

Современные системы компьютерного тестирования

Нардюжев Виктор Иванович — доцент кафедры информационно-компьютерных дисциплин Обнинского института атомной энергетики, кандидат технических наук. E-mail: vin@kaluga.ru
Нардюжев Иван Викторович — доцент кафедры автоматизированных систем управления Обнинского института атомной энергетики, кандидат технических наук. E-mail: vin@kaluga.ru

Анализ функций, технической структуры и программного обеспечения современных систем компьютерного тестирования

Проведём анализ, рассматривая основные задачи современных систем компьютерного тестирования.*

* Нардюжев В.И., Нардюжев И.В. Тестирование на компьютерах через Интернет//Труды Центра тестирования. Вып. 2. М.: Прометей, 1999. С. 139–157.

Создание и хранение тестовых заданий

Задача создания и хранения тестовых заданий уже давно решается с помощью компьютеров. Даже в системах, ориентированных исключительно на бланковое тестирование, ведение банка тестовых заданий на компьютере считается необходимым*, так как позволяет:

- обрабатывать исходный авторский материал;
- вносить необходимые изменения и дополнения;
- назначать и поддерживать реквизиты тестовых заданий;
- выполнять автоматизированную компоновку тестов;
- проводить анализ результатов тестирования;
- хранить и передавать информацию.

* Хлебников В.А., Сальников Н.Л., Нардюжев В.И., Нардюжев И.В. Автоматизированная система государственного централизованного тестирования// Проблемы информатизации высшей школы. Бюллетень 3–4 (9–10) 1997. М.: ГосНИИ системной интеграции, 1998. С. 30– 37.

Для эффективного решения задачи создания и хранения тестовых заданий в системе компьютерного тестирования должны быть реализованы следующие базовые функции:

а) Работа с разными типами тестовых заданий:

- с множественным выбором;
- с множественным ответом;
- с дополнением;
- со свободным ответом;
- на установление соответствия;
- на выявление последовательности.

б) Назначение весовых коэффициентов заданиям и ответам.

в) Проверка орфографии, поиск по ключевым словам и реквизитам.

г) Использование в тестовых заданиях форматированного текста, графики, звука, видеофрагментов, активных программных элементов.

д) Подготовка шаблонов, импорт/экспорт содержимого тестовых заданий.

е) Архивация и восстановление необходимой информации.

ж) Формирование контрольной распечатки банка.

Поддержка различных типов тестовых заданий усложняет программное обеспечение, тем не менее большинство систем компьютерного тестирования общего назначения реализуют эту функцию, используя элементы графического интерфейса операционной системы и учитывая ограничения выбранного формата. Системы, рассчитанные на конкретные тесты, поддерживают лишь необходимые им типы заданий (например, только задания с множественным вы-

бором).

Функция назначения весовых коэффициентов заданиям и ответам (например, установка уровня трудности и присуждение штрафа) реализуется либо на уровне эксперта, либо в процессе статистического анализа результатов тестирования и расчёта параметров тестовых заданий. При отсутствии этой функции значительно сужается область применения системы компьютерного тестирования, так как становится невозможным её использование для разработки высококачественных тестов.

Проверка орфографии помогает разработчикам тестов избежать досадных ошибок и опечаток, однако лишь немногие системы из-за трудности реализации предлагают её в качестве встроенной функции (внешней считается проверка в отдельном текстовом редакторе на этапе подготовки задания).

Поиск по ключевым словам и реквизитам необходим при работе с крупными банками тестовых заданий, поэтому он есть в большинстве систем компьютерного тестирования общего назначения, при этом используются как внутренние механизмы и алгоритмы самой системы, так и стандартные сервисы операционной системы или офисного пакета.

Преимуществом компьютерных тестов является возможность использования графики, анимации, звука, видеофрагментов и активных элементов, так как уже сегодня важную роль в жизни тестируемого играют музыка и кинематограф, радио и телевидение, компьютеры и быстро растущая глобальная сеть Интернет. Прототипы мультимедийных тестов появились достаточно давно, однако теперь речь идёт о промышленных объёмах, когда число тестируемых достигает миллионов*.

* Bennett R.E. Reinventing Assessment: Speculations on the Future of Large-Scale Educational Testing. Princeton, NJ: Educational Testing Service, 1997. College Board 1997–98 Annual Report. New York, NY: College Entrance Examination Board, 1998. Educational Testing Service 1998 Annual Report. Princeton, NJ: Educational Testing Service, 1998.

Мультимедийные элементы, как правило, создаются с помощью отдельного программного обеспечения на специальном оборудовании и поступают в систему компьютерного тестирования уже в виде готовых файлов. Полученные файлы можно привязывать к конкретным тестовым заданиям и использовать либо через стандартные средства операционной системы, либо через дополнительные программные компоненты. Единственной проблемой при этом остаётся большой размер файлов. В разной степени поддержка мультимедийных элементов реализована во всех системах компьютерного тестирования.

Процесс разработки тестового задания является достаточно трудоёмким, поэтому особое значение получают механизмы сокращения трудозатрат, такие, как использование шаблонов и импорт/экспорт содержимого разработанных ранее (возможно, в другой системе) тестовых заданий. Из-за отсутствия однозначного соответствия между форматами большинство систем готовы импортировать лишь отдельные элементы (тексты, картинки, звуки, видеофрагменты), но не целые задания.

Аналогичные трудности возникают при попытке преобразовать готовые тесты в формат HTML для представления через Интернет: внешний облик тестовых заданий значительно меняется, а часть функционала просто теряется. Поэтому в выгодном положении оказываются системы, изначально работающие в формате HTML и учитывающие его ограниченные возможности.

Функции архивации и восстановления банка тестовых заданий реализованы во многих современных системах компьютерного тестирования. Большой объём ценной информации и её возможная распределённость по дискам локальной сети затрудняют применение простых внешних механизмов типа: регулярная упаковка всего массива данных. Поэтому используются внутренние процедуры, отслеживающие историю изменения банка тестовых заданий и способные выполнить возврат к предыдущему состоянию в случае проведения восстановительных работ.

Операция формирования контрольной распечатки банка тестовых заданий характерна для систем бланкового тестирования, в которых авторы тестовых материалов не имеют

непосредственного доступа к электронному архиву*.

* Хлебников В.А., Сальников Н.Л., Нардюжев В.И., Нардюжев И.В. Автоматизированная система государственного централизованного тестирования // Проблемы информатизации высшей школы. Бюллетень 3–4 (9–10) 1997. М.: ГосНИИ системной интеграции, 1998. С. 30–37.

Такой подход является крайне неэффективным (требуются значительные временные и материальные затраты, актуальность бумажного варианта теряется при первом же изменении в электронном аналоге, любые правки делаются дважды — сначала на бумаге, потом на компьютере), однако он сохранился в некоторых системах компьютерного тестирования.

Завершая анализ функций, связанных с созданием и хранением тестовых заданий, отметим необходимость их тесной интеграции с такими подсистемами, как компоновка тестов и анализ статистической информации.

Компоновка тестов

Современные компьютерные тесты лишь автоматизируют существующий процесс. Они аналогичны предлагаемому на бумаге: используют те же модели поведения, измеряют те же характеристики, основаны на тех же типах заданий. В ряде случаев остаётся спорной их экономическая эффективность. Тем не менее они необходимы, чтобы создать базу для движения вперёд, для перехода к новым тестам и средствам разработки*.

* Hawkest T. Experiment in Computer-Assisted Assessment at Warwick Mathematics Institute // Interactions Journal, v. 2, № 3, 1998. Bennett R.E. Reinventing Assessment: Speculations on the Future of Large-Scale Educational Testing. Princeton, NJ: Education Testing Service, 1997.

Первоначально это будут экспертные системы, которые по указанным параметрам заданий будут предлагать наилучший (с точки зрения теории) шаблон. В определённый момент инструментарий разработчика станет столь изощрённым, что сможет формировать задания и тесты самостоятельно по ходу тестирования*.

* Там же.

Заметим, что сегодня процесс создания тестов достаточно трудоёмок*.

