

# Программная реализация «системы поддержки принятия решений для психологического и педагогического тестирования»

**П. Н. Думин\***,  
выпускник кафедры прикладной информатики и мультимедийных технологий факультета информационных технологий Московского городского психолого-педагогического университета

Представлена программная реализация системы поддержки принятия решений для психолого-педагогического тестирования. В основе предложенной технологии тестирования лежит возможность использования марковских сетей в качестве модели, описывающей динамику прохождения тестов, параметры которой возможно идентифицировать по экспериментальным данным. Система ориентирована на тестирование детей дошкольного возраста на предмет готовности к обучению в школе. Основными преимуществами системы является возможность получения вероятностной оценки принадлежности испытуемого к одной из диагностируемых групп, а также расчет дифференцирующей силы для непредъявленных тестов. На основе этих оценок специалист, проводящий процедуру тестирования, может выбрать близкую к оптимальной последовательность тестов, предъявление которых в выбранном порядке позволит за наименьшее время дифференцировать самый вероятный диагноз. Оценка вероятностей диагноза и построение прогноза происходит с использованием байесовского подхода.

**Ключевые слова:** марковские модели, психологическое тестирование, система поддержки принятия решений.

Компьютерное тестирование все более широко используется в психодиагностике по разным причинам, среди которых можно упомянуть интенсификацию процесса тестирования, а также тот факт, что опытный психолог за меньшее время получает больше оснований для принятия решения (в случае диагностирующего теста).

В большинстве случаев системы тестирования представляют собой реализованные в

компьютерном варианте методики. Отличие от классической формы проведения тестирования заключается в том, что результаты тестирования подсчитываются автоматически, а оценка результатов тестирования производится на основе каких-либо условий (количественного сравнения результата с нормами, выделенными для конкретного теста).

Современные технологии тестирования позволяют не только строго оценивать ре-

---

\*dumin.pn@gmail.com

зультаты тестирования, но и прогнозировать различные показатели процедуры, такие как динамика прохождения тестов, получение вероятностных характеристик результатов тестирования, дифференцирующую силу предъявленных тестов.

Система поддержки принятия решения, описанная далее, после предъявления очередного теста выдает вероятностную оценку принадлежности тестируемого к одной из групп, предусмотренных предъявленным тестом, а также рекомендации, касающиеся выбора следующего теста из имеющейся батареи.

В данном исследовании рассматривалась проблема тестирования дошкольников на предмет готовности к школе. Здесь и далее под тестированием понимается процедура, позволяющая разбить испытуемых на две группы – готовых и не готовых к школе. Как правило, чтобы определить, готов ли ребенок к школе, психолог использует ряд методик, после чего, в зависимости от успешности выполнения заданий, делает соответствующий вывод. Обычно психологу приходится использовать последовательно множество единообразных методик, что приводит к утомлению испытуемого. Такое состояние тестируемого может сильно исказить реальные результаты и, как следствие, привести к неверному решению специалиста. Чтобы избежать подобных ситуаций, целесообразно предъявлять не все тесты из подготовленной батареи, а только те, которые будут обладать наибольшей дифференцирующей способностью.

Среди факторов, влияющих на тестирование, помимо правильности ответов на предъявленные задания также можно выделить динамику прохождения этой процедуры. Сочетания определенных значений данных факторов позволяют получить дополнительную информацию об уровне измеряемой способности испытуемого. Процесс тестирования в данном случае может быть описан как марковская модель с непрерывным временем. Подробное описание предложенной модели тестирования приведено в статье «Концепция системы поддержки принятия решений для психологического тестирования» [4].

Стратегия лица, принимающего решение об оценке результатов тестирования, обуслов-

лена необходимостью обеспечения наибольшей дифференциации самого вероятного диагноза за наименьшее время. Таким образом, в качестве следующего необходимо предъявить тот тест, после прохождения которого испытуемого с наивысшей вероятностью можно будет отнести к одной из диагностируемых групп.

Определенный уровень развития измеряемых характеристик позволяет сделать вывод о готовности испытуемого, прошедшего тестирование, к школе [3] Ниже дано описание тестов, выбранных для системы.

Методика «Эталоны» (автор О.М. Дьяченко) направлена на изучение уровня развития восприятия и содержит задачи, требующие соотнесения формы предметов с заданными образцами (эталонами). В качестве входных данных для программного модуля, реализующего процедуру прохождения данной методики, следует рассматривать 4 карточки с заданиями. Процедура тестирования подразумевает последовательное выполнение заданий теста. При выполнении каждого задания испытуемый на одной из картинок входного набора отмечает элементы картинки, которые, по его мнению, являются правильным выполнением задания (т.е. фигуры, которые соответствуют эталону). Выходным параметром для данной методики является количество верно отмеченных фигур.

