

Е.В. Грязнова, А.Е. Краснов

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ БАЛЛЬНО-РЕЙТИНГОВОГО ОЦЕНИВАНИЯ КАЧЕСТВА ОБУЧЕНИЯ В ВЫСШЕЙ ШКОЛЕ

Аннотация. Предметом исследования данной статьи является балльно-рейтинговая система оценивания качества обучения в высшей школе. Интерес представляет анализ проблем, связанных с внедрением этой системы в высшей школе нашей страны не только на качественном, но и количественном уровнях. Актуальность подобного исследования обосновывается тем, что в условиях становления информационного общества в системе традиционного образования оказалось сложным реализовать высокий уровень динамичности, доступности, вариативности, дифференцируемости и контроля образовательного процесса, особенно на уровне бакалавриата, которые требуются в условиях нового общества знаний. Образование рассматривается как сложная динамическая развивающаяся система.

В основе методологии исследования лежит рейтинговая технология оценивания качества функционирования сложных систем, в том числе качества знаний в системе образования. Анализируются два метода оценивания степеней сложности дидактических и экспертных модулей: метод эвристических суждений и метод согласованных суждений.

Рейтинговая технология оценивания качества знаний в системе образования, путем выставления, так называемого, рейтинга в многобалльной шкале значительно снижает вероятность ошибки оценивания. Для выбранного случая 5-и и 100 балльных шкал оценивания на уровне вероятности ошибки РОШ = 0,0025 (не более одной ошибки на четыреста случаев) обоснованы следующие правила перевода рейтингов в оценки: 2 — рейтинги до 48 баллов; 3 — рейтинги лежат в пределах 52-68 баллов; 4 — рейтинги лежат в пределах 72-88 баллов; 5 — рейтинги лежат в пределах 92-100 баллов. В диапазонах 48-52, 68-72 и 88-92 баллов обоснованные оценки качества функционирования системы (например, знаний обучаемых) на уровне вероятности ошибки РОШ = 0,0025 отсутствуют. Здесь необходимо проводить дополнительное оценивание, например, в случае образования — переэкзаменовку, если студент пожелает повысить свою оценку, или принимать решения: «в пользу студента», «в пользу преподавателя». Вероятность ошибки оценивания заключена в пределах, обусловленных обоснованностью выбора степеней сложности экспертных модулей. Данные степени сложности, в свою очередь, зависят от сложности соответствующих системных (дидактических в случае системы образования) модулей.

Ключевые слова: балльно-рейтинговая система, рейтинг, компетентности, ошибка оценивания, модульно-рейтинговое оценивание, дидактический модуль, экспертный модуль, метод эвристических суждений, метод согласованных суждений, сложная система.

Введение

Целью данной статьи является анализ проблем, связанных с внедрением балльно-рейтинговой системы в высшей школе нашей страны не только на качественном, но и количественном уровнях. Россия, особенно в советские времена, отличалась высоким качеством образования на всех уровнях. Но пришло время, когда новации вступают в противоречие с традициями согласно всеобщим законам развития.

Результат этого противостояния может быть как положительным, так и отрицательным, но в любом случае перемены неизбежны. И в зависимости от того насколько правильно и своевременно будут выявлены, а затем решены проблемы данного преобразования, и будет зависеть конечный результат.

Традиционная система образования в нашей стране в определенный момент развития информационного общества перестала вписываться в мировую образовательную систему, что и потре-

бовало ее реформирования. Основной причиной послужила информатизация, охватившая все сферы общественной жизни. В традиционном образовании оказалось сложным реализовать высокий уровень динамичности, доступности, вариативности, дифференцируемости и контроля образовательного процесса, особенно на уровне бакалавриата, которые требуются в условиях нового общества знаний.

Мировой опыт показывает, что подобные недостатки нивелируются с введением в образование балльно-рейтинговой системы. Действительно, становится очевидным, что классическая форма экзамена зачастую нерезультативна, т.к. объем информации с каждым годом увеличивается, преподавателям проверить студента на знание всего пройденного курса за один экзаменационный ответ невозможно, происходит усреднение студентов, занимающихся регулярно и тех, кто готовится от случая к случаю, только в период сессии. Кроме того, внедрения балльно-рейтинговой системы требует реализация Болонского соглашения, что делает этот процесс неизбежным.

