия управленческих решений в социальной сфере».

общепринятой концепции «медицины, основанной на доказательствах» (evidence-based medicine, ЕВМ) [1]. Следуя основным положениям этой концепции, можно утверждать, что в медицинской практике должны рекомендоваться к широкому применению только те методы диагностики и лечения заболеваний, эффективность и безопасность которых подтверждена в исследованиях, выполненных с единых научно-методологических позиций.

При этом существует настоятельная потребность в разработке автоматизированных медицинских приложений (АМП), способных обрабатывать в интерактивном пользовательском режиме имеющиеся и поступающие данные, а также позволяющих создавать диагностические системы по предложениям (гипотезам) эксперта-медика, подтверждаемым формализованными статистическими методами. АМП должны обладать способностью подгружать (интегрировать в себя) вновь создаваемые диагностические системы или независимые модули для их использования в научных изысканиях и реальной диагностической практике.

Программная среда Matlab в решении медицинских задач

К медицинским задачам чаще всего относят задачи общей и дифференциальной диагностики, прогнозирования, выбора стратегии и тактики лечения и др. Эти задачи практически всегда имеют неоднозначное, "нечеткое" решение. Специалист-медик на основе имеющихся знаний формулирует гипотезу по интересующей его задаче, а затем подтверждает гипотезу экспериментом или наблюдением.

В медицине и здравоохранении достаточно распространено использование различных статистических методов, в частности при принятии решений в задачах оценки состояния здоровья, прогноза, выбора стратегии, тактики, профилактики и лечения. Существует множество статистических критериев для подтверждения или опровержения выдвигаемых экспертом-медиком гипотез, среди которых отметим основные [2].

1. Критерий Стьюдента – параметрический критерий, для проверки гипотезы о средних значениях двух случайных наблюдений с нормальным распределением.

2. Критерий Вилкоксона – один из самых известных инструментов непараметрической статистики, являющийся альтернативой критерия Стьюдента для двух независимых наблюдений.

3. Критерий Хи-квадрат – непараметрический критерий, для проверки гипотезы о принадлежности распределения к заданному типу. Критерий можно применять более чем к двум выборкам.

При подтверждении выдвинутой гипотезы и определения значимых признаков перед экспертом встает задача создать на их основе систему диагностики, пригодную для применения на практике врачом. При этом использование специалиста по экспертным системам не всегда возможно.

В настоящее время имеется определенное количество способов синтеза экспертных систем, разработанных на основе методов эволюционного программирования, дерева решений, кластеризации, множественной линейной регрессии, а также с помощью генетических алгоритмов, нейронных сетей и нечеткой логики [3].

Существует значительное число унифицированных программных продуктов по статистике, анализу данных и синтезу экспертных систем (например, SPSS, «Statistica for Windows», MS Excel, PolyAlalyst, Matlab и многие другие), имеющих широкую область применения.

Из перечисленных выше программных средств выделим PolyAlalyst [3] и Matlab [4,5], в первую очередь в связи с тем, что в них реализованы современные алгоритмы построения и обучения нейронных сетей и средств нечеткой логики. В основу PolyAlalyst заложена высокоэффективная технология искусственного интеллекта Data Mining. При обработке исходных данных она позволяет обнаруживать многофакторные зависимости и придавать им функциональный вид. При хорошем интерфейсе и ряде других плюсов одной из слабых сторон с позиции развития PolyAlalyst является закрытость этого программного продукта, не позволяющая интегрировать в него другие методы анализа, – например, статистические критерии. В отличие от PolyAlalyst программная среда Matlab открытая, а это позволяет реализовать сколь угодно сложные алгоритмы анализа данных. Кроме того, Matlab имеет приложение GUI [4] – графический интерфейс пользователя, позволяющий создавать гибкие приложения с удобным интерфейсом, в частности ориентированном и на медицинского работника.