Методика «Схематизация» (автор Р.И. Бардина) направлена на диагностику уровня развития образного мышления. Ребенку предлагается найти адресат письма. В поисках нужного объекта ребенок должен сориентироваться в лабиринте дорожек, представленном ему в виде схемы. Средством ориентировки служат указанные в «адресе» повороты и материальные объекты – деревья, цветы, грибы. Ребенку предъявляется 10 задач в порядке возрастания сложности: лабиринт становится не только все более разветвленным, но постепенно включает в себя необходимость предвосхищающего действия и сочетание двух систем признаков. Методика имеет балльную систему оценки, стандартизована. В результате проведения методики выносится заключение о соответствии уровня развития наглядно-образного мышления возрастной норме.

Методика «Систематизация» (автор Н. Б. Венгер) направлена на выявление уровня развития действий логического мышления. Ребенок должен распределить геометрические фигуры в поля матрицы, соотнеся их с формой и размером других фигур, стоящих по краям этой матрицы.

В качестве входных данных программного модуля, реализующего процедуру прохождения данной методики, следует рассматривать 9 картинок с заданиями. Процедура тестирования подразумевает последовательное выполнение заданий теста, при выполнении которых испытуемым на каждой картинке из входного набора отмечаются конкретные элементы картинки, которые, по мнению ребенка, являются правильным выполнением задания (т. е. места, где должны быть расположены фигуры). Выходным параметром для данной методики является количество верно отмеченных фигур.

Методика «Сапожки» (Н.И. Гуткина) позволяет исследовать обучаемость ребенка, т. е. проследить, как он пользуется для решения задач правилом, с которым раньше никогда не встречался. Трудность предлагаемых задач постепенно возрастает за счет введения в них объектов, по отношению к которым можно применить усвоенное правило только после осуществления необходимого процесса обобщения. Используемые в методике задачи построены таким образом, что для их решения требуется осуществить эмпириче-

ское или теоретическое обобщение. Под эмпирическим обобщением понимается умение классифицировать предметы по существенным признакам или подводить под общее понятие. Под теоретическим обобщением понимается обобщение на основе содержательной абстракции, когда ориентиром служит не конкретный отличительный признак, а факт наличия или отсутствия отличительного признака независимо от формы его проявления. Таким образом, методика «Сапожки» позволяет исследовать обучаемость детей, а также особенности развития процесса обобщения.

### Программная реализация системы

Программная реализация СППР выполнена в среде графического программирования LabView 2010. На определенном этапе (для идентификации параметров построенных моделей) были использованы электронные таблицы Excel.

Для разделения рабочего пространства специалиста-психолога, проводящего тестирование, и испытуемого было принято решение о распределении функционала СППР на несколько компьютеров с использованием технологии «клиент-сервер». Основные модули системы расположены на сервере (машине эксперта), модули предъявления тестов расположены на машине клиента (за которой находится испытуемый). Схема функционирования СППР изображена на рис. 1.



Рис. 1. Принципиальная схема функционирования СППР

В системе содержатся 4 теста. Один из них дихотомический – результат его решения может быть либо верным, либо неверным. Остальные тесты имеют несколько уровней успешности выполнения, однако в связи с особенностями архитектуры системы приводятся к аналогичной форме оценки, т. е. задания этих тестов также оцениваются либо как выполненные верно, либо как выполненные неверно. Реализуется это посредством сложения интенсивностей переходов между объединенными состояниями. На рисунке 2 изображен фрагмент модели теста «Эталоны».

Как видно, тест имеет три уровня успешности выполнения заданий. Из состояния  $S_0$  (начало тестирования) испытуемый может переместиться в состояние  $S_{1,0}$ , если задание выполнено неверно, в состояние  $S_{1,1}$  – если задание выполнено частично верно, и в состояние  $S_{1,2}$  – если задание выполнено полностью верно.

Чтобы иметь возможность запрограммировать данный тест, необходимо сократить количество состояний. Это можно сделать путем их объединения. При объединении состояний интенсивности, в них входящие, складываются между собой. Таким образом, модель, используемая в системе, будет иметь следующий вид (рис. 3).

При прохождении тестирования на клиентской машине запущена программа, которая находится в режиме ожидания служебной команды. После получения команды от сервера происходит запуск соответствующего теста. Каждый предъявляемый тест представляет собой последовательность картинок, на которых расположены LED-контроллеры. Для решения задачи необходимо активировать один или несколько контроллеров в зависимости от задания.

