

О.В. Стукач, канд. техн. наук
(Томский политехнический университет, Томск)

СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ОЦЕНКИ КАЧЕСТВА ПОДГОТОВКИ СПЕЦИАЛИСТОВ ПО ИНФОРМАЦИОННЫМ ТЕХНОЛОГИЯМ

Предложена новая методика оценки качества подготовки специалистов по информационным технологиям, основанная на использовании карт контроля качества и модифицированном методе экспертных оценок. Рассмотрены особенности тестирования специалистов во время обучения. На примере показаны преимущества разработанной методики.

Модифицированный метод экспертных оценок

Определяющим фактором конкурентоспособности любой организации является качество подготовки специалиста. На эту тему написано немало работ, например, [1-2], но до количественной оценки качества подготовки авторы так и не дошли. Традиционные способы оценки успеваемости – по пятибальной или рейтинговой системе не годятся для подготовки разработчиков программного обеспечения (ПО).

Особенностью подготовки выпускников по специальности "разработчик программного обеспечения" (IT engineer) является то, что он должен не только иметь соответствующие знания по своей специальности, но и уметь работать в команде. Современные программные комплексы создаются большим числом активно взаимодействующих между собой разработчиков, поэтому в процессе выполнения проекта важны не сколько знания и умения каждого разработчика, сколько средний уровень команды. Кроме того, следует учитывать, что задача оценки качества ПО не имеет достоверного правильного ответа. Поэтому обычные тестовые задачи зачастую не подходят для обучения разработчиков современного ПО.

Решение в числовой шкале задачи оценки уровня подготовки специалистов необходимо работодателям для отбора специалистов при приеме на работу. Следствием этого является существенное снижение затрат фирм на разработку ПО и понимание того, на каком уровне специалист может решать поставленные задачи. Поэтому проблема разработки численной методики оценки качества подготовки весьма актуальна, особенно в условиях обострения рыночной конкуренции.

Использование метода экспертных оценок позволяет решить триединую задачу:

- продемонстрировать наличие нестандартных творческих способностей;
- оценить уровень знаний каждого специалиста;
- показать, насколько специалист справляется с работой в команде разработчиков.

Предложена новая методика тестирования и оценки успеваемости специалистов ПО, краткое описание которой заключается в следующем. Специалисту, выступающему в роли эксперта, предлагается ответить на ряд вопросов, оценивая несколько возможных альтернатив. Например, можно предложить оценить три исходных текста программ по возможным критериям: краткость кода, оригинальность алгоритма, выполнение требования минимального объема памяти, выполнение требования минимального времени исполнения, абсолютная автономность и так далее. Оценивание выполняется в нечисловой шкале, чтобы исключить возможную необъективность. Например, могут использоваться следующие оценки: отлично, очень хорошо, хорошо, удовлетворительно, посредственно, плохо. При проверке работ нечисловая шкала заменяется числовой (от нуля до единицы) так, как это сделано в известном методе Делфи [3]. Далее вычислялось среднее значение и дисперсия по ответам на вопросы каждого эксперта.

Работа в команде означает, что поставленная задача решается всеми экспертами. Поэтому из оценок каждого эксперта выводился так называемый "средний эксперт", оценки которого представляли собой средние оценки всех экспертов. В этом случае мы моделировали классический метод экспертных оценок. Ясно, что средний эксперт не всегда удовлетворяет предъявляемым к нему требованиям. В большинстве случаев некачественная оценка может попасть в оценки среднего эксперта. Средний эксперт позволяет всего лишь увидеть тенденции в оценках всей группы экспертов, но его оценки – это не гарантия приближения к истине.

Из формальной теории измерений известно, что истинность измерений определяется путем сравнения с эталоном. В методе экспертных оценок эталона не существует, так как нет абсолютно истинного результата измерения. В нашем случае роль эталона выполнял преподаватель, уровень знаний которого принимался за тот, который должен быть достигнут каждым экспертом и, следовательно, средним экспертом.