Matlab представляет собой интерактивную систему, основным объектом которой является массив, не требующий явного указания размерности. Обладая множеством специализированных функций, данный программный продукт может быть использован для автоматизации требуемых вычислений практически в любой области науки и техники. При этом он содержит средства построения моделей на основе нечеткой логики и гибридных сетей. Кроме того, по скорости выполнения вычислений его возможности превосходят потенциалы многих подобных систем. Matlab – операционная среда и язык программирования. Одна из наиболее важных сторон этого программного обеспечения состоит в том, что на его языке могут быть написаны программы для многократного использования. Пользователь может создать специализированные функции и программы, который оформляются в виде *m*-файлов.

Все существующие и вновь создаваемые ППП рассматриваемой системы функционируют в едином операционном пространстве. Это обстоятельство обуславливает наличие ряда следующих преимуществ:

каждый ППП обеспечивает робастность и высокую точность вычислений, что является следствием многолетних исследований разработчиков системы Matlab;

имеется возможность анализа, коррекции и применения *m*-функций в качестве шаблонов, написанных на языке Matlab, для разработки новых приложений;

каждый ППП может быть использован в рамках системы Matlab на любой вычислительной платформе.

В среде Matlab [5] реализован алгоритм синтеза нечеткой логики и нейронной сети в виде так называемой *гибридной сети* или *адаптивной нечеткой нейронной системы вывода ANFIS* (Adaptive Neuro-Fuzzy Inference System).

В настоящей работе в качестве основного средства разработки АМП «Medical Toolbox» выбрана программная среда Matlab. Структурная схема Matlab-приложения представлена на рис. 1.

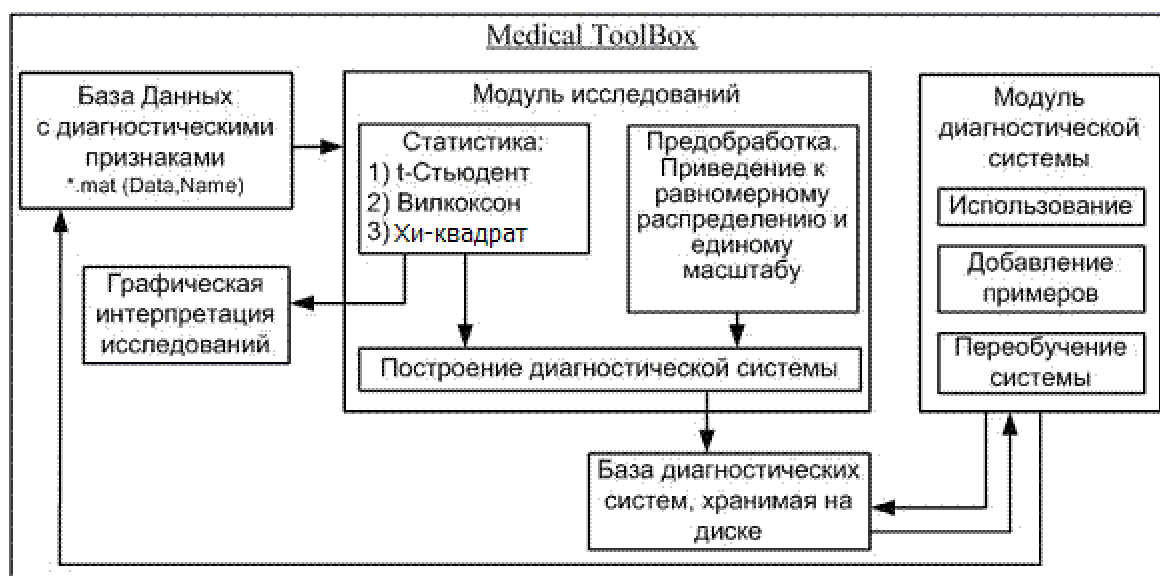


Рис. 1. Структура Matlab-приложения «Medical Toolbox».

Matlab-приложение состоит из двух основных модулей: модуля исследований и модуля диагностической системы, для которых на диске хранятся две базы: база данных с диагностическими признаками и база с диагностическими системами.