Еще со времен существования дореволюционных гимназий механизм оценивания качества обучения в отечественном образовании основывался на пятибалльной системе отметок. Как и любая другая, она имеет определённые достоинства и недостатки.

Анализ имеющиеся недостатков традиционной системы контроля качества знаний, приводит к постановке вопроса о поиске принципиально иного подхода к оцениванию, который способствовал бы индивидуализации учебного процесса, повышению мотивации обучаемых и их самостоятельности. В этой ситуации практический интерес представляет балльно-рейтинговая система оценивания, поскольку она является более гибкой и может включать в себя различные элементы других систем.

Основные положения балльно-рейтинговой системы изложены в методических рекомендациях Министерства образования и науки РФ. Она обеспечивает:

- упорядочение, прозрачность и расширение возможностей применения различных видов и форм текущего и промежуточного контроля качества процесса и результатов обучения;
- формализацию процесса оценивания;
- реализацию индивидуального подхода в образовательном процессе;

Достоинства	Недостатки
1. Отметки интуитивно понятны.	1. Отметка играет не стимулирующую, а только констатирующую роль.
2. Система проста в употреблении.	2. Учитывается лишь успешность усвоения стандартного набора знаний, причем усвоения большей частью экстенсивного (выучил – воспроизвел).
3. Отметки удобны для конкурсов, вступительных экзаменов, статистики, отчётности.	3. В области профильного обучения оценки становятся неравнозначными.
4. Система позволяет оценить творческие способности студента в силу отсутствия жестких критериев контроля.	4. Возможна большая относительная ошибка оценивания знаний.
	5. Система выполняет функцию внешнего контроля успешности обучения: не предполагает оценки учеником собственных действий, сопоставления его самооценки с внешней оценкой.
	6. Система затрудняет индивидуализацию обучения: невозможно оценить реальные достижения каждого конкретного студента в сравнении с его предыдущими результатами.
	7. Система является малоинформативной (вырожденной): по отметке часто нельзя судить о действительном уровне знаний.
	8. Оценка носит часто субъективный характер.

- формирование у обучаемых мотивации к систематической работе — аудиторной и самостоятельной;
- стимулирование обучаемых к освоению образовательных программ на базе объективности и дифференциации оценки результатов их учебной работы;
- рейтингование обучаемых по степени формирования компетенций, включающих как результаты освоения дисциплин, так и их творческую (академическую) активность;
- объективную базу отбора обучаемых для продолжения учебного процесса (магистратура, аспирантура), прохождения семестрового обучения за рубежом и трудоустройства выпускников;
- корректировку преподавателями учебного процесса и оказания воспитательного воздействия на студента.

Можно выделить следующие преимущества балльно-рейтинговой системы оценивания (как для обучаемого, так и для преподавателя):

Для обучаемого	Для преподавателя
1. Возможность распоряжаться своим временем самому.	1. Рациональное планирование учебного процесса, организации индивидуальной и творческой работы учащихся.
2. Проведение постоянной самодиагностики и самоконтроля учебных достижений.	2. Стимулирование эффективного обучения старшекурсников.
3. Возможность выбора порядка выполнения учебных заданий, самостоятельное планирование работы.	3. Возможность своевременно вносить коррективы в организацию учебного процесса.
4. Сравнение уровня своих знаний с уровнем других учащихся.	4. Объективная оценка выполнения каждым учащимся всех учебных заданий.
5. Наличие соответствующих прав (освобождение от зачёта, экзамена и т.д.).	5. Возможность объективно определять итоговую оценку по предмету.

Балльно-рейтинговая система основана на применении простейших шкал (например, той же традиционной 5-ти балльной системы) для оценивания уровня освоения отдельных компетенций и накопления этих оценок в виде баллов, в свою очередь, характеризующих освоение совокупности компетенций в процессе обучения.