Вычислялись среднее значение и дисперсия для оценок каждого эксперта, среднего и эталонного экспертов. Это позволяет сравнить уровень знаний экспертов друг относительно друга и со средним экспертом. Один из результатов проведенных исследований показан на рис. 1. Здесь использовались следующие числовые оценки: отлично – 1, очень хорошо – 0,75, хорошо – 0,625, удовлетворительно – 0,5, посредственно – 0,25, плохо – 0.

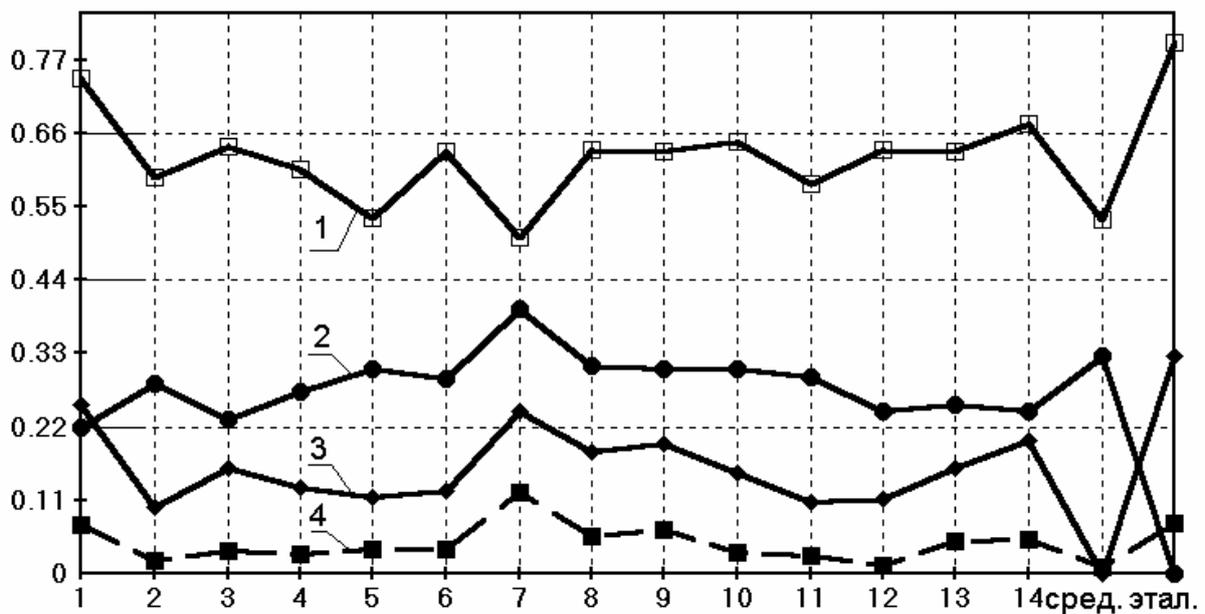


Рис. 1. График 1 – среднее значение оценок каждого эксперта; 2 – отклонение оценки эксперта от оценки эталонного эксперта; 3 – отклонение оценки эксперта от оценок среднего эксперта; 4 – дисперсия оценок экспертов.

По оси абсцисс отмечены номер эксперта, средний и эталонный эксперт. По графикам можно наблюдать отклонение оценок экспертов от эталона и среднего. Ближе всего к среднему значению второй и пятый эксперт. Довольно далеки в своих оценках первый, седьмой и 14-й. Видно, что оценка среднего эксперта не гарантирует приближение к истинному значению – оценке эталонного эксперта. Его средняя оценка на графике 1 существенно отличается и от оценки среднего эксперта, и от оценки каждого. Это говорит о том, что знания эталонного эксперта существенно выше среднего.

По графикам разностей можно определить, насколько велико отклонение оценок каждого эксперта от эталонного (график 2) и от среднего (график 3). На рис. 1 средняя оценка первого эксперта наиболее приближена к эталонному, что говорит о том, что его знания самые высокие в группе. Наиболее выделяется из группы седьмой эксперт, оценки которого далеки от среднего и эталонного и который может быть признан неуспешным. Об этом говорит и график дисперсии оценок (график 4) – у седьмого эксперта она наибольшая.