Модуль исследований имеет три основных составляющих: статистический анализ данных (критерии Стьюдента, Вилкоксона и Хи-квадрат), предобработка данных и алгоритм построения диагностической системы. Под предобработкой подразумевается увеличение информативности данных и приведение их к общей шкале измерения по методу, предложенному в работе [6]. Диагностическая система создается по алгоритму, описанному в работе [7]. В создании системы могут участвовать не более 10 признаков, номера которых пользователь может задать сам или они выберутся автоматически в зависимости от используемого критерия.

Модуль диагностической системы позволяет работать врачу с ранее созданными диагностическими системами, пополняя их по ходу работы

новыми примерами или данными, а при необходимости эти системы и переобучая.

Экранные формы Matlab-приложения «Medical Toolbox» на примере системы диагностики бронхиальной астмы

Реоэнцефалография (РЭГ) – это метод исследования пульсового кровенаполнения тканей головного мозга, основанный на его способности изменять свое сопротивление переменному току высокой частоты в зависимости от количества протекающей через него крови. С физической точки зрения реограмма отражает комплексное сопротивление живых тканей. Одними из основных показателей являются реографические индексы (РИ OR, РИ OL, РИ FR, РИ FL), отражающие интенсивность кровотока в исследуемой области. В общем метод дает 28 диагностических признаков. Целью настоящего исследования явилось создание диагностической системы по выявлению БА на основе использования показателей регионарного мозгового кровотока.

Прежде чем начать работать в приложении, необходимо данные привести к форме представления, принятой в среде Matlab. Для этого надо определить набор числовых параметров, подлежащих анализу. Данные могут быть представлены в виде прямоугольной таблицы, где каждая запись представляет собой отдельный объект или состояние объекта, а каждое поле – свойства или признаки всех исследуемых объектов. Последнее поле является выходным признаком, подлежащим определению (диагноз, метод

лечения). Если данные размещены в нескольких связанных между собой таблицах, то все равно необходимо привести их в прямоугольную форму. Затем данные нужно сохранить в текстовом файле и открыть его в среде Matlab как матрицу с именем Data. Аналогично можно экспортировать название признаков в матрицу строк Name. Матрицы должны быть сохранены в файле с расширением *.mat, что позволит открыть его в приложении «Medical Toolbox» (рис. 1).

Изначально данные РЭГ по здоровым и больным пациентам хранятся в двух матрицах, которые затем объединяют в файле REG.mat, причем здоровых маркируют значением –1, а больных 1.

При запуске модуля исследований (рис. 2) открывается окно «Статистика» (рис. 3).

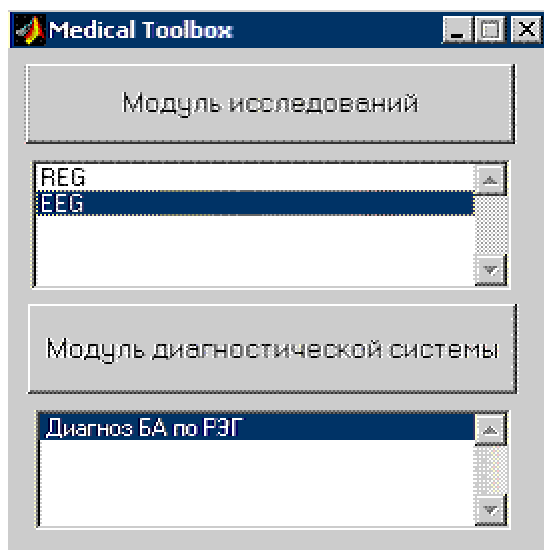


Рис. 2. Экранная форма Matlab-приложения «Medical Toolbox».