После того как ответ на задачу дан, необходимо нажать кнопку «Следующий вопрос». При нажатии на эту кнопку в массив записываются текущий ответ тестируемого и происходит инициализация следующего задания. Перед инициализацией свойство True/False всех контроллеров меняется на False (контроллеры становятся неактивными). Структура хранения файлов, требуемых для предъявления тестов, представляет собой каталоги с соответствующими тестам названиями. Проверка правильности выполнения задания осуществляется путем сравнения логических переменных, соответствующих LED-контроллерам (значения определенного числа которых были изменены при выполнении задания – активации контроллеров), с заранее известными ответами, хранящими-

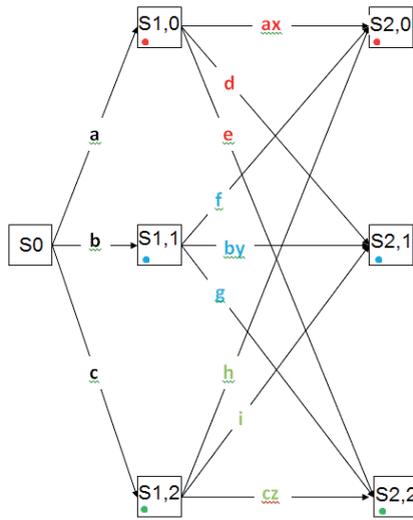


Рис. 2. Фрагмент модели для теста «Эталоны»

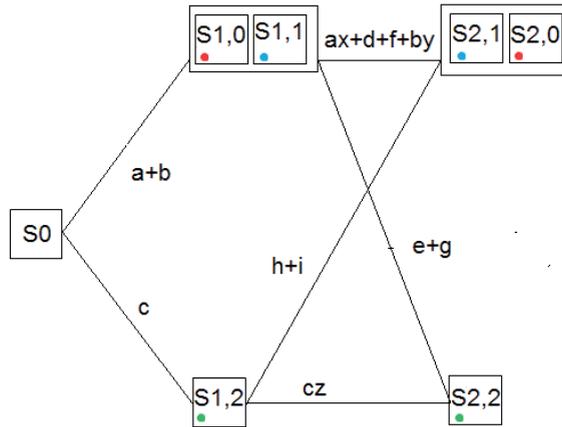


Рис. 3. Модель теста, адаптированная для использования в системе

ся в текстовом файле. Во время прохождения теста также сохраняются временные показатели, за которые испытуемый дал ответ на предъявленное задание. Время, затраченное на решение задания, определяется разностью между моментом времени, когда задание было инициализировано, и моментом, когда испытуемый нажал кнопку перехода к следующему заданию.

При последовательном решении предъявляемых заданий в системе ведется запись состояний. Происходит это после перехода испытуемого к следующему заданию теста.

Для внесения в подпрограмму предъявления новых тестов пользователю необходимо указать путь к файлам теста, к файлам, содержащим правильные ответы на задания, а также перераспределить LED-контроллеры на форме предъявления тестов. Каждый LED-контроллер имеет собственный идентификатор. При перераспределении контроллеров для оформления задания необходимо соблюдать соответствие индикаторов файлу с ответами на задание.

Как упоминалось выше, система функционирует по сети. При передаче данных используется протокол TCP. На клиентской машине ожидается служебная команда запуска теста. После получения команды вызывается соответствующий тест и соединение закрывается. Подпрограмма тестирования на выходе предоставляет результаты, время решения заданий

и историю состояний, в которых находился испытуемый. Эти данные записываются в соответствующие массивы. Далее массивы форматируются определенным образом для удобства передачи на сервер. После форматирования TCP-порт открывается, на сервере запускается процедура приема данных. Массивы, содержащие в себе результаты процедуры тестирования, отправляются на сервер для дальнейшей обработки и подсчетов необходимых оценок. Получение данных на стороне сервера производилось аналогичным образом. Массивы, переданные с клиента, разбирались определенным образом (для этого было применено форматирование массивов). После разбора данные сохранялись и передавались для расчетов следующим подпрограммам.

Работа с системой со стороны специалиста, проводящего тестирование, заключается в выборе требуемых тестов на главной форме. Помимо кнопок, вызывающих соответствующие тесты, на форме расположены гистограммы вероятностей, и кнопка, вызывающая форму ввода параметров моделей. После выбора теста на клиентскую машину отправляется служебная команда, обрабатывая которую подпрограмма клиента предъявляет соответствующий тест. После получения информации о пройденном тесте запускается подпрограмма подсчета дифференцирующей силы непредъявленных тестов. Значения вероятностей, полученные после данной про-

цедуры, используются для построения гистограмм прогноза.