Рейтинг (англоязычный — *rating*) — дословно оценка, отнесение к тому или иному классу. Как правило, рейтинг выражается в виде отношения набранных баллов к максимально возможному (процент освоения раздела или всей дисциплины). В линейных системах баллы и рейтинги — синонимы. В нелинейных системах оценивания удобнее пользоваться рейтингами, а не баллами.

Компетентности представляют собой характеристики качества подготовки обучающихся, оценка которых не может быть в полной мере стандартизована. Они с трудом поддаются измерениям. Трудность здесь видится в том, что компетентность нельзя трактовать как сумму предметных знаний и умений. Это — приобретаемое в результате обучения новое качество, увязывающее знания и умения со спектром интегральных характеристик качества подготовки, в том числе и со способностью применять полученные знания и умения на практике. Поэтому возникает задача создания междисциплинарных (комплексных) измерителей, требующих при оценке результатов обучения использования методов многомерного шкалирования и специальных методов интеграции оценок отдельных характеристик обучаемых.

Конечно, данные показатели достаточно поверхностны. Чтобы понять сущность изучаемого

явления, необходим системный подход, в рамках которого при реализации многопланового и дорогостоящего проекта одной из важнейших проблем является оценивание качества функционирования сложных систем, на которых базируется проект.

Основы рейтингового оценивания

При реализации многопланового и дорогостоящего проекта одной из важнейших проблем является оценивание качества функционирования сложных систем, на которых базируется проект.

Сложной называется система, состоящая из многих функционально связанных частей, не только наследующая свойства этих частей, но и обладающая новыми свойствами, которые не были ни у одной из частей системы. При этом сложные системы обладают также свойствами устойчивости и потенциальной возможности развития. Образование попадает под это определение. Действительно, в процессе образования изучается достаточно ограниченный круг теорий на ограниченном количестве практических задач. Однако человек становится способным решать новые, ранее не встречавшиеся задачи и создавать новые теории.

Как правило, для оценивания качества функционирования даже сложных систем специалисты используют простые шкалы, выраженные лингвистическими или числовыми переменными. Например, система функционирует в режиме: плохо, удовлетворительно, хорошо, отлично. Или, например, знания обучающегося оценены как: 2, 3, 4, 5. В данных примерах состояние функционирования системы относится к четырем разным классам или множествам.

Однако совершенно ясно, что необходим некоторый точный инструмент отнесения режима функционирования системы к конкретному классу. Для сложных систем цена ошибки может быть очень высока (неустраняемая авария, помеха в карьере уникального специалиста и т.п.).

Для сложных систем в качестве инструмента принято выбирать рейтинговую технологию оценивания качества их функционирования. Технология заключается в том, что для оценивания качества функционирования системы приглашаются несколько экспертов, которые одновременно по одной и той же номинальной шкале с помощью одинаковых наборов правил (программ) проводят оценивание системы. Несмотря на одинаковые программы оценивания, каждый эксперт подмечает

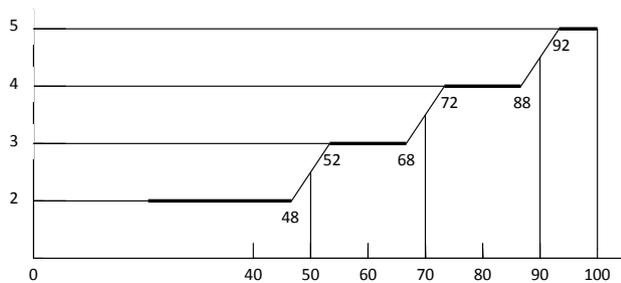


Рис. 1. Шкала перевода рейтингов в номинальные оценки

ет «свои свойства» системы. Далее все оценки экспертов складываются, а результат переводится в номинальную шкалу.

Сразу же заметим, что перевод в номинальную шкалу путем усреднения оценок экспертов (деления суммарных оценок на число экспертов) неправилен, т.к. мы имеем дело со сложными системами, свойства которых не являются суммами отдельных свойств.