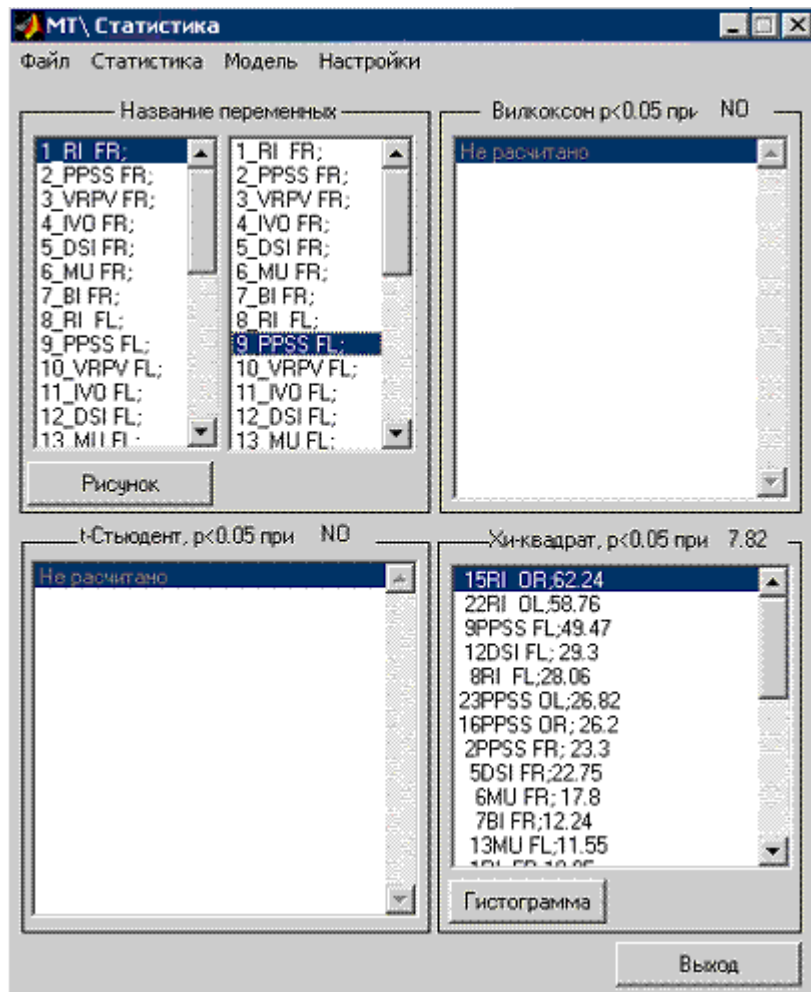


Рис. 3. Экранная форма окна «Статистика».

В разделе «Название переменных» отражаются названия всех признаков, имеющих в базе, в форме двух столбцов с порядковыми номерами. Этот раздел позволяет графически отображать (рис. 4) распределение

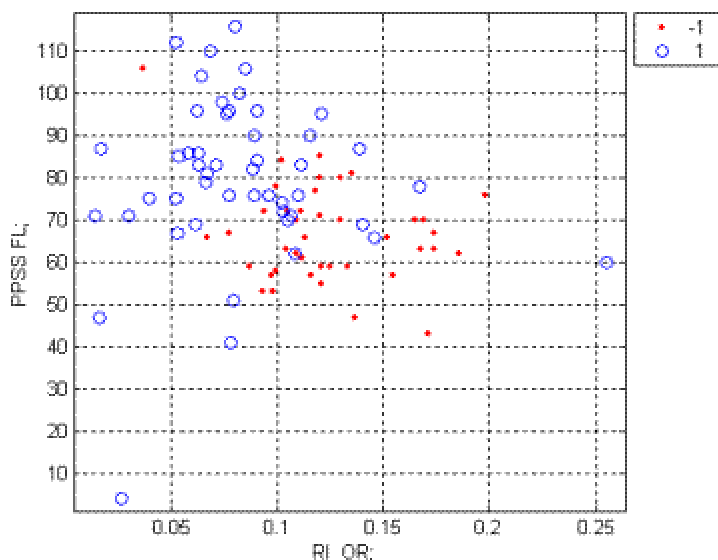


Рис. 4. Экранная форма распределения двух признаков относительно выходного признака (диагноза).