Генерация системы уравнений реализована в отдельной подпрограмме. На вход поступают значения параметров модели, на выходе – массив, состоящий из строк.

Перед запуском тестирования в системе происходит интегрирование систем уравнений с сохраненными ранее параметрами для заложённых тестов.

На вход подпрограмме поступают параметры моделей, время начала и окончания интегрирования, шаг и начальное значение. Далее происходит генерация системы уравнений и ее интегрирование. Результаты процедуры записываются в виде таблиц с указанием теста и диагностируемой группы, для которых получены вероятности пребывания в различных состояниях модели.

Оценки вероятностей, полученные на данном этапе, используются для подсчета показателей готовности к школе и построения прогноза дифференцирующей силы теста (рис. 4). Перед предъявлением первого теста (когда оценки, подсчитанные системой, отсутствуют) для подсчета используются априорные вероятности принадлежности испытуемого к одной из двух диагностируемых групп. После процедуры подсчета оценок на главной форме приложения строятся гистограммы распределения вероятностей.

Подобные гистограммы строятся для каждого непредъявленного теста. Дифференцирующая сила определяется специалистом на

основе разницы между вероятностью испытуемого относиться к одному из классов при условно верном и условно неверном решении очередного теста. Для того чтобы строить такие прогнозы, в системе используются паттерны, содержащие в себе условные значения времени и правильности выполнения заданий. При исследовании на экспериментальной группе было зафиксировано время, за которое испытуемые, готовые к школе, проходили тест, верно решая предъявляемые задания. Для паттерна, на основе которого строится прогноз дифференцирующей силы при условно верном прохождении теста, использовались усредненные значения времени успешного решения заданий. Для паттерна, прогнозирующего дифференцирующую силу при условно неверном прохождении теста, показатели времени были значительно увеличены, а задания условно были решены неверно.

Для дополнения системы методиками специалисту необходимо выполнить следующие шаги:

- 1) выбрать методику тестирования;
- 2) внести в систему данные о количестве вопросов каждого теста, содержащегося в выбранной методике;
- 3) в соответствии со структурой хранения уже существующих в системе методик оформить файлы с вопросами и ключами каждого задания теста;
- 4) ввести в программу идентифицированные параметры моделей, созданных для каждого теста и для каждой группы испытуемых.



Рис. 4. Прогноз дифференцирующей силы теста. Две пары гистограмм для двух возможных исходов выполнения теста

### Идентификация моделей

Для идентификации параметров модели были использованы электронные таблицы Excel. Созданные модели были описаны СДУ Колмогорова и проинтегрированы в Excel. Для решения этой задачи был использован метод Эйлера второго порядка – для него возможно закодировать схему интегрирования. После этого этапа устанавливаются начальные значения искоемых параметров модели и запускается процедура оптимизации. Если значения, полученные при первой итерации, не удовлетворяют требованиям, на начальные значения параметров накладываются соответствующие ограничения, процедура запускается повторно. Окончательное решение об останове процедуры поиска параметров принимается при достижении оптимальных значений. Полученные в ходе этапа оптимизации значения интенсивностей переходов заносятся в систему в качестве параметров моделей в специально спроектированный для этого модуль.

После подсчета теоретических вероятностей нахождения в различных состояниях модели мы получили возможность сравнить частоты пребывания, полученные в модели, с реальными частотами, полученными в ходе исследования на экспериментальной группе. На рисунке 5 изображена гистограмма распределения предсказанных и реальных частот пребывания в состояниях модели.

Для контрольных моментов времени подсчитывается квадрат разности между теоретическими и экспериментальными значениями вероятностей пребывания в разных состояниях. Полученная статистика Пирсона минимизировалась встроенным методом оптимизации в Excel. Подсчитанные значения (интенсивности переходов между состояниями) использовались в качестве параметров моделей и были внесены в систему.

Несмотря на то что вероятности выполнить последние задания близки по значениям, группа испытуемых, готовых к школе, быстрее справилась с тестированием в целом, показывая высокие результаты с самого начала тестирования (вероятность решить первое задание теста у них составила около 0,4; для этого же момента времени вероятность правильного выполнения другой группы составила примерно 0,1). Все задания первой группой были решены уже к моменту времени, соответствующего 400 секундам, в то время как другая группа справилась с заданиями теста значительно позже (к 600 секундам). Это позволяет говорить, что подтверждается предположение о различии динамики выполнения заданий между диагностируемыми группами и мы имеем возможность прогнозировать результаты выполнения непредъявленных тестов, опираясь на паттерны, построенные для них.