Равномерная дисперсия свидетельствует о том, что сильных противоречий в группе относительно проделанной работы не наблюдается. При росте дисперсии оценки необходимо пересмотреть, либо скорректировать. По дисперсии можно также судить о том, насколько эксперты способны работать в команде. При сильном разбросе дисперсии можно говорить о том, что принимаемое общее решение группой экспертов вряд ли будет правильным, хотя некоторые эксперты при этом могут даже быть абсо-

лютно точны в своих оценках. По равномерной дисперсии можно судить о том, насколько эксперты честно и добросовестно выполняли свою работу.

Результаты проведенных исследований позволяют решить задачу выделения из группы наиболее результативных в своих оценках экспертов. На этом основании было предложено формировать из выделяющихся в нескольких группах экспертов (например, первый эксперт) отдельные группы. По дисперсии мы судили о том, насколько эти эксперты совместимы в принятии единых решений. Добившись равномерной дисперсии, нам удалось сформировать единую команду экспертов.

Контрольные карты

Многие организации видят свое развитие в сотрудничестве с университетами. Возврат инвестиций в образование достигается ими за счет повышения качества образования. В этом случае возникает задача контроля качества процесса обучения.

Управление качеством (Total Quality Measurement) широко используется во многих технологических процессах в современном производстве и сфере услуг [4]. Система обеспечения качества необходима для того, чтобы непрерывно контролировать состояние процесса и гарантировать качество продукции или услуг, а при необходимости вмешаться в процесс при его отклонениях от нормального режима и предотвратить тем самым потери от выпуска некачественной продукции. Средства управления качеством могут использоваться почти в любом производственном процессе для определения и выбора проблемы, диагностики возможных причин возникновения проблемы и выработки решений и рекомендаций по устранению проблемы [5].

Предложено применить концепции управления качеством в системе подготовки специалистов по информационным технологиям. Для этого использовались статистические методы управления качеством для мониторинга, оценки и выработки решений по управлению образовательным процессом на основе контрольных карт [5]. Построение карт проводилось с помощью статистического пакета STATISTICA [6].

На рис. 2 представлен пример контрольной X-R карты. Карта состоит из двух гистограмм, X-карты отклонений и R-карты размахов. По гистограмме X-карты судят о том, подчиняется ли распределение оценок нормальному закону распределения. Если закон распределения далек от нормального, объем выборки следует увеличить. Процесс обучения представлен в виде изменяющийся во времени кривой на картах отклонений и размахов. Как видно из графиков, изменение может происходить в пределах, называемых границами регулирования (штриховые линии). Эти границы автоматически рассчитываются по известным из [4] формулам. Также их можно искусственно установить по оценкам обучаемых, то есть тем пара-

метрам, которые бы определяли такие знания и навыки у разработчиков ПО, которые хотели бы видеть в них работодатели. Средняя линия представляет собой усреднение по оценкам. По отклонениям от среднего можно судить о качестве процесса образования. На X-карте можно видеть точки, выходящие за границы регулирования. Для оценки качества производственного процесса это означает, что необходимо вмешаться в процесс, так как он становится неконтролируемым. В нашем случае это может означать, что ряд экспертов демонстрирует более высокие знания, чем средний эксперт. Желательно, чтоб этих точек становилось как можно больше. В примере на рис. 2 за контрольными пределами оказался эталонный эксперт. Серия точек, то есть состояние процесса, при котором последовательные точки лежат по одну сторону от средней линии свидетельствует о том, что на процесс действует однонаправленная причина, связанная с недостаточной проработкой учебного материала. Следует вернуться к рассмотренной теме и изучить ее еще раз более основательно. Характерно, что на R-карте не наблюдается серии точек в том месте, где она присутствует на X-карте. Поэтому по R-карте трудно судить о качестве процесса обучения, поскольку в данные включен эталонный эксперт, результаты которого влияют на общую картину обучения группы. Следует либо временно исключить эталонного эксперта из данных, либо строить карту отклонений оценок от эталонных отдельно.