двух признаков относительно выходного признака (диагноза) при нажатии кнопки «Рисунок».

Matlab позволяет редактировать рисунки прямо в рабочей области, поэтому без особых усилий можно менять как цвет, так и форму отображения данных.

В модуле для анализа данных и подтверждения выдвигаемых гипотез рассчитываются один па-

раметрический (критерий Стьюдента) и два непараметрических (критерий Вилкоксона и Хи-квадрат) критерия. Их расчет осуществляется через меню «Статистика». Расчет любого критерия отражается строками (порядковый номер признака, название признака, значение критерия) в рабочей области окна «Статистика» в порядке убывания значимости, а также выводится через командную строку, откуда значения критерия можно скопировать.

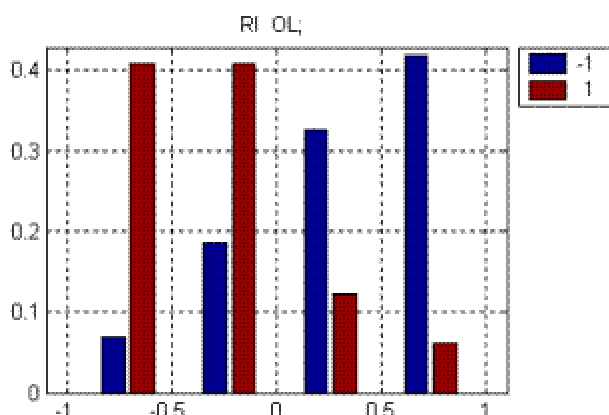


Рис. 5. Экранная форма распределения признака относительно диагноза в виде гистограммы.

существовании закономерностей в признаке, а сам рисунок наглядно демонстрирует характер закономерностей. Перед запуском диагностической системы данные необходимо подвергнуть предобработке, используя меню «Модель/Предобработка». Диагностическая система создается в окне «Построение диагностической системы» (рис. 6) из меню «Модель/ Создать».

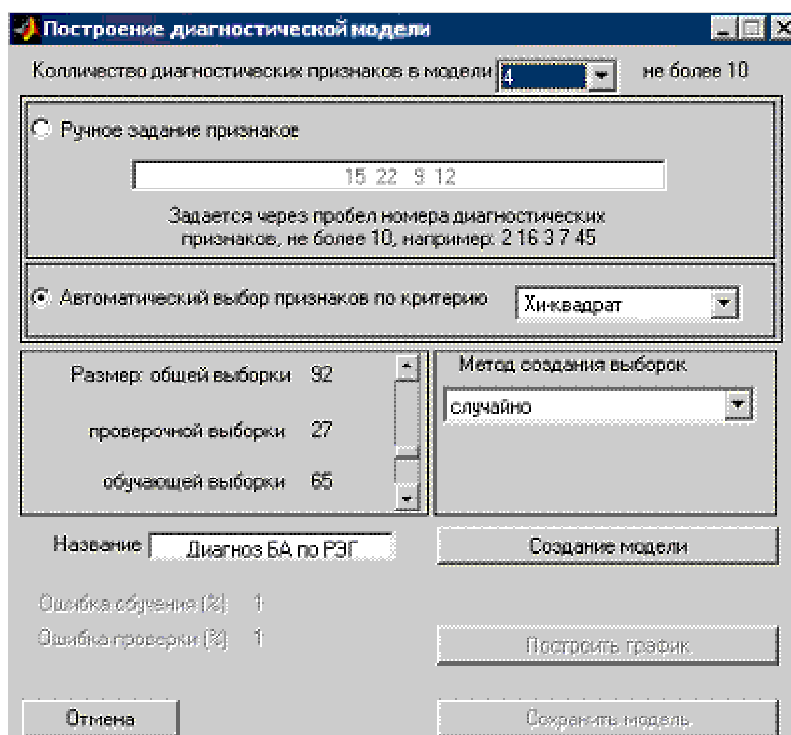


Рис. 6. Экранная форма окна «Построение диагностической системы».