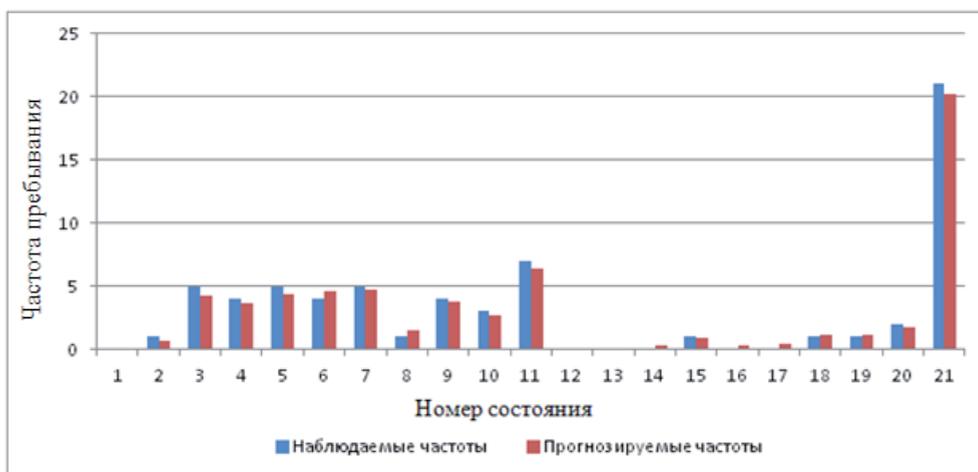


Рис. 5. Гистограмма распределения частот пребывания в состояниях модели

Программно реализована система поддержки принятия решений, базирующаяся на концепции тестирования, использующей сети Маркова как модели прохождения тестов. Среди преимуществ системы можно выделить обеспечение специалиста-психолога вероятностными оценками принадлежности ис-

пытываемого к одной из диагностируемых групп, а также возможность получения прогнозов дифференцирующей силы непредъявленных тестов. Последняя выражается в оценках вероятностей диагнозов при условии принадлежности испытуемого к различным диагностируемым группам.

#### Литература

1. Бардина Р.И и др. Диагностика умственного развития детей старшего дошкольного возраста (от 5 до 7 лет). М., 1996.
2. Крамер Г. Математические методы статистики. М., 1976.
3. Куравский Л.С., Баранов С.Н., Корниенко П.А. Обучаемые многофакторные сети Маркова и их применение для исследования психологических

характеристик // Нейрокомпьютеры: разработка и применение. 2005. № 12.

4. Куравский Л.С., Марголис А.А., Юрьев Г.А., Мармалюк П.А. Концепция системы поддержки принятия решений для психологического тестирования // Психологическая наука и образование. 2012. № 1.

## Software implementation of "decision support system for psychological and pedagogical testing"

P. N. Dumin,

*Graduate, Chair of Applied Informatics and Multimedia Technologies, Department of Informational Technologies, Moscow State University of Psychology and Education*

We present a software implementation of decision support system for psychological and pedagogical testing. The proposed testing technology is based on use of Markov nets as a model describing the dynamics of tests completion, parameters of which may be identified from experimental data. The system is focused on testing readiness for school in preschool children. The main advantages of the system are possibility of probabilistic assessment of subject's affiliation to one of the diagnostic groups and computation of the differentiating power for tests not being presented. Based on these estimates, an expert conducting the test will be able to choose tests order close to the optimal, which will allow to differentiate the most likely diagnosis in the shortest time. Estimation of the diagnosis probability and prognosis is done using the Bayesian approach.

**Keywords:** Markov model, psychological testing, decision support system.

#### References

1. Bardina R.I i dr. Diagnostika umstvennogo razvitiya detej starshego doshkol'nogo vozrasta (ot 5 do 7 let). M., 1996.
2. Kramer G. Matematicheskie metody statistiki. M., 1976.
3. Kuravskij L.S., Baranov S.N., Kornienko P.A. Obuchaemye mnogofaktornye seti Markova i ih primenenie dlja issledovaniya psihologicheskikh

harakteristik // Nejrokomp'jutery: razrabotka i primenenie. 2005. № 12.

4. Kuravskij L.S., Margolis A.A., Jur'ev G.A., Marmaljuk P.A. Konceptija sistemy podderzhki prinjatija reshenij dlja psihologicheskogo testirovanija // Psihologicheskaja nauka i obrazovanie. 2012. № 1.