* Castle R.A. The Relative Efficiency of Two-Stage Testing Versus Traditional Multiple Choice Testing Using Item Response Theory in Licensure. Lincoln, NE: University of Nebraska, 1997.

А если учесть важность решений, принимаемых по результатам тестов, число специалистов высокой квалификации, задействованных в их разработке, необходимость непрерывного обновления и контроля за качеством тестовых материалов, то можно прийти к выводу, что «определённый момент» ещё очень далёк.

В современных системах компьютерного тестирования для компоновки тестов реализуются следующие функции:

- a) Отбор заданий из банка по указанному критерию.
- b) Компоновка различных тестов:
 - традиционных;
 - со случайной перестановкой заданий;
 - со случайной перестановкой ответов;
 - адаптивных.
- c) Установка проходного балла.
- d) Печать скомпонованных тестов.

При отборе заданий из банка в качестве критерия могут выступать: название предмета, номер темы, имя автора, дата поступления в архив, частота пользования, уровень трудности и т.д. Над отобранными заданиями можно выполнять различные операции: удаление из банка, передача на экспертизу, помещение в тест и т.п.

Программная реализация функции отбора зависит от принятой в системе схемы хранения самих тестовых заданий и их реквизитов:

- *объектно-ориентированная* — тестовое задание лежит в отдельном файле, содержащем все необходимые реквизиты;
- *реляционная* — реквизиты находятся в специальной таблице и связаны с тестовыми заданиями по ключевому полю;
- *табличная* — тестовое задание и его реквизиты занимают одну строку, но разные поля таблицы.

Объектно-ориентированная схема является логически ясной, но очень медленной, так как предполагает перебор большого количества отдельных файлов. Реляционный подход быстрее, но требует специальных механизмов для поддержания целостности связей. Табличный вариант самый быстрый, но наименее гибкий, поскольку вся информация должна храниться в одной громадной таблице.

Современные системы компьютерного тестирования сильно различаются по методам компоновки тестов. Традиционный метод предполагает, что разработчик самостоятельно расставляет задания по тесту, опираясь на их реквизиты. При этом формируется либо новый автономный документ, либо набор ссылок на задействованные задания. Параллельно составляется таблица ключей и весовых коэффициентов.

Для проведения крупномасштабного тестирования требуется большое количество тестов, близких по характеристикам. В этом случае используют метод случайной перестановки типовых тестовых заданий. Чем больше заданий, тем реже они повторяются в разных тестах. Важно отметить следующие особенности метода:

- чтобы не нарушать содержательную структуру теста, нужно сохранять своё положение, меняясь только с ему аналогичными;
- перестановка может выполняться как на этапе подготовки теста, так и в ходе тестирования (в бланковом варианте это было невозможно);
- при расчёте статистических характеристик необходимо учитывать факт использования задания в разных тестах.

Для повышения секретности дополнительно можно применить случайную перестановку ответов, трудность реализации которой зависит от формата хранения тестовых заданий. В любом случае необходимо фиксировать все выполненные перестановки. Многие системы делают это автоматически.

Особенно популярным в последнее время становится компьютерное адаптивное тестирование (Computerized Adaptive Testing — CAT), при котором компьютер отбирает задания в зависимости от предыдущих ответов, подстраивая тест к уровню тестируемого. Компьютерные адаптивные тесты считают самым значительным практическим достижением образовательного тестирования за последние два десятилетия. Они улучшают мотивацию, сокращают время тестирования, требуют меньше заданий для каждого экзаменуемого, при этом не снижают точность измерения*.

* Mislavy R.I. Test Theory Reconcived. CSE Technical Report, № 376. Los Angeles, CA: CRESST/ Education Testing Service, 1994. Mislavy R.I. Almond R.G. Graphical Models and Computerized Adaptive Testing. CSE Technical Report, № 434. Los Angeles, CA: CRESST/ Education Testing Service, 1997.

Основой для современных компьютерных адаптивных тестов является теория ответа на задание (Item Response Theory— IRT), учитывающая уровень тестируемого и трудность задания. Несмотря на упрощённую природу и строгие требования к независимости параметров, пользователи модели IRT-CAT разработали методики, гарантирующие её применимость в практических приложениях.

Отметим, что адаптивные тесты komponуются в процессе тестирования, нуждаются в большом банке тестовых заданий и требуют сложных расчётов.

Функция установки проходного балла реализована в большинстве систем компьютерного тестирования. Она полезна при подходе типа «сдал/не сдал». Некоторые сложности возникают, когда тестирование проходит по нескольким тестам разной суммарной трудности. В этом случае необходимо использовать дополнительные компенсирующие коэффициенты.

Печать скомпонованных тестов выполняется при параллельной работе бланкового и компьютерного вариантов тестирования, например, для проведения сравнительных исследований. Понятно, что возможности этой функции ограничены и не распространяются на мультимедийные и адаптивные тесты.

Подводя итоги, нужно сказать о тесной связи функций компоновки тестов с подсистемами создания и хранения тестовых заданий, проведения сеансов тестирования, обработки ответов тестируемых, сбора и анализа статистической информации, передачи данных по каналам связи и обеспечения секретности.

Регистрация и подготовка тестируемых

Можно выделить три основных фактора, влияющих на реализацию функции регистрации тестируемых:

- время и место регистрации;
- объём интересующей информации;
- способ передачи данных.

Регистрацию можно провести предварительно или непосредственно перед тестированием. Первый вариант позволяет использовать для регистрации выделенные помещения, отдельную технику и специальный персонал, что существенно повышает эффективность всей работы. По второму варианту тестируемому надо самостоятельно заполнить регистрационную форму на экране компьютера. В любом случае, получив необходимые сведения, система выдаёт идентификационный номер, по которому осуществляется привязка регистрационной информации к данным о сеансе тестирования.

Какие сведения об участниках интересуют организаторов, зависит от целей тестирования. Например, в тестовом самоконтроле регистрация вообще нужна, а при тестировании в масштабах страны с анализом влияющих факторов, наоборот, требуется полноценная анкета. В такой ситуации в современной системе компьютерного тестирования общего назначения принимается одна из возможных схем:

- присвоить каждому сеансу тестирования индивидуальный номер для использования в отдельной информационной системе учёта участников;
- перед присвоением индивидуального номера запросить минимальный набор данных об участнике, а отдельную информационную систему учёта вести лишь при необходимости;
- предоставить средство для разработки нужных регистрационных форм. В последнее время широкое распространение получают новые способы передачи регистрационных данных: по телефону, факсу, электронной почте, через Web-страницу. Такие возможности в первую очередь необходимы организаторам крупномасштабного тестирования, число участников которого может составить несколько миллионов человек в год*.

* College Board 1997–98 Annual Report. New York, NY: College Entrance Examination Board, 1998. Educational Testing Service 1998 Annual Report. Princeton, NJ: Education Testing Service, 1998.

Помимо регистрации важным вопросом является подготовка тестируемых к прохождению компьютерных тестов:

- выпуск информационных бюллетеней, отвечающих на вопросы: «Как проходит тестирование?», «Из каких разделов и заданий состоит тест?», «Каковы временные ограничения?», «Какой стратегии следует придерживаться?», «Каким образом подсчитываются баллы?», «Как лучше подготовиться?», «Где зарегистрироваться?», «Как узнать свой результат?» и т.д.;
- распространение электронных курсов обучения основным приёмам работы за персональным компьютером, необходимым для выполнения типовых тестовых заданий;
- проведение репетиционного тестирования в обстановке, максимально приближенной к сдаче реального компьютерного теста.

Учитывая важность подготовки тестируемых, некоторые современные системы компьютерного тестирования позволяют сформировать специальные репетиционные тесты, которые содержат элементы обучения, реагируют на правильные ответы, разъясняя суть ошибки, и

имеют менее жёсткие временные ограничения.

Проведение сеансов тестирования на компьютерах

Проведение сеансов тестирования на компьютерах стало возможным благодаря многолетним крупномасштабным исследованиям самых разных аспектов перехода от традиционных бумажных тестов к компьютерным. Например, одно из таких исследований*, проведённое по заказу ETS (Educational Testing Service — Службы тестирования в образовании), показало отсутствие зависимости между компьютерной грамотностью и результатами выполнения компьютерного варианта TOEFL (Test Of English as a Foreign Language — Тест по английскому как иностранному) после прохождения электронного курса обучения основным приемам работы за компьютером.

* Taylor C., Jamieson J., Eignor D., Kirsch I. The Relationship Between Computer Familiarity and Performance on Computer-based TOEFL Test Tasks. TOEFL Research Rep., № 61. Princeton, NJ: Educational Testing Service, 1998.

В рамках другого исследования* изучалось влияние вариантов тестирования (бланкового и компьютерного) на выполнение учащимися средней школы тестовых заданий (с множественным выбором и письменных). Компьютерный вариант теста специально делали максимально близким к бумажному оригиналу. Экспериментальная группа, работавшая больше за компьютерами, значительно лучше выполнила задания, предполагающие развёрнутый письменный ответ, показав на заданиях с множественным выбором результаты, аналогичные тем, что получила контрольная группа.

* Russell M., Haney W. Testing Writing on Computers: An Experiment Comparing Student Performance on Tests Conducted via Computer and via Paper-and-Pencil// Education Policy Analysis Archives, v. 5, № 3, 1997.

На основе полученных данных был сделан вывод, что всё большее число учащихся привыкает к письму за компьютером, а крупномасштабные исследования по-прежнему ориентированы на использование бумажных тестов, что приводит к недооценке реальных способностей учеников. Это при том, что уже в 1993 году по информации NCEES (National Center for Educational Statistics— Национальный центр статистики в образовании США) 68,9% учащихся использовали компьютер в школе, а 45,8% служащих — на рабочем месте. По данным анкетирования*, проведённого среди 90 тыс. человек, сдававших TOEFL в апреле 1996 года, высокую степень знакомства с компьютером продемонстрировали 50% опрошенных, среднюю — 34%, низкую — лишь 16%. Таким образом, компьютерное тестирование становится не только выгодной альтернативой, но и реальной необходимостью.

* Eignor D., Taylor C., Kirsch I., Jamieson J. Development of a Scale for Assessing the Level of Computer Familiarity of TOEFL Examinees. TOEFL Research Rep., № 60. Princeton, NJ: Educational Testing Service, 1998.
Kirsch I., Jamieson J., Taylor C., Eignor D. Computer Familiarity Among TOEFL Examinees. TOEFL Research Report, № 59. Princeton, NJ: Educational Testing Service, 1998.

Стремительное расширение круга пользователей заставляет разработчиков систем компьютерного тестирования постоянно совершенствовать функции, связанные с проведением сеансов тестирования на компьютерах:

- a) Предъявление теста
 - с дискеты;
 - с жёсткого диска автономного компьютера;
 - через локальную сеть;
 - через Интернет.
- b) Начало, приостановка, продолжение и завершение сеанса тестирования.
- c) Показ отведённого, истраченного и оставшегося времени.
- d) Отображение количества заданий в тесте.
- e) Работа с тестовыми заданиями разных типов.
- f) Оперативная реакция на выбранный ответ.

g) Пропуск отдельных заданий.

Вариант предъявления теста с дискеты привлекателен своей простотой, но имеет существенные ограничения на пространство, выделяемое для хранения ведущей программы, тестовых заданий и ответов тестируемого, что затрудняет использование сложного (а значит, объёмного) программного обеспечения и мультимедийных элементов (особенно звуковых и видеофрагментов).

Большая ёмкость жёсткого диска автономного компьютера снимает ограничения дискетного варианта. Однако в этом случае пользователю потребуются дополнительные носители информации (набор дискет или компактный диск) и отдельная процедура, чтобы установить необходимые компоненты теста.

При работе в классе установка и обслуживание теста на каждом компьютере требуют значительных трудозатрат, для сокращения которых имеет смысл объединить компьютеры в локальную сеть и разместить программное обеспечение и тестовые задания на общедоступном диске файлового сервера. Работоспособность указанной схемы зависит от пропускной способности сети и производительности сервера.

Вариант предъявления теста через Интернет допускает множество реализаций с разными ответами на ключевые вопросы:

- Нужно ли клиенту до начала тестирования получать с сервера и устанавливать на своём компьютере дополнительное программное обеспечение?
- Необходимые тестовые задания загружаются предварительно или по ходу тестирования?
- Обработка ответов и подсчёт набранных баллов производится на локальном компьютере или на удалённом сервере?

Получение дополнительного программного обеспечения и его установка значительно усложняют процедуру тестирования, но могут существенно расширить возможности базовой операционной системы и стандартного просмотрщика Web-страниц. При этом критическими являются следующие параметры:

- объём программного модуля, получаемого с сервера;
- пропускная способность каналов связи;
- количество клиентов.

Предварительная загрузка тестовых заданий позволяет пользователю:

- выбрать удобное время, когда всё работает быстрее и стоит дешевле;
- получить необходимые материалы, отключиться от глобальной сети и спокойно работать над тестом, не беспокоясь об оплате каждой минуты подключения;
- избавиться от томительного ожидания появления на экране следующего тестового задания;
- совместно и многократно использовать (например, в локальной сети) загруженные ранее файлы.

В свою очередь загрузка тестовых заданий по ходу тестирования:

- даёт возможность быстрее начать работу и оценить необходимость её продолжения (например, в случае репетиционного тестирования);
- является единственно возможной, когда заранее неизвестно, какое задание будет следующим (например, при адаптивном тестировании);
- уменьшает объём передаваемой информации и оплату графика, загружая лишь реально используемые задания и программные компоненты.

Место обработки ответов и подсчёта набранных баллов определяется:

- целями тестирования (самоконтроль или аттестация);
- сложностью процедуры (например, для оценки неформального задания может потребоваться мнение эксперта);
- необходимостью централизации данных (например, для анализа результатов апробации новых тестовых материалов).

Завершая анализ варианта предъявления теста через Интернет, отметим его перспек-

тивность. Доказательство — развитие глобальной сети в последние годы: увеличение капиталовложений, привлечение ведущих разработчиков аппаратных средств и программного обеспечения, быстрый рост числа конечных пользователей, наращивание пропускной способности каналов, формирование стандартов, проведение новых исследований и т.д.

При любом способе предъявления теста необходимо реализовать функции начала, приостановки, продолжения и завершения сеанса тестирования. Чтобы начать работу с тестом, пользователь должен указать свой идентификационный номер, полученный в момент регистрации, а система должна его проверить.

Выполнение теста в любой момент может быть приостановлено:

- при завершении одного из разделов;
- по желанию тестируемого с разрешения инструктора;
- по причине сбоя аппаратуры или программного обеспечения.

Для нормального продолжения работы с тестом необходимо знать состояние сеанса в момент приостановки: перечень выполненных заданий, данные на них ответы, истраченное время и т.д. Непредсказуемость аварийных приостановок заставляет систему периодически сохранять состояния всех активных сеансов на надёжном носителе информации. При этом отдельным вопросом является правильный выбор периода — частое сохранение приводит к снижению производительности системы, редкое — может стать причиной потери важных промежуточных данных.

Завершение сеанса тестирования может быть добровольным (по желанию тестируемого, с разрешения инструктора) или принудительным (например, если закончилось отведённое время). В любом случае возврат к работе над тестом, как правило, невозможен. Получив сигнал о завершении сеанса, система либо на месте подсчитывает и сообщает набранные баллы, если это возможно, либо формирует протокол тестирования для отправки на обработку удалённому серверу.

Во многих тестах ключевой является продолжительность работы. Поэтому в большинстве современных систем компьютерного тестирования реализованы показ, учёт и контроль отведённого, истраченного и оставшегося времени. Временные ограничения могут быть назначены не только всему тесту, но и отдельным разделам или заданиям.

Для планирования своего времени тестируемому полезно знать количество заданий и процент выполненной работы, но эти показатели не всегда известны самой системе. Например, при адаптивном тестировании число предъявленных заданий зависит от стабильности получаемых ответов.

Работа с тестовыми заданиями разных типов требует наличия достаточных вычислительных ресурсов и необходимой программной поддержки на каждом компьютере, за которым проходят сеансы тестирования. Поэтому ещё на этапе создания тестов нужно учесть реальные возможности конечных пользователей по установке и обслуживанию сложных программно-технических комплексов.

Наличие оперативной реакции на выбранный ответ зависит от её

- необходимости (она требуется при самоконтроле или тренировке, но не нужна для аттестации);
- возможности (система сумеет прокомментировать каждый вариант при множественном выборе, но не способна самостоятельно быстро оценить эссе, предназначенное для проверки экспертами). Допустимость выполнения тестовых заданий в произвольном порядке определяется авторами теста с учётом его типа. Например, пропуск отдельных заданий и возврат к ним позже могут быть запрещены, когда
 - тестовые задания специально отсортированы по уровню трудности для определения максимальных возможностей тестируемого;
 - есть достаточно автономные разделы (иногда называемые *субтестами*) со своими правилами и временными ограничениями;
 - используется алгоритм адаптивного тестирования, при котором каждое следующее задание выбирается в зависимости от ответа на текущее.

Рассмотрев функции, связанные с проведением сеансов тестирования на компьютерах, можно сделать вывод о наличии большого числа возможных вариантов реализации технической структуры и программного обеспечения современной системы компьютерного тестирования.

Обработка ответов и подсчёт набранных баллов

Процедура обработки ответов и подсчёта набранных баллов получает информацию из подсистем компоновки тестов и проведения сеансов тестирования, поэтому неудивительно, что некоторые её особенности уже обсуждались выше.

В системе компьютерного тестирования обработка ответов может быть:

- локальная или удалённая;
- формальная или экспертная;
- оперативная или отложенная.

Обработка считается *удалённой*, если она выполняется за пределами места проведения сеансов тестирования, а для передачи всех необходимых данных удалённому серверу используются распределённые вычислительные сети или носители информации, подлежащие физической транспортировке.

Обработка становится *экспертной*, когда невозможно простое сравнение с ключами и требуется привлечение специалистов. Например, в одном из разделов теста по русскому языку как иностранному на вопрос задания тестируемый должен дать развёрнутый устный ответ*.

* Нардюжев В.И., Нардюжев И.В. HTML-тесты — новый вид тестов на компьютерах // Развитие системы тестирования в России. Всероссийская научно-практическая конференция: Тезисы докладов. Москва, 25–26 ноября 1999 г. Ч. 3. М.: Прометей, 1999. С. 42–44.

Обработка является *отложенной*, если нельзя показать её результаты сразу после завершения сеанса тестирования. Причиной может быть удалённость сервера, необходимость обращения к эксперту, сложность алгоритма расчёта набранных баллов и так далее.

Для самоконтроля типовой считается локальная формальная оперативная обработка ответов. Она наиболее простая в реализации, самая быстрая и очень удобная в использовании.

Крупномасштабное аттестационное тестирование ориентировано в первую очередь на удалённую (централизованную) формальную (без привлечения множества экспертов), отложенную (тщательную и комплексную) обработку ответов большого числа (до нескольких миллионов) тестируемых.

Любая обработка заканчивается подсчётом набранных баллов. В самом простом варианте за каждый правильный ответ даётся один балл. Более сложные схемы дополнительно учитывают уровни трудности тестовых заданий с помощью весовых коэффициентов и назначают штраф за грубые ошибки. Помимо абсолютного балла некоторые системы сообщают рейтинг, то есть указывают положение каждого тестируемого относительно других участников.

Сбор и анализ статистической информации

Сбор и анализ статистической информации обеспечивают обратную связь, без которой тестирование перестаёт быть научным методом контроля знаний и оценки учебных достижений.

Условно всю статистику можно разделить на две части:

- информация об участниках тестирования;
- данные о качестве тестовых материалов.

При работе с первой частью система компьютерного тестирования, как правило, позволяет:

- а) Выбрать объект статистической обработки.
- б) Для выбранного объекта получить в цифровом виде:
 - количество участников по категориям;
 - распределение тестируемых по числу набранных баллов;

- процент правильных ответов на каждое задание теста.
- с) По полученным цифровым данным построить графики.
- d) Произвести сравнительный анализ результатов тестирования различных объектов статобработки.
- e) Увидеть список тестируемых, с возможностью произвольной сортировки и быстрого поиска.
- f) Распечатать статистические формы и списки (или их части) на принтере.
- g) Экспортировать интересующий блок данных для анализа в специальном статистическом пакете.

При сборе данных о качестве тестовых материалов с помощью системы компьютерного тестирования чаще всего определяются:

- a) Минимальное, среднее, максимальное значение, стандартное отклонение и коэффициент вариации полученных тестовых баллов.
- b) Номера тестов, в которых использовалось задание.
- с) Количество тестируемых, выполнявших тест.
- d) Статистические параметры задания:
 - уровень трудности;
 - точно-бисериальная корреляция;
 - дифференцирующая способность.
- e) Работа дистракторов:
 - частота выбора ответов всеми (слабейшими и сильнейшими);
 - отклонения частот от среднего значения;
 - наличие отклонений от нормы.
- f) Независимость заданий в тесте.

Представленный выше перечень функций сбора и анализа статистической информации связан с расчётом базовых величин и даёт возможность автору или тестологу быстро перейти от полученной в отчёте цифры к находящемуся за ней тестовому заданию. Для определения комплексных характеристик, таких, как надёжность, валидность, эффективность тестов, используются специальные статистические пакеты.

Ключевым моментом является возможность быстрого получения нужных данных от подсистем создания и хранения тестовых заданий, компоновки тестов, регистрации и подготовки тестируемых, проведения сеансов тестирования на компьютерах, обработки ответов и подсчёта набранных баллов. Для её успешной реализации в некоторых системах используется метод поэтапной загрузки всей необходимой информации в единую реляционную базу данных с параллельным расчётом базовых величин.

Передача данных по каналам связи

В последние годы значительно меняются цели, технологии и научные основы тестирования в образовании. Конкуренция заставляет производителей повышать продуктивность и качество тестов, учитывать разнородность населения и новые потребности рынка, принимать на вооружение новейшие технологии и последние достижения науки.

Появилась возможность регистрироваться по телефону и на Web-сервере, платить по кредитной карточке, тестироваться в относительно небольшом и удобном центре, получать результаты в момент завершения сеанса.

Организаторы и центры тестирования теперь могут обмениваться вопросами теста и ответами тестируемого в электронном виде, аналогичным образом отправлять результаты теста заинтересованным учреждениям.

Ряд тестовых программ уже работают по такой схеме, а в ближайшие несколько лет их число значительно возрастёт. Поэтому сейчас исследуются самые разные аспекты перехода на новые технологии передачи данных.

Например, в штате Пенсильвания в 1997 году провели эксперимент*, чтобы ответить на вопросы: «Каковы отличия тестирования через Интернет от традиционных бумажных те-

стов?», «Ставит ли предоставление теста через Интернет какую-либо подгруппу в невыгодное положение?», «Как можно обеспечить информационную безопасность?», «Во что обойдётся реализация тестирования через Интернет на уровне штата?»

* Bicanich E., Slivinski T., Hardwicke S.B., Kapes J.T. Internet-Based Testing: A Vision or Reality? // T.H.E. Journal (Technological Horizons in Education), v. 25, № 12, 1997. P. 61–65.

В эксперименте приняли участие около 400 студентов из 14 учебных заведений. Использовались тесты, составленные из банка тестовых заданий с множественным выбором. Программное обеспечение для тестирования через Интернет предоставила фирма WebTester. С помощью контрольных групп рассчитывалась ретестовая эквивалентность — около 360 студентов получили как бумажные, так и предоставленные через Интернет наборы заданий в качестве тестов достижений в конце учебного курса.

По результатам эксперимента исследователи сделали вывод, что вариант предоставления теста не оказал влияния на показатели учащихся. 75% тестируемых высказались за тестирование через Интернет, несмотря на то, что 68,5% имели небольшой опыт работы в сети. Администраторы отметили, что новые технологии требуют меньше времени на подготовку и проведение тестирования, особенно на анализ результатов. Среди важных факторов успеха названы: способ подключения учебного заведения к сети, качество и цена предоставляемых поставщиком услуг передачи данных по каналам связи, возможности используемых компьютеров.

Приняв во внимание затраты на разработку новых тестов, проведение тестирования и обработку результатов (без учёта стоимости компьютеров, программного обеспечения и оплаты доступа к сети), исследователи определили для данного случая пороговое число участников, когда тестирование через Интернет будет экономически эффективнее бумажного — 375 человек.

Однако реализация проектов на уровне штата столкнётся с серьёзными трудностями (широкий спектр используемых аппаратных и программных средств, разнообразие схем подключения к сети, работа через множество независимых поставщиков услуг, отсутствие подготовленных кадров на местах, недостаток опыта в организации подобных мероприятий).

Помимо организационных трудностей реализации крупномасштабного тестирования через Интернет существуют чисто технические проблемы. Связаны они с пропускной способностью каналов, с мощностью серверов и рабочих станций, с оптимальностью алгоритмов передачи, приёма и обработки больших объёмов информации. Простое наращивание ресурсов негативно скажется на экономической эффективности применения новых технологий, поэтому специалисты ищут другие решения, часто прибегая к различным методам моделирования*.

* Abdulla G. Analysis and Modeling of World Wide Web Traffic. Blacksburg, VA: Virginia Polytechnic Institute and State University, 1998. Abrams M., Williams S., Abdulla G., Patel S., Ribles R., Fox E. Multimedia Traffic Analysis Using Chitra 95. Blacksburg, VA: Computer Science Department, Virginia Polytechnic Institute and State University, 1995. Brakmo L.S., Peterson L.L. Experiences with Network Simulation. University of Arizona, 1996. Heidemann I., Obraczka K., Touch I. Modeling the Performance of HTTP Over Several Transport Protocols. Los Angeles, CA: The University of Southern California, 1997.

Обеспечение секретности

Снижение цен и рост пропускной способности каналов позволят быстро и качественно передавать и обрабатывать большие объёмы информации, сделают возможным перевод в электронный формат всего технологического цикла: поиск интересующего вуза, регистрация заявки, получение материалов для подготовки, тренировка, тестирование, пересылка результатов, отправка свидетельств и образцов выполнения заданий теста, разрешение спорных вопросов и так далее.

Подобное развитие событий требует тщательной подготовки*. Учёт результатов тестирования в процессе принятия важных решений ведёт к увеличению числа попыток сдать эк-

замен нечестным путём: познакомиться с содержимым теста заранее или воспользоваться услугами «суфлёра».

* Cotton G.D. High-Tech Approaches to Breaching Examination Security. Clearwater, FL: Loss Prevention Schroeder Measurement Technologies, Inc., 1997.

Современные технологии позволяют за короткое время с помощью доступного по цене оборудования украсть содержимое любого теста с минимальным риском быть пойманным. Арсенал международного промышленного шпионажа доступен сегодня каждому. Это фотокамеры в наручных часах и зажигалках, устройства звукозаписи в пачке из-под сигарет, имеющие чувствительный микрофон и способные записывать 3 часа без остановки, видеокамеры в футляре для очков или на галстучке, устройства передачи видеосигнала на расстояние в 20 миль, портативные перехватчики электромагнитного излучения, ультрафиолетовые ручки и так далее.

В условиях хранения, представления, передачи и обработки тестовых материалов в электронном виде экзаменационный шпионаж становится реальной угрозой. Эксперты предлагают ряд контрмер*, среди которых — детекторы металла, источники шума для подавления возможной передачи, изоляция помещений, обучение персонала методам визуального контроля. Относительно простые меры безопасности могут сохранить значительные капиталовложения в разработку тестов.

* Cotton G.D. High-Tech Approaches to Breaching Examination Security. Clearwater, FL: Loss Prevention Schroeder Measurement Technologies, Inc., 1997.

Особое внимание вопросам информационной безопасности уделили специалисты ETS, разрабатывая в 1996 году типовой проект технического оснащения компьютерных классов для проведения тестирования: отдельный бокс с рабочей станцией для каждого тестируемого, специальная комната с параболическим стеклом для инструктора, установленные в классе видеокамеры и микрофоны, журнал регистрации удостоверений личности, изолированная локальная сеть и сервер, хранящий данные в зашифрованном виде*.

* Educational Testing Service, 1998 Annual Report. Princeton, NJ: Educational Testing Service, 1998.

Современные системы компьютерного тестирования вносят свой вклад в комплекс управленческих, организационных и программно-технических мер обеспечения информационной безопасности, реализуя базовые возможности:

а) Защищать паролями:

- банк тестовых заданий;
- скомпонованные тесты;
- используемые программы.

б) Шифровать:

- банк тестовых заданий;
- скомпонованные тесты;
- реквизиты и результаты тестируемых;
- информацию о сеансах тестирования;
- статистические данные.

с) При выполнении теста блокировать доступ к окружающей среде, например, в многопрограммной операционной системе.

В ближайшем будущем эксперты ожидают широкое распространение электронных средств обучения*, позволяющих перейти от разовых экзаменов в конце учебного курса к контролю знаний на всём его протяжении. При этом существенно повысится объективность оценки учебных достижений и станет практически невозможным какой-либо обман.

* Bennett R.E. Reinventing Assessment: Speculations on the Future of Large-Scale Educational Testing. Princeton, NJ: Educational Testing Service, 1997.

Классификация современных систем компьютерного тестирования

Проведённый анализ показал большое разнообразие современных систем компьютерного тестирования. Для унификации теоретических исследований прикладных методов необходимо разделить всё множество систем на классы*, выбрав основные свойства в качестве классификационных признаков.

* Майоров С.А., Новиков Г.И., Алиев Т.И., Махарев Э.И., Тимченко Б.Д. Основы теории вычислительных систем. М.: Высшая школа, 1978.

Назначение

Все системы компьютерного тестирования можно разделить на *проблемно-ориентированные* и *общего назначения*. Проблемно-ориентированные системы имеют конкретную область применения:

- международные сравнительные исследования;
- мониторинг качества образования в масштабах страны;
- лицензирование и государственная аккредитация учебных заведений;
- аттестация учащихся и студентов, учителей и преподавателей;
- проверка профессиональной пригодности специалистов;
- сертификация новых учебных программ и методик;
- самоконтроль учащихся в процессе обучения;
- подготовка к сдаче итоговых экзаменов;
- апробация тестовых материалов.

Системы общего назначения, наоборот, претендуют на успешную работу сразу в нескольких областях. Однако существенное различие в условиях эксплуатации, например, для мониторинга качества образования в масштабе всей страны и для самоконтроля отдельных учащихся в процессе обучения значительно затрудняет создание эффективной многоцелевой системы.

Поэтому ориентация на проблему характерна для крупномасштабного тестирования, где на первое место выходит эффективное решение конкретных задач, а широкая область применения свойственна коммерческим продуктам, так как привлекает больше покупателей.

Техническая структура

В ходе анализа были рассмотрены основные задачи современных систем компьютерного тестирования, которые условно можно разбить на три группы:

- формирование банка тестовых заданий и компоновка тестов;
- проведение сеансов тестирования и передача данных;
- обработка ответов, подсчёт баллов, сбор и анализ статистики. Функции первой группы, как правило, выполняются на однопроцессорных персональных компьютерах, объединённых локальной сетью и образующих довольно однородную многомашинную вычислительную систему. Разнообразие технических структур наблюдается при реализации функций второй и третьей групп.

а) *Автономные компьютеры без выхода в Интернет*. В этом варианте рабочие станции (РС) не имеют эффективных средств оперативной связи между собой и с удалённым тест-сервером (ТС). Поэтому установка программ, обновление тестов и сбор статистики выполняются отдельно для каждого компьютера с помощью съёмных носителей информации (дискет или компактных дисков), что снижает оперативность и повышает трудоёмкость обслуживания сеансов тестирования.

б) *Автономные компьютеры, один имеет выход в Интернет*. Вертикальная связь обеспечивает доступ к удалённому тест-серверу для получения программ и тестов, позволяет централизованно обрабатывать ответы тестируемых. Однако при наличии изолированных компьютеров, как и в случае а), значительную часть работы необходимо выполнять через съёмные носители информации, резко увеличивая временные и трудовые затраты.

с) *Автономные компьютеры, каждый имеет выход в Интернет.* Подобная схема даёт возможность рабочим станциям непосредственно общаться с тест-сервером, оперативно получать исходные материалы и быстро возвращать итоговые данные. Эффективность схемы зависит от пропускной способности каналов и производительности центрального сервера.

д) *Локальная сеть без выхода в Интернет.* В отличие от варианта а) локальная сеть позволяет рабочим станциям совместно использовать установленные один раз на файл-сервере программы и тесты, упрощая их обновление и значительно экономя дисковое пространство, облегчает сбор и анализ статистики за счёт ведения единой базы данных.

е) *Локальная сеть, один компьютер имеет выход в Интернет.* В этом случае любые обновлённые программы и тесты, полученные на тест-сервере, сразу становятся доступными всем рабочим станциям.

Аналогично данные, собранные в локальной сети, можно тут же через Интернет отправить на обработку центральному серверу.

ф) *Локальная сеть, каждый компьютер имеет выход в Интернет.* При некоторой избыточности данная схема имеет ряд важных свойств:

- эффективно использует пропускную способность внешних каналов, так как загруженные один раз через Интернет материалы затем совместно используются в локальной сети;
- обеспечивает максимальную оперативность передачи данных, непосредственно связывая отправителя и получателя;
- равномерно распределяет функциональную нагрузку между локальной сетью, рабочей станцией и удалённым тест-сервером.

Развитие глобальной сети Интернет и улучшение технического оснащения учебных заведений и центров тестирования* способствуют появлению новых распределённых систем компьютерного тестирования, техническую структуру которых можно отнести к классу ф). Для тестирования большого числа отдельных пользователей глобальной сети, разбросанных по всему миру, применяются системы класса с). Основные достоинства систем этих классов— централизация данных, оперативность обработки, простота администрирования, что очень важно при большом количестве сеансов тестирования.

* Coley R., Cradler J., Engel P. Computers and Classrooms: The Status of Technology in U.S. Schools. Princeton, NJ: Educational Testing Service, 1997.

Соответственно системы классов е) и б) можно отнести к продуктам переходного периода. Например, в большинстве случаев будет достаточно специальной программы, чтобы всем рабочим станциям локальной сети предоставить выход в Интернет через единственный совместно используемый модем, получив тем самым функциональный эквивалент технической структуры класса ф).

Постепенно сдают свои позиции системы компьютерного тестирования без выхода в Интернет, оставляя за собой нишу самоконтроля учащихся в процессе обучения и подготовки к сдаче итоговых экзаменов.

Алгоритм тестирования

По данному признаку современные системы компьютерного тестирования можно разделить на *линейные и адаптивные*. При линейном алгоритме набор заданий в тесте определён заранее. В случае адаптивного тестирования система отбирает задания в зависимости от предыдущих ответов, подстраивая тест к уровню тестируемого. Иногда встречаются смешанные варианты, когда одна часть теста линейная, а другая адаптивная.

Поддержка разнотипных заданий

К важным свойствам относится способность системы работать с разными типами тестовых заданий: с множественным выбором, с множественным ответом, с дополнением, со свободным ответом, на установление соответствия, выявление последовательности. По этой причине набор поддерживаемых типов тестовых заданий следует принять в качестве признака

классификации.

Реализация мультимедийных элементов

Преимуществом компьютерных тестов является возможность использовать форматированный текст, графику, анимацию, звук, видеофрагменты и активные программные элементы. Поэтому целесообразно классифицировать системы по перечню реализованных мультимедийных элементов.

Обработка ответов тестируемых

Процедура обработки ответов тестируемых может быть локальной или удалённой, формальной или экспертной, оперативной или отложенной. В зависимости от выбранного варианта обработки любая система компьютерного тестирования попадает в один из шести возможных классов.

Уровень секретности

Данный признак делит все системы на три класса — с низким, средним и высоким уровнем секретности. Перед системами первого класса не ставится задача ограничения доступа к какой-либо информации, поэтому они годятся только для самоконтроля и подготовки. Системы второго класса рассчитаны на принятие организационных мер и использование специальных программных компонент для обеспечения режима секретности на уровне паролей, шифрации, блокировки доступа к окружающей среде и т.д. Работа с системами третьего класса предполагает привлечение специалистов в области информационной безопасности для оснащения помещений, подбора программных и технических средств, обучения персонала*.

* Методические материалы по теме «Информационная безопасность». Дальневосточный государственный университет. Кафедра компьютерных систем. <http://lemoi-www.dvgu.ru/lect/>. Смелянский Р.Л. Конспект лекций по курсу «Система передачи данных и сети ЭВМ». Московский государственный университет. Кафедра АСВК. <http://cmc.cs.msu.su/curr/cn/>. Спесивцев А.В., Вегнер В.А., Крутяков А.Ю., Серёгин В.В., Сидоров В.А. Защита информации в персональных ЭВМ. М.: Радио и связь, МП Веста, 1992.

Программное обеспечение

В последнее время наибольшее распространение получили 32-разрядные операционные системы семейства Windows, поэтому неудивительно, что на работу под их управлением ориентировано большинство систем компьютерного тестирования. Реже встречаются разработки для MacOS и Unix. Программы для DOS и Win 16 либо перешли под Win32, либо морально устарели и прекратили своё существование.

Некоторые современные системы компьютерного тестирования тесно взаимодействуют не только с операционной системой, но и с HTTP-серверами, просмотрщиками HTML-страниц, офисными пакетами и т.д. Отсутствие единых стандартов вынуждает разработчиков ориентироваться на один продукт в каждой из областей, порождая тем самым дополнительное деление на классы.

Простота использования

Безусловно, современные системы компьютерного тестирования стремятся быть простыми в использовании. Препятствуют этому комплексность решаемых задач, большое число выполняемых функций, значительные объёмы данных. Достичь цели помогают: чёткое и эффективное разделение на функциональные подсистемы, ориентация на стандарты в информационной индустрии, единый стиль оформления программ, документация для каждой группы пользователей, возможность потренироваться (автору в работе с банком тестовых заданий, инструктору в управлении сеансами тестирования, администратору в установке системы, тестируемому в использовании мыши и клавиатуры).

Разработчики современных систем компьютерного тестирования

Информационной основой для проведённого анализа и предложенной классификации современных систем компьютерного тестирования являются:

- аналитические статьи, технические отчёты, научные труды авторитетных организаций и экспертов, опубликованные в последние годы за рубежом;
- материалы, размещённые в Интернете на Web-серверах крупнейших организаторов централизованного тестирования, ведущих разработчиков и основных распространителей программного обеспечения в данной области, включая описание технологий, перечень возможностей систем, демонстрационные версии, обучающие программы и рекламные ролики (см. ниже);
- пятилетний опыт реальной работы над автоматизацией отечественного централизованного тестирования, активное и непосредственное участие в которой принимали авторы (см. ниже).

Сегодня в мире существует множество организаций, разрабатывающих системы компьютерного тестирования, подробные сведения о которых можно получить через Интернет. Назовём некоторые из них.

Крупномасштабное тестирование

ETS (Educational Testing Service — Службы тестирования в образовании) — частная некоммерческая организация [<http://www.ets.org/>], созданная в США в 1947 году. Разрабатывает, готовит и ежегодно централизованно администрирует несколько (в 1995 году порядка 9) миллионов тестов (учебных достижений и вступительных) в 180 странах мира. Вопросами компьютерного адаптивного тестирования занимается с 1970 года. Планирует увеличить долю компьютерных тестов до 10%. На данный момент предлагает в компьютерном варианте:

- SAT (Scholastic Aptitude Test) I: Reasoning Test — тест для определения общих вербальных и математических способностей выпускников школ, поступающих в университеты, независимо от выбранной специальности [<http://www.college-board.org/>]. Компьютерный вариант ещё не получил широкого распространения, но у него хорошие перспективы.
- GRE (Graduate Record Examinations) General Test — основной тест для аспирантов всех специальностей, кроме медицины, бизнеса и права [<http://www.gre.org/>], результаты которого рекомендуется представлять при поступлении в аспирантуру большинства университетов США.
- GMAT (Graduate Management Admission Test) — тест для оценки подготовленности претендентов, поступающих в аспирантуру на специальности, связанные с бизнесом [<http://www.gmat.org/>]. Результаты теста используют свыше 1500 послевузовских программ обучения менеджменту по всему миру.
- TOEFL (Test of English as a Foreign Language) — тест по английскому языку как иностранному [<http://www.toefl.org/>], необходимый студентам и аспирантам, ранее обучавшимся на другом языке, для поступления в колледжи и университеты США и Канады.
- PRAXIS I: Academic Skills Assessments — сертификационный тест для оценки профессиональной пригодности начинающих преподавателей и учителей [<http://www.teachin-gandleaming.org/>]. Активно используется в 35 американских штатах.
- NBPTS (National Board for Professional Teaching Standards) Assessments — сертификационная оценка учителей математики в средних и старших классах американских школ [<http://www.teachingand-leaming.org/>]. Часть двухдневного теста предьявляется на компьютере.

Переходя на компьютерное тестирование, ETS проводит многочисленные научные исследования и эксперименты, готовит информационные материалы для тестируемых, родителей, учителей, руководителей, разъясняя выгоды от внедрения новых технологий, распространяет обучающие и демонстрационные программы (ETS Computer-Based Testing Demo, ETS Next Frontier, One-on-One with the SAT, PowerPrep, TOEFL Sampler, AccuPlacer), содержащие описание технологии тестирования и образцы тестов, успешно внедряет регистрацию

на Web-сервере, сообщение результатов заинтересованным организациям через Интернет, а тестируемым по телефону.

Однако ETS не является единственной в своём роде. Аналогичные услуги студентам и родителям, школам и колледжам, профессиональным ассоциациям и правительственным учреждениям, бизнесу и промышленности оказывает, например, корпорация ACT [<http://www.act.org/>] — независимая некоммерческая организация, созданная в 1959 году при Университете Айовы для поддержки программы тестирования абитуриентов американских колледжей (American College Testing Program). Сегодня она обслуживает клиентов по всему миру и поддерживает свыше 100 различных тестовых программ, получив за 1998 год около 1,8 млн заявок на тестирование и приняв на работу свыше тысячи новых сотрудников.

Крупнейшей в мире сетью из 2 500 центров компьютерного тестирования в 150 странах мира располагает Sylvan Prometric [<http://www.sylvan-prometric.com/>] — подразделение корпорации Sylvan Learning Systems, специализирующееся на компьютерных тестах:

- для сертификации в области информационных технологий (по заказам 60 ведущих компаний, среди которых AutoDesk, Compaq, CompTIA, Computer Associates, Corel, IBM, Lotus, Microsoft, Novell, Oracle, Sybase, Sun Microsystems и другие);
- для выдачи профессиональных лицензий (в таких важных отраслях, как здравоохранение, финансы, страхование, управление недвижимостью);
- для приёма в высшие учебные заведения (GRE, GMAT, TOEFL).

На данный момент в компьютерных центрах Sylvan Prometric проведено почти 13 млн сеансов тестирования на 25 языках (включая русский), из них свыше 3,5 млн в 1998 году.

Коммерческие продукты

Question Mark Computing Ltd. — основана в Лондоне в 1998 году. Компания [<http://www.qmark.com/>] предлагает свою систему компьютерного тестирования Question Mark в пяти вариантах: DOS, Windows, Macintosh, Web и Perception. Система успешно используется коммерческими организациями и учебными заведениями в 50 странах мира, отличается широкой областью применения и простотой освоения.

LXR (Logic eXtension Resources) — калифорнийская фирма, созданная в 1981 году [<http://www.lxrtest.com/>], распространяет систему LXR*TEST для ведения банка тестовых заданий, автоматизированной компоновки тестов, проведения сеансов тестирования на бумаге и за компьютером, обработки ответов, подсчёта баллов и анализа результатов. Система, число установок которой по всему миру превысило 28 тыс., имеет три редакции: Personal, Professional и Scoring для Windows и Macintosh. Сейчас готовится вариант для WWW.

Virtual Learning Tech-nologies — американская компания со штаб-квартирой в Бермингеме [<http://www.vlea-ming.com/>] для проведения сеансов тестирования через Интернет рекомендует свою систему CyberExam, которая позволяет легко и быстро готовить разнотипные задания с мультимедийными элементами, использовать адаптивный алгоритм или компоновать тесты со случайной перестановкой заданий, проводить сеансы тестирования с помощью обычного просмотрщика Web-страниц, немедленно получать результаты, для более тщательного анализа импортировать данные в статистические пакеты.

Среди разработчиков и распространителей программного обеспечения для систем компьютерного тестирования нужно упомянуть такие компании, как

- Assessment Systems Corporation [<http://www.assess.com/>];
- Asymetrix Learning Systems [<http://www.asymetrix.com/>];
- Bobrow Test Preparation Services [<http://www.bobrowtest.com/>];
- Chariot Software Group [<http://www.chariot.com/>];
- Cogent Computing Corporation [<http://www.cogen-tcorp.com/>];
- EBSCO Industries [<http://www.ecmtest.com/>];
- Edutest [<http://www.edu-test.com/>];
- Kamico Instructional Media [<http://www.kamico.com/home.htm>];
- Knowlton & Associates [<http://www.studio-ide.com/>];

- Macromedia [<http://www.macromedia.com/>];
- Peterson's [<http://www.petersons.com/>];
- SinartLite Software [<http://www.smartlitesoftware.com/>];
- Scholastic Testing Systems [<http://www.test-prep.com/>];
- Schroeder Measurement Technologies [<http://www.smtorg.com/>];
- Test.com [<http://www.test.com/>];
- Topgrade Software [<http://www.topgrade.com/>];
- WebTester [<http://www.webtester.com/>].

Отечественные системы

С 1996 года Центр тестирования выпускников общеобразовательных учреждений РФ для решения задач создания и хранения тестовых заданий, компоновки тестов и подготовки оригинал-макетов, обработки ответов тестируемых и подсчёта набранных баллов, печати сертификатов, сбора и анализа статистической информации, передачи данных по каналам связи использует автоматизированную систему АвтоТест*, разработанную при активном участии авторов работы с учётом технологии проведения централизованного тестирования на бланках.

* Нардюжев И.В. Принципы построения автоматизированной системы государственного централизованного тестирования/ Сборник научных трудов № 12 кафедры АСУ. Обнинск: ИАТЭ, 1998. Хлебников В.А., Сальников Н.Л., Нардюжев В.И., Нардюжев И.В. Автоматизированная система государственного централизованного тестирования/ Проблемы информатизации высшей школы. Бюллетень 3–4 (9–10) 1997. М.: ГосНИИ системной интеграции, 1998. С. 30–37. Хлебников В.А., Сальников Н.Л., Нардюжев В.И., Нардюжев И.В. Компьютерные и телекоммуникационные технологии в системе централизованного тестирования /Международное сотрудничество в образовании: Материалы научно-практической конференции. СПб.: Издательство СПбГТУ, 1998.

Система позволяет при оформлении тестовых заданий с множественным выбором использовать форматированный текст, списки, таблицы, сложные математические формулы, схемы и рисунки, назначать весовые коэффициенты, проверять орфографию, искать по ключевым словам и реквизитам, готовить шаблоны, импортировать и экспортировать текст и графику, архивировать части банка и формировать контрольные распечатки. По состоянию на 1999 год банк содержит свыше 45 000 тестовых заданий по 31 учебной дисциплине и хранится в изолированной локальной сети, доступ к которой имеет ограниченный круг ответственных лиц.

Система обеспечивает автоматизированную компоновку тестов по методу случайной перестановки типовых тестовых заданий, формируя готовые к печати оригинал-макеты и составляя таблицы ключей и весовых коэффициентов. Для кампаний тестирования 1997–1999 гг. было скомпоновано более 1000 тестов.

При проведении тестирования в единые сроки по всей России необходимо быстро обрабатывать большое количество талонов с ответами. До 1997 года ввод информации выполняли операторы, теперь это задача сканеров и специального программного обеспечения. После формального сравнения ответов тестируемых с ключами и подсчёта набранных баллов результаты поступают в базу данных, которая является источником информации для всех дальнейших операций: выдачи списков, печати сертификатов, сбора статистики. Информационные базы данных с результатами централизованного тестирования 1997–1999 гг. содержат соответственно 142, 231 и 380 тыс. записей.

Для получения на их основе оперативной статистики используется подсистема, которая позволяет выбрать объект и определить для него количество тестируемых, распределение набранных баллов и полученных оценок, процент правильных ответов на задание теста, по цифровым данным построить графики, сравнить результаты тестирования разных объектов, просмотреть список тестируемых с возможностями сортировки и быстрого поиска, распечатать статистические формы и списки на принтере, передать данные в специализированный статистический пакет для детального анализа.

Ежегодно по итогам централизованного тестирования выпускается отчёт, основой которого являются статистические формы системы АвтоТест:

- с информацией об участниках тестирования на уровне страны, региона, представительства и учебного заведения (23 типа);
- с данными (результаты классического и дистракторного анализа тестов и отдельных заданий) о качестве тестовых материалов (5 типов).

В кампаниях тестирования 1998–1999 гг. активно использовалась новая телекоммуникационная подсистема. Её основой стала корпоративная сеть Центра тестирования и его региональных представительств— РосТелеТест. В этой сети работают FTP-сервер для региональных представительств и WWW-сервер с общедоступной информацией по вопросам централизованного тестирования. Электронная почта используется в качестве резервного канала связи.

Перед телекоммуникационной подсистемой стоят три задачи:

- Информационная поддержка деятельности Центра тестирования — любой человек, имеющий доступ к сети Интернет, может прочитать приказы Министерства образования, посмотреть обобщённые статистические данные по результатам предыдущих кампаний, ознакомиться с расписанием и условиями тестирования, связаться с региональными представительствами.
- Оперативное предоставление организаторам тестирования в регионах инструкций, тестовых материалов, программного обеспечения, данных о результатах тестирования, статистической информации по всей России и конкретному региону.
- Обеспечение быстрой передачи файлов с ответами тестируемых из региональных представительств на FTP-сервер Центра тестирования для удалённой формальной отложенной обработки. Выполнение в регионах самой трудоёмкой операции ввода информации с бланков ответов позволяет значительно уменьшить нагрузку на Центр тестирования и ускорить всю процедуру обработки.

Пока телекоммуникационная подсистема успешно решала поставленные перед ней задачи. В кампаниях централизованного тестирования 1998 — 1999 гг. через Интернет на FTP-сервере в автоматическом режиме было обработано свыше 54 300 и 95 700 протоколов тестирования соответственно из 15 ведущих представительств Центра тестирования выпускников общеобразовательных учреждений РФ.

В процессе работы над системой АвтоТест в составе небольшого творческого коллектива авторами работы предложены и реализованы в виде алгоритмов, методик и программно-технических комплексов научно-практические решения в следующих областях:

- ведение банка тестовых заданий и компоновка бумажных тестов;
- организация высоконадёжной и эффективной работы аппаратных и программных средств в корпоративной сети РосТелеТест;
- проведение репетиционного и централизованного тестирования на бланках с обработкой ответов через Интернет;
- поддержка больших по объёму баз данных с реквизитами, ответами и результатами участников тестирования;
- получение оперативной статистики по всей России и по каждому региональному представительству;
- формирование ежегодного статотчёта по итогам тестирования;
- оценка качества тестов и тестовых заданий на основе классического и дистракторного статистического анализа.

Эти решения достаточно универсальны и проверены на практике, поэтому могут использоваться при создании подобных автоматизированных систем в других организациях. Например, для Головного центра тестирования граждан зарубежных стран по русскому языку разработана аналогичная система для ведения центральной базы данных о лицах, прошедших тестирование, их результатах и выданных им документах, для получения оперативной и годовой статистики, для анализа качества тестовых материалов.

Очевиден значительный прогресс в области автоматизации отечественного централизованного тестирования на бланках в то время, когда весь мир начинает переходить на компьютерное тестирование, осознав его перспективность и убедившись в его преимуществах. В период перехода проводятся исследования, создаются методики, разрабатываются программы, идёт подготовка инфраструктуры, обучение персонала. О масштабе работ можно судить по деятельности иностранных организаций и компаний, названных выше. На фоне достижений зарубежных коллег отечественные успехи выглядят весьма скромно. Такое положение дел лишний раз доказывает актуальность данного исследования.

Требования к информационно-вычислительной системе компьютерного тестирования

Приведённый анализ, рассмотренные зарубежные примеры и имеющийся отечественный опыт указывают на необходимость разработки информационно-вычислительной системы компьютерного тестирования (ИВС КТ) — системы, состоящей из одного или нескольких компьютеров и набора программ, обеспечивающих весь технологический цикл от создания тестов и проведения тестирования на автономных компьютерах, в локальной сети, через Интернет до статистического анализа результатов и оценки качества тестовых материалов.

ИВС КТ должна обладать следующими свойствами:

- комплексность (для решения задач создания и хранения тестовых заданий, компоновки тестов, регистрации участников, проведения сеансов тестирования на компьютерах, обработки ответов и подсчёта набранных баллов, сбора и анализа статистической информации, передачи данных по каналам связи и обеспечения секретности);
- производительность (для получения, хранения, обработки и передачи больших объёмов информации и параллельного выполнения множества функций при наличии ограниченных вычислительных ресурсов);
- надёжность (для предоставления услуг одновременно большому количеству людей с редкими сбоями и возможностью быстро обнаружить и устранить неисправность и полностью восстановить основные информационные базы данных, насчитывающие десятки и сотни тысяч записей);
- масштабность (для расширения диапазона возможных применений: тестирование на автономных компьютерах, в локальной сети, через Интернет с оперативной или отложенной обработкой ответов одного человека или сотен тысяч тестируемых);
- защищённость (для обеспечения уровня секретности, соответствующего важности решений, принимаемых по результатам тестирования);
- простота (для снижения финансовых и временных затрат на установку, освоение и обслуживание всего программно-технического комплекса).

В условиях роста популярности варианта предъявления компьютерных тестов через Интернет, при котором функции проведения сеансов тестирования распределены между рабочими станциями и тест-сервером и предполагают интенсивный обмен данными по каналам связи в режиме «запрос-ответ», целесообразно в качестве показателя эффективности ИВС КТ выбрать среднее время отклика на запрос при ограничениях на производительность системы.

Постановка задач исследования

Перечислим исходные данные:

- результаты анализа функций, технической структуры и программного обеспечения современных систем компьютерного тестирования и предложенная классификация;
- рассмотренные зарубежные примеры и имеющийся отечественный опыт автоматизации бланковой технологии;
- определение информационно-вычислительной системы компьютерного тестирования, требования к ней и показатель её эффективности.

Задачи исследования:

- Создать концептуальную модель ИВС КТ, разработать аналитическую модель центрального блока обработки заявок, поступающих на тест-сервер, и имитационную модель телекоммуникационной подсистемы, исследовать с помощью моделей работоспособность различных вариантов ИВС КТ.

- Разработать и исследовать центральные алгоритмы взаимодействия рабочих станций и тест-сервера в рамках ИВС КТ.

- Разработать и исследовать информационное обеспечение подсистем автоматизированной компоновки компьютерных тестов, проведения сеансов тестирования и статистического анализа результатов.

- Разработать и исследовать варианты программной реализации основных модулей ИВС КТ.

- Разработать методику синтеза ИВС КТ и дать рекомендации по её экспериментальным испытаниям.

- Внедрить разработанные математические модели, центральные алгоритмы и программное обеспечение ИВС КТ в ведущих организациях, занимающихся централизованным тестированием в России.