Для проверки полученных результатов все имеющиеся данные разбиваются на проверочную и обучающую выборки. Размер и характер деления определяются пользователем. По величине ошибок, вычисляемых при

обсчете системой обучающей и проверочной выборок, пользователь судит о том, насколько адекватна и статистически значима построенная система.

Заключение

Анализируя структуру приложения «Medical Toolbox», можно сделать вывод, что она достаточно сложна. Но, несмотря на это, приложение обладает удобным и понятным интерфейсом, что значительно упрощает процесс взаимодействия пользователя с программным продуктом. Пользователю даже не обязательно пользоваться клавиатурой, особенно в режиме диагностирования.

Все множество данных скрыто и отображается только в форме названий признаков, тем самым снижается уровень информационной нагрузки на пользователя. Гибкий интерфейс позволяет использовать пакет либо для статистического анализа данных и интерпретации результатов в наглядной форме, либо как диагностическую систему для практикующих врачей.

Рассмотренная диагностическая система (рис. 7) соответствует представлениям эксперта о том, что при диагностике БА по реоэнцефалограмме необходимо использовать реографические индексы, два из них (РИ OR, РИ OL) входят в систему.

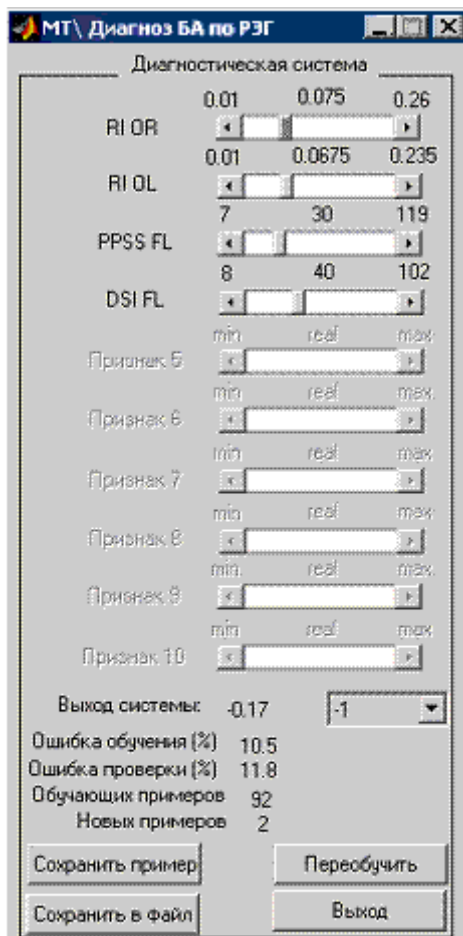


Рис. 7. Экранная форма системы диагностики бронхиальной астмы по реографическим индексам.

ЛИТЕРАТУРА

1. Основы эпидемиологии и статистического анализа в общественном здоровье и управлении здравоохранением. – М.: Москов. медакадемия им. И.М. Сеченова, 2003.
2. *Леонов В. П.* Наукометрия статистической парадигмы экспериментальной биомедицины // Вестник Томского гос. университета. – 2002. – № 275. – С. 17-24.
3. *Арсеньев С.* Извлечение знаний из медицинских баз данных. <http://www.megaputer.ru>.
4. *Еремин Е.Л., Еремина В.В., Семичевская Н.П., Шевко Д.Г.* Алгоритмы и S-модели гибридных систем адаптивного управления. – Благовещенск: АмГУ, 2005.
5. *Круглов В.В., Дли М.И., Голунов Р.Ю.* Нечеткая логика и искусственные нейронные сети. <http://www.neuroproject.ru>.
6. *Ежов А.А., Шумский С.А.* Нейрокомпьютинг и его применение в экономике и бизнесе. <http://www.neuroproject.ru>.
7. *Безруков Н.С., Еремин Е.Л.* Построение и моделирование адаптивной нейронечеткой системы в задачах медицинской диагностики // Информатика и системы управления. – 2005. – №2(10). – С. 36-46.