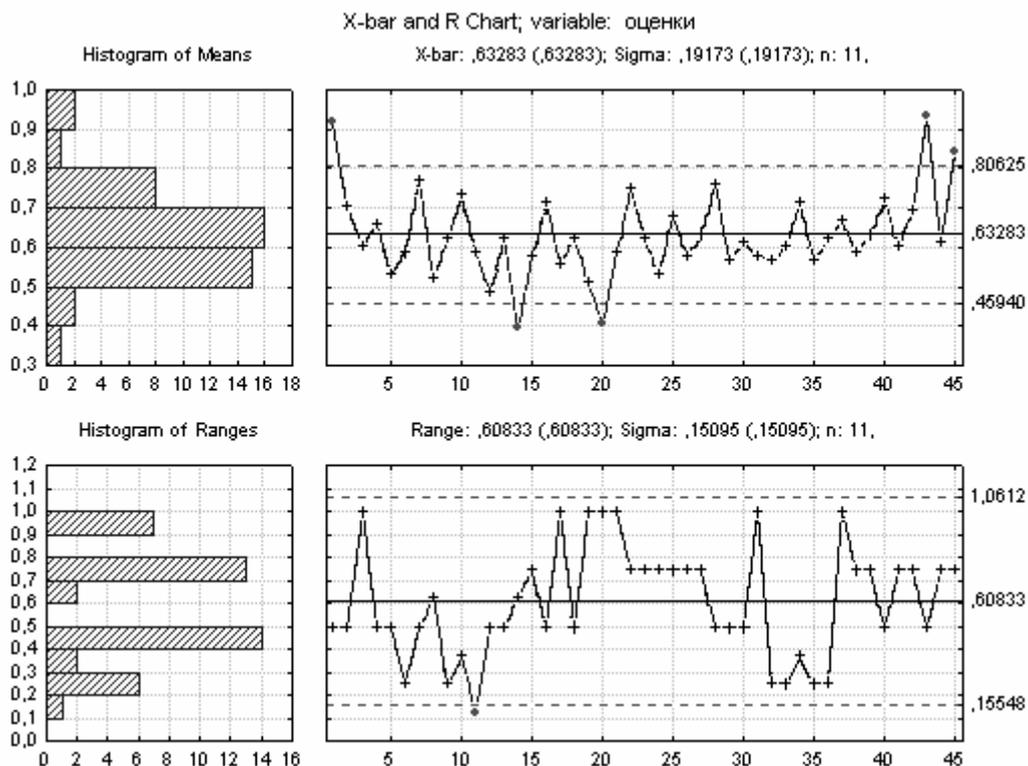


Рис. 2. Пример X-R карты для контроля качества подготовки разработчиков ПО

На рис. 3 изображена контрольная X-карта отклонений средних оценок экспертов от оценок эталонного. Для сравнения оценки эталонного эксперта также включены в карту. По отклонениям от средней линии можно судить о том, через какое время оценки экспертов начинают приближаться к оценкам эталонного, то есть насколько быстро группа приобретает знания. В приведенном примере изображена карта входного контроля, то есть на время начала обучения.

Следует обратить также внимание на то, что оценки первого эксперта наиболее близки к оценкам эталонного. Это видно из обеих карт на рис. 2 и 3. Следовательно, уровень знаний первого эксперта самый высокий в группе. По мере приобретения знаний все больше точек будет приближаться к средней линии. Это даст возможность провести расслоение, то есть заняться поиском группы, имеющей одинаковые средние оценки, что будет означать повышение качества работы специалистов в команде.