Для наглядности рассмотрим следующий пример. Предположим, что для оценивания функционирования системы приглашены $N = 20$ равнозначных экспертов. В качестве номинальной шкалы выбрана числовая 5-и балльная шкала (2, 3, 4, 5). Тогда диапазон суммарной экспертной оценки будет заключен в пределах: 40 ÷ 100 баллов. Введя 100 балльную суммарную шкалу, мы достигли высокой точности оценивания состояний системы, так как по правилам статистики абсолютная ошибка σ_3 любого эксперта в 1 балл в суммарной оценке приводит к ошибке $\sigma_\Sigma = \sigma_3 / \sqrt{N} = 1 / \sqrt{20} = 1 / 4,47 \approx 0,2$.

Задача теперь заключается в правильном переводе 100 балльной шкалы оценок в номинальную 5-и балльную шкалу (см. рис. 1).

Здесь сразу же возникает проблема — точки 50, 70 и 90 суммарной 100 балльной шкалы являются неопределенными (неустойчивыми) при переводе их в точки номинальной 5-и балльной шкалы. Данные неустойчивые точки соответствуют случаям, когда мнения экспертов разделились поровну (10 на 10). Например, 10 экспертов дали оценки 3, а 10 — 4. Что будет, если 11 экспертов дадут оценки 3, а 9 — 4. В этой ситуации один эксперт может ошибиться на 1 балл. Вероятность ошибки (один случай из 20) равна $P_{ош} = 1/N = 1/20 = 0,05 = 5\%$. В результате точка 70 суммарной 100 балльной шкалы сместится к точке $11 \cdot 3 + 9 \cdot 4 = 69$. Очевидно, что в этом случае ошибочно (с вероятностью 5%) выставлять оценку 3 системе в номинальной шкале. Чтобы снизить вероятность ошибки будем выставлять оценку 3, ког-

да суммарная оценка экспертов равна $12 \cdot 3 + 8 \cdot 4 = 68$. В этом случае ошибка оценивания (если эксперты независимы, т.е. между ними нет договоренности) равна вероятности двух независимых событий (каждый из двух экспертов ошибся) $P_{ош} = (1/N)^2 = (0,05)^2 = 0,0025$ или 0,25%. Тем самым, мы в 20 раз уменьшили ошибку оценивания качества функционирования сложной системы.

Таким образом, в рассмотренном примере режимы функционирования системы с 0,25% вероятностью ошибки (менее 3-х случаев из 1000) разбиваются на следующие классы (см. рис. 1):

- 2 — экспертные оценки до 48 баллов;
- 3 — экспертные оценки лежат в пределах 52 ÷ 68 баллов;
- 4 — экспертные оценки лежат в пределах 72 ÷ 88 баллов;
- 5 — экспертные оценки лежат в пределах 92 ÷ 100 баллов.

Проблема заключается в том, что в диапазонах 48 ÷ 52, 68 ÷ 72 и 88 ÷ 92 обоснованные оценки качества функционирования системы отсутствуют — имеются люфты оценивания. В данных диапазонах вероятность ошибки оценивания возрастает в 20 раз (с 0,25% до 5%).

В случае образования любой обучающийся, набравший, например, количество баллов из диапазона 68 ÷ 72 имеет гарантированную тройку, но может по праву апеллировать к переэкзаменовке, чтобы повысить свою оценку до 4-х, ссылаясь на 5% ошибку, т.е. на один ошибочный случай из 20 (именно его случай — экзаменовало 20 экспертов и один вполне мог ошибиться всего на 1 балл).

Основы модульно-рейтингового оценивания

Совершенно ясно, что использовать 20 экспертов для оценивания эффективности функционирования сложных систем практически невозможно. Например, даже на крупной кафедре 20 преподавателей не могут одновременно экзаменовать группу студентов. Однако возможно применить модификацию давно известного в теории стационарных процессов принципа эргодичности: судить о возможных состояниях системы, наблюдая ее во времени. Это означает, что вместо N экспертов мы должны пригласить одного эксперта, но он должен применить для оценивания качества функционирования системы N различных правил (тестов, зачетов, брифингов, собеседований и т.п. в случае образования), естественно, в разное время наблюдения системы.

По каждому из N правил проводится оценивание качества функционирования системы в выбранной номинальной шкале оценок, затем все N оценок складываются, а суммарные оценки переводятся в номинальную шкалу.

С точки зрения перевода суммарных оценок в номинальную шкалу мы имеем те же проблемы, что и при использовании N экспертов. Однако возникают новые проблемы, связанные с применением N различных правил или модулей экспертного оценивания качества функционирования системы:

- 1) поскольку система наблюдается во времени, то каждый экспертный модуль имеет дело, вообще говоря, с системами разного уровня сложности — система развивается. Следовательно, оценки, выставленные каждым модулем экспертного оценивания должны иметь свой специфический «вес», учитывающий сложность системы на текущем временном уровне;
- 2) в динамических и развивающихся системах их будущие состояния во многом предопределяются начальными состояниями. Следовательно, особое значение имеют первые модули экспертного оценивания, так называемого, входного контроля системы (принцип Тагучи). Специфика рассмотренных проблем применительно к образованию с очевидностью требует соблюдения следующих 4-х правил:
 - 1) изучение конкретных дисциплин должно строиться на основе изучения в различные «моменты» (интервалы) времени ряда M дидактических модулей (учебно-практических пособий) или разделов данных модулей, на которые опирается изучаемая дисциплина ($M \leq N$).
 - 2) изучение каждого дидактического модуля должно сопровождаться контролем качества его освоения с помощью модуля экспертного оценивания соответствующего « t »-го временного уровня ($t = 1, 2, \dots, T$, где $M \leq T \leq N$).
 - 3) контроль качества знаний каждым модулем экспертного оценивания последующего временного уровня « t » должен учитывать накопленные знания, проконтролированные модулем экспертного оценивания предыдущего временного уровня « $t-1$ ».
 - 4) первый модуль экспертного оценивания должен во многом предопределять как будущую траекторию обучения (какие дидактические модули изучать), так и набор модулей экспертного оценивания качества получаемых знаний.

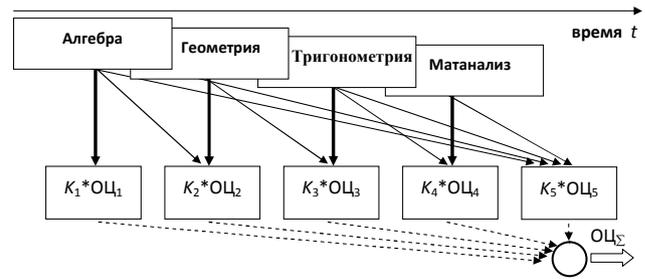


Рис. 2. Этапы модульно-рейтингового оценивания компетентностей

Перечисленные правила попросту говорят о том, что процесс контроля качества образования должен строиться в хронологическом порядке логической увязки изучаемых материалов — каждый последующий контроль строится на основе предыдущих. Не зря бытует всем известная поговорка «Повторенье — мать ученья».

Снова рассмотрим наглядный пример. Дисциплина «Математика» является фундаментальной основой практически всех специальных дисциплин, изучаемых в Университете. Условно предположим, что она базируется на изучении $M = 4$ основных дидактических модулей: 1 — «Алгебра»; 2 — «Геометрия»; 3 — «Тригонометрия»; 4 — «Матанализ». Будем проводить оценивание качества знаний по отдельным разделам данных дидактических модулей по 5-и балльной номинальной шкале. При этом выберем 5 модулей экспертного оценивания компетентностей соответствующих дидактических модулей и дисциплины в целом на 5-ти временных уровнях (см. рис. 2).

Введем числа Kt ($t = 1, 2, 3, 4, 5$), описывающие количество различных правил оценивания качества функционирования системы на « t »-ом временном уровне (различных заданий: контрольные работы, тесты, вопросы рубежного контроля). На рис. 2 «жирные» стрелки, идущие от дидактических модулей к модулям экспертного оценивания, означают оценивание знаний на каждом временном уровне по данным дидактическим модулям, «тонкие» стрелки означают учет знаний по модулям предыдущих временных уровней, а «пунктирные» стрелки соответствуют комплексному учету знаний.

Для комплексного оценивания воспользуемся рейтинговой технологией оценивания качества знаний на основе введения 100 балльной суммарной шкалