

ОБЩАЯ ПСИХОЛОГИЯ

УДК 371.263

С. А. Сухушин

ЭКСПЕРТНАЯ СИСТЕМА ДЛЯ ВЫРАБОТКИ РЕКОМЕНДАЦИЙ ПО УЛУЧШЕНИЮ КАЧЕСТВА ПЕДАГОГИЧЕСКИХ И ПСИХОЛОГИЧЕСКИХ ТЕСТОВ

Томский государственный педагогический университет

Тестирование все чаще используется как метод объективного контроля знаний, хотя метод обладает как преимуществами, так и недостатками по сравнению с традиционными методами контроля. К преимуществам, в частности, относят следующие факторы:

- возможность автоматизации как проведения тестирования, так и проверки результатов;
- объективное оценивание и одинаковые условия для всех испытуемых и др. [1]

Важным фактором является качество контрольно-измерительных факторов, в данном случае тестов. Методы современной тестологии дают возможность не только получать объективные оценки уровня знаний обучаемых, но и оценивать качество тестов и отдельных тестовых заданий. Результат достигается за счет применения методов математической статистики, и существует несколько подходов для оценки уровня знаний обучаемых и качества тестовых заданий:

- применение классической теории тестов;
- применение однопараметрической модели Раша;
- применение двухпараметрической модели Бирнбаума;
- применение трехпараметрической модели Бирнбаума.

В представленном списке модели расположены от более простых к более сложным с точки зрения применяемого математического аппарата. Особенностью модели Раша и различных моделей Бирнбаума является их построение на основе использования логистической функции. При этом модель Раша является частным случаем двухпараметрической модели Бирнбаума, которая, в свою очередь, является частным случаем трехпараметрической модели.

Более подробное описание различных моделей можно найти в специальной литературе [2]. Здесь перечислим основные параметры, которые позволяют рассчитать каждая из моделей.

Наиболее часто в нашей стране используется **классическая теория тестов**. Данная модель также является исторически самой первой и основана на расчете ряда коэффициентов по матрице тестирования, содержащей информацию – правильно или неправильно каждый испытуемый отвечает на каждое тестовое

задание. Подход позволяет рассчитывать следующие параметры.

Уровень знаний испытуемого. Этот показатель классически определяется как доля правильных ответов, данных испытуемым на тестовые задания. Модель может быть усложнена путем введения весовых коэффициентов значимости тех или иных заданий на основе показателя сложности задания.

Сложность (трудность) тестового задания. Этот коэффициент показывает, насколько сложно то или иное задание, и определяется отношением испытуемых, правильно выполнивших данное задание, к общему числу испытуемых. Обратим внимание, что формальное название характеристики противоположно ее смыслу. То есть на самом деле задание тем легче, чем выше ее показатель сложности.

Коэффициент дискриминации тестового задания. Показывает, насколько хорошо данное задание различает «сильных» и «слабых» испытуемых. Задания с низким коэффициентом дискриминации считаются диагностически бесполезными и поэтому некачественными.

Коэффициент надежности. Данный коэффициент отражает надежность измерения, т.е. степень постоянства результатов.

Значимость (валидность) тестового задания. Данный коэффициент показывает, насколько результаты ответа испытуемыми на данное задание соответствуют их индивидуальным баллам. Задания с низким показателем валидности подлежат исключению или замене.

Основой для построения **однопараметрической модели Раша** также служит матрица тестирования. В основе модели лежит априорное положение о зависимости вероятности ответа определенного испытуемого на определенное задание от **уровня знаний испытуемого** и от **сложности задания** (1). Обработка матрицы тестирования по модели Раша – это подбор показателей уровня знаний каждого испытуемого и сложности каждого задания так, чтобы реальная матрица, полученная по результатам тестирования, как можно более ближе соответствовала закону

$$p_{ij} = \frac{1}{1 + \exp(\delta_i - \theta_j)}, \quad (1)$$

где p_{ij} – вероятность успешного выполнения j -го задания со сложностью δ_j , i -м испытуемым, с уровнем подготовленности θ_i .

Несмотря на то, что в основной формуле используется два параметра δ и θ , модель называется однопараметрической, так как вероятность успешного выполнения задания реально зависит не от самих параметров δ и θ , а от их разности. Единица измерения и называется **логитом**. При одном логите ($\delta = 1$ и $\theta = 1$) вероятность успеха $p = 0,5$, т.е. вероятность выполнения «среднего» задания «средним» участником должна быть равна 0,5. Функция такого типа в математике называется **логистической**.

В отличие от классической теории тестов, **сложность тестового задания** и **уровень знаний испытуемого** представлены в линейной шкале, а не порядковой, что дает возможность объективно измерять не только, какое из двух заданий сложнее, но и на сколько одно задание сложнее другого. Модель позволяет оценивать и качество тестовых заданий, на основе соответствия ответов на него модели требуемому распределению, задаваемому логистической функцией.

Двухпараметрическая модель Бирнбаума основана на предположении, что на вероятность ответа конкретного испытуемого на конкретное задание влияет также **коэффициент дискриминации**. Зависимость в данной модели следующая:

$$p_{ij} = \frac{1}{1 + \exp(-d_j(\theta_i - \delta_j))}, \quad (2)$$

где d_j – коэффициент дискриминации. Подобно коэффициенту дискриминации из классической теории тестов, коэффициент дискриминации в двухпараметрической модели Бирнбаума показывает, насколько хорошо задание различает «сильных» и «слабых» испытуемых, отличие состоит в том, что данный коэффициент дискриминации представлен в линейной шкале.

Трехпараметрическая модель Бирнбаума, кроме того, учитывает **вероятность угадывания** ответа на тестовое задание. Данный коэффициент показывает, насколько легко испытуемые могут угадать правильный ответ, исходя из формулировки задания, не обладая необходимыми знаниями. Такая ситуация может возникнуть, например, при неграмотном подборе дистракторов (вариантов ответа) на задание закрытого типа. Зависимость подчиняется следующей формуле:

$$p_{ij} = c_j + (1 - c_j) \frac{1}{1 + \exp(-d_j(\theta_i - \delta_j))}, \quad (3)$$

где c_j – вероятность угадывания.

На самом деле, каждая из моделей имеет свои достоинства и недостатки. Классический подход позволяет рассчитать больше всего параметров, но цен-

ность таких оценок не очень высокая, так результат представлен только в порядковой шкале. Модель Раша гарантирует линейность оценки сложности задания, но не позволяет оценить коэффициент дискриминации и вероятность угадывания. Переход к моделям Бирнбаума решает эту проблему, но лишает модель «линейности» [2].

Независимо от выбранной модели, обработка требует довольно глубокого изучения самих моделей, а также принципов взаимодействия различных параметров друг на друга. Поэтому мы считаем целесообразным создание экспертной системы, которая, с одной стороны, использовала бы совокупность моделей для расчета различных параметров, а с другой – давала бы понятные рекомендации по качеству теста и тестовых заданий [1] для автора на естественном языке.

Экспертная система – это сложный программный комплекс, аккумулирующий знания специалистов в конкретных предметных областях и тиражирующие этот эмпирический опыт для консультаций менее квалифицированных пользователей [3]. Обобщенная структура экспертной системы представлена на рис. 1.

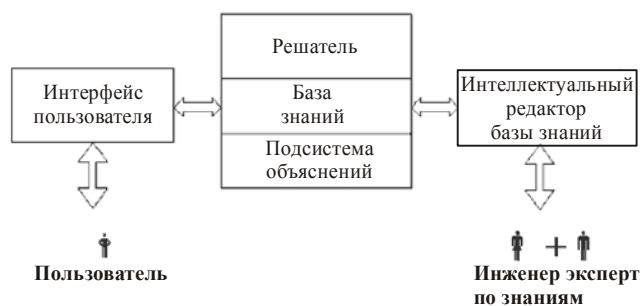


Рис. 1. Обобщенная структура экспертной системы

Рассмотрим отдельные блоки экспертной системы, а также определим, что будет выполняться в каждом блоке непосредственно в предполагаемой нами экспертной системе.

Интерфейс пользователя – комплекс программ, реализующих диалог пользователя с экспертной системой как на стадии ввода информации, так и получения результатов [3]. В этом блоке предполагается оболочка для ввода и хранения входных данных, из которых непосредственно внутри оболочки будут вычисляться необходимые коэффициенты по различным моделям оценки результатов тестирования, а также некоторые дополнительные параметры, такие, как сведения об испытуемых (например: факультет, специальность и курс, если испытуемые студенты) и сведения о тесте (дисциплина, по которой проводится тестирование и тип тестирования (педагогическое или психологическое)). Этот блок также предполагает вывод выходных данных, т.е. рекомендации по тесту и тестовым заданиям.

База знаний – ядро экспертной системы, совокупность знаний предметной области, записанная на машинный носитель в форме, понятной эксперту и пользователю (обычно на некотором языке, приближенном к естественному) [3]. В данном случае база

знаний будет содержать набор правил, по которым будут оцениваться результаты тестов.

Решатель – программа, моделирующая ход рассуждений эксперта на основании знаний, имеющихся в базе знаний [3]. Данный блок будет строить рекомендации по тесту исходя из правил, хранящихся в базе знаний.

Подсистема объяснений – программа, позволяющая пользователю получить ответы на вопросы: «Как была получена та или иная рекомендация?» и «Почему система приняла такое решение?» [3].

Интеллектуальный редактор базы знаний – программа, представляющая инженеру по знаниям возможность создавать БЗ в диалоговом режиме [3]. Посредством данного блока будет осуществляться редактирование правил для более тонкой настройки и усовершенствования экспертной системы.

Все экспертные системы классифицируются по нескольким критериям. Критерии классификации экспертных систем представлены на рис. 2.



Рис. 2. Классификация экспертных систем

Рассмотрим подробнее пункты классификации и определим, к какому пункту относится наша экспертная система по каждому критерию.

Классификация по типу решаемой задачи

– **Интерпретация данных.** Это одна из традиционных задач для экспертных систем. Под интерпретацией понимается определение смысла данных, результаты которого должны быть согласованными и корректными. Обычно предусматривается многовариантный анализ данных.

– **Диагностика.** Под диагностикой понимается обнаружение неисправности в некоторой системе. Неисправность – это отклонение от нормы. Такая трактовка позволяет с единых теоретических позиций рассматривать и неисправность оборудования в технических системах, и заболевания живых организмов, и всевозможные природные аномалии. Важной спецификой является необходимость понимания функциональной структуры («анатомии») диагностирующей системы.

– **Мониторинг.** Основная задача мониторинга – непрерывная интерпретация данных в реальном масштабе времени и сигнализация о выходе тех или иных параметров за допустимые пределы. Главные проблемы – «пропуск» тревожной ситуации и инверсная задача «ложного» срабатывания. Сложность этих проблем в размытости симптомов тревожных ситуаций и необходимость учета временного контекста.

– **Проектирование.** Проектирование состоит в подготовке спецификаций на создание «объектов» с заранее определенными свойствами. Под спецификацией понимается весь набор необходимых документов, чертеж, пояснительная записка и т.д. Основные проблемы здесь – получение четкого структурного описания знаний об объекте и проблема «следа». Для организации эффективного проектирования и, в еще большей степени, перепроектирования необходимо формировать не только сами проектные решения, но и мотивы их принятия. Таким образом, в задачах проектирования тесно связываются два основных процесса, выполняемых в рамках соответствующей экспертной системы: процесс вывода решения и процесс объяснения.

– **Прогнозирование.** Прогнозирующие системы логически выводят вероятные следствия из заданных ситуаций. В прогнозирующей системе обычно используется параметрическая динамическая модель, в которой значения параметров «подгоняются» под заданную ситуацию. Выводимые из этой модели следствия составляют основу для прогнозов с вероятностными оценками.

– **Планирование.** Под планированием понимается нахождение планов действий, относящихся к объектам, способным выполнять некоторые функции. В таких экспертных системах используются модели поведения реальных объектов с тем, чтобы логически вывести последствия планируемой деятельности.

– **Обучение.** Системы обучения диагностируют ошибки при изучении какой-либо дисциплины с помощью ЭВМ и подсказывают правильные решения. Они аккумулируют знания о гипотетическом «ученике» и его характерных ошибках, затем в работе способны диагностировать слабости в знаниях обучаемых и находить соответствующие средства для их ликвидации. Кроме того, они планируют акт общения с учеником в зависимости от успехов ученика с целью передачи знаний.

Экспертные системы (ЭС) также различают по принципу **поддержки временных параметров**:

– **Статические ЭС** разрабатываются в предметных областях, в которых база знаний и интерпретируемые данные не меняются во времени. Они стабильны.

– **Квазидинамические ЭС** интерпретируют ситуацию, которая меняется с некоторым фиксированным интервалом времени.

– **Динамические ЭС** работают в сопряжении с датчиками объектов в режиме реального времени с непрерывной интерпретацией поступаемых данных.

По типу интеграции экспертной системы с программными продуктами других типов различают:

– **Автономные ЭС** работают непосредственно в режиме консультаций с пользователем для специфически «экспертных» задач, для решения которых не требуется привлекать традиционные методы обработки данных (расчеты, моделирование и т. д.).

– **Гибридные ЭС** представляют программный комплекс, агрегирующий стандартные пакеты прикладных программ (например, математическую статистику, линейное программирование или системы управления базами данных) и средства манипулирования знаниями. Это может быть интеллектуальная надстройка над ППП или интегрированная среда для решения сложной задачи с элементами экспертных знаний [3].

Нашу экспертную систему можно типизировать следующим образом. По решаемой задаче – это система интерпретации данных, потому как в основе ле-

жит интерпретация данных о результатах тестирования на основании статистических моделей. Система является квазидинамической, так как база данных результатов тестирования (матрица тестирования) периодически заполняется (во время каждого сеанса тестирования). По степени интеграции с другими программами система является гибридной, так как помимо выработки рекомендаций по тестам необходимо в рамках поставленной задачи определять множество параметров по различным моделям анализа результатов тестирований. Следовательно, конечный продукт предполагает совместную работу ЭС, математического блока расчета параметров и средства управления базами данных.

Такая ЭС позволит преподавателям без дополнительных знаний тестологии создавать качественные тесты на основе рекомендаций, генерируемых в автоматическом режиме.

Литература

1. Минин М. Г. Диагностика качества знаний и компьютерные технологии обучения. Томск: Изд-во ТГПУ, 2000. 216 с.
2. Маслак А. А. Измерение латентных переменных в социально-экономических системах. Славянск-на-Кубани: Изд-во СГПИ, 2006. 333 с.
3. Экспертные системы: структура и классификация. <http://www.ssti.ru/>

Поступила в редакцию 25.12.2008

УДК 612. 821.3: 51-76

Ю. В. Бушов, М. В. Светлик, Е. П. Крутенкова

КОРРЕЛЯЦИЯ ИНТЕЛЛЕКТА И ТОЧНОСТИ ВОСПРИЯТИЯ ВРЕМЕНИ С ВЫСОКОЧАСТОТНОЙ ЭЛЕКТРИЧЕСКОЙ АКТИВНОСТЬЮ МОЗГА

Томский государственный университет

В последние годы пристальное внимание исследователей привлекает высокочастотная электрическая активность мозга, так называемый гамма-ритм. Частота этого ритма варьирует по данным разных авторов [1–3] от 30 до 80, 200 Гц и более, а амплитуда не превосходит 5–10 мкВ. Присутствие этого ритма обнаружено не только у человека, но и у животных [4]. Имеются данные о том, что основную роль в электрогенезе гамма-ритма частотой от 30 до 80 Гц играют постсинаптические потенциалы, а колебания более высокой частоты являются отражением суммарной синхронизированной импульсной активности нейронов [4]. Согласно популярной в настоящее время гипотезе, гамма-ритм у человека играет ключевую роль в обеспечении когнитивных процессов. Это подтверждают полученные в последнее время данные о связи гамма-ритма с процессами восприятия [5], внимания [1], сознания [6] и обработки семантической информации [7]. Имеются данные о том, что ампли-

туда и частота этого ритма зависят от состояния человека и вида выполняемой когнитивной задачи [8, 9]. Поскольку частотные параметры гамма-ритма близки к нейронной активности, полагают, что он отражает активность нейронных сетей. Считают [6], что именно на частоте гамма-ритма происходит синхронизация активности и функциональное объединение пространственно удаленных популяций нейронов при осуществлении сознательной деятельности. Все это позволяет предположить наличие определенной связи между активностью головного мозга на частоте гамма-ритма, с одной стороны, и показателями интеллекта и успешностью выполнения различных видов интеллектуальной деятельности – с другой. Вместе с тем анализ литературы показал, что специальные исследования в этом направлении фактически не проводились. В связи с этим, целью настоящего исследования явилось изучение взаимосвязи показателей интеллекта и точности восприятия времени с