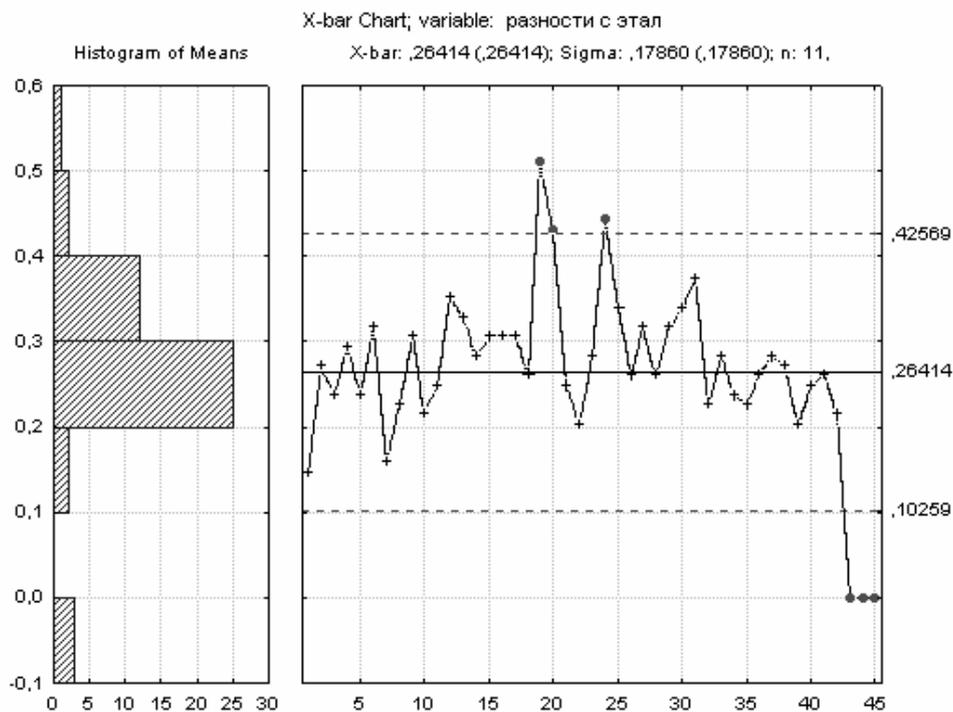


Рис. 3. Пример контрольной X-карты отклонений оценок экспертов и эталонного эксперта

Таким образом, анализ контрольных карт позволяет корректировать процесс обучения разработчиков ПО во времени и дает информацию о том, насколько качественно усваивается учебный материал. Включение в набор данных оценок эталонного эксперта позволяет оценить степень подготовленности обучаемых и определить уровень знаний каждого из них.

Заключение

Быстрый рост экономического сектора, связанного с информационными технологиями и возрастающая конкуренция требуют от организаций качественных решений все более сложных задач в сжатые сроки. Использование статистических методов контроля качества в современных условиях является самым доступным и доказательным способом оценки уровня подготовки разработчиков ПО. В данной работе показано, что модифицированный метод экспертных оценок позволяет проводить доказательный отбор специалистов и позволяет определить, насколько они умеют работать в команде. Метод контрольных карт позволяет оценить качество обучения непосредственно в его процессе и, следовательно, позволяет вовремя скорректировать его.

ЛИТЕРАТУРА

1. *B. Hergenhahn, M. Olson*, "An Introduction to Theories of Learning". Englewood Cliffs, NJ: Prentice-Hall, 1997.
2. *G. Bengu, W. Swart*, "A computer-aided, total quality approach to manufacturing education in engineering". In IEEE Transactions on Education. Vol. 39, pp. 415-422, August 1996.
3. *Силич В.А., Силич М.П.* Системный анализ и исследование операций: Учеб. пособие. – Томск: Изд-во ТПУ, 2000. – 97 с.
4. Статистические методы повышения качества / Под ред. Х. Кумэ. – М.: Финансы и статистика, 1990. – 301 с.
5. *R. Rada, K. Hu*, "Patterns in Student – Student Commenting". In IEEE Transactions on Education. Vol. 45. No. 3, pp. 262–267, August 2002.
6. *Боровиков В.П., Боровиков И.П.* STATISTICA - Статистический анализ и обработка данных в среде Windows. – М: Информационно-издательский дом "Филинь", 1997. – 608 с.