

Адаптивное тестирование знаний в системе «телетестинг»

Шмелёв А.Г.
Бельцер А.И.
Ларионов А.Г.
Серебряков А.Г.

Центром тестирования «Гуманитарные технологии» (МГУ) в течение последнего времени создана и развивается Интернет-ориентированная система компьютерного тестирования «Телетестинг» (авторское свидетельство №10267 в агентстве «Роспатент»). В 1997–1999 гг. с использованием этой системы были проведены три всероссийские компьютерные олимпиады с одноимённым названием, в которых приняли участие почти 16000 человек из 70 различных городов РФ и ближнего зарубежья.

В 1999 году была впервые внедрена автоматическая обработка файлов-протоколов в режиме FTP-связи: на FTP-сервере для каждого регионального центра создавалась отдельная директория с личным паролем доступа; после появления в этой директории файла с исходными протоколами (ответами на пункты теста) специализированная программа на сервере немедленно обрабатывала протоколы и создавала файл-таблицу, которую пользователь мог немедленно буксировать на свой компьютер – в ходе одного двухстороннего сеанса подключения в режиме FTP. В 1999 году этим сервисом успешно воспользовались уже 42 региональных центра, что объясняется, очевидно, организационными и психологическими удобствами такой формы обработки результатов. Важно подчеркнуть, что региональные организаторы могут ознакомиться с тестовыми данными, раньше, чем такая возможность предоставляется организаторам олимпиады в Москве. Таким образом, благодаря современным технологиям мы получаем новый важный источник повышения достоверности тестирования и доверия к объективности подобного способа оценки знаний.

Новым инструментом повышения не только достоверности, но прежде всего точности компьютерного тестирования знаний стал впервые внедрённый в ходе «Телетестинг-99» САТ-подход (принятая на западе аббревиатура для «Computer Adaptive Testing»). Осуществить переход к САТ в системе «Телетестинг» удалось на основе созданного за три года банка заданий (свыше 6000 по основным школьным предметам) с эмпирически измеренным уровнем трудности (для нашего алгоритма уровень трудности определялся с точностью до порядковой шкалы с тремя градациями).

Ниже даётся краткое описание принципов адаптивного алгоритма, который был реализован в ходе олимпиады «Телетестинг-99»:

1. Адаптивность выражается в изменении относительных пропорций при предъявлении лёгких, средних и трудных заданий в зависимости от числа правильных ответов зарегистрированных в ходе сеанса тестирования.

2. Адаптивность сочетается с принципом тематической репрезентативности: равномерной представленности различных тематических разделов учебной программы в тесте.

3. Адаптивность сочетается с принципом «лестничного алгоритма» – предъявление заданий с систематическим нарастанием уровня трудности. Вначале предъявляются лёгкие задания, затем – средние и, если испытуемый успешен на предыдущих уровнях, – трудные.

4. Подготовка варианта. Для каждого испытуемого перед тестированием программа создаёт уникальный вариант, включающий «лишние» задания (по 50% лёгких, трудных и заданий средней сложности от количества, которое планируется реально предъявить).

5. Начальный цикл. На каждом уровне) программа предъявляет всем испытуемым один и тот же начальный цикл заданий, независимо от качества их ответов, т.е. адаптивность в ходе начального цикла отключена. Это К лёгких заданий – по числу тематических разде-

лов. Из каждого тематического раздела программа подбирает задания случайным образом.

6. Начиная с заданий $K+1$ после каждого очередного ответа, программа тестирования уже включает адаптивность и учитывает правильность ответов – для того, чтобы определить обоснованность так называемого «досрочного перевода» испытуемого на более высокий уровень. На каждом шаге оценивается значимость различий между числом правильных (А) и неправильных (В) ответов по формуле, имеющей приближённое распределение Хи-квадрат с одной степенью свободы:

$$\chi^2 = \frac{(A-B)^2}{A+B}$$

При значениях $\chi^2 > 4$ (что приближенно соответствует пятипроцентному уровню ошибки в отвержении гипотезы равновероятной встречаемости правильных ответов и ошибок) осуществляется досрочный (до исчерпания заданий данного уровня) перевод испытуемого на более высокий уровень трудности.

7. Подсчёт очков. За решение задач на более высоком уровне даётся больше очков. Принятые размеры начисления очков отражают пропорции в вероятности случайного правильного ответа при принятом нами уровне трудности заданий (вероятности случайного правильного ответа).

Таким образом, мы считаем, что разработали и применили один из самых осторожных (мягких) алгоритмов адаптивности, учитывающий только возможность досрочного перехода на более высокий уровень с нижнего и среднего уровней трудности. Алгоритм даёт высокую корреляцию с обычным неадаптивным методом формирования случайных вариантов, но повышает точность оценивания для крайних испытуемых (слабо и хорошо подготовленных), так как обе крайние категории получают больше заданий релевантного им уровня трудности.

Ещё один значимый эффект – повышение достоверности в результате САТ, так как в этом случае исключается быстрое изучение банка заданий путём простого «прощёлкивания» вариантов на компьютере (таким образом, можно узнать лишь «лёгкие» задания, а трудные и часть средних оказываются неизученными).

Любая научная или практическая лаборатория компьютерного обучения может апробировать наш подход в действии, если организуют у себя пункт проведения олимпиады «Телетестинг». Заявки принимаются по адресу электронной почты – tt2000@ht.aha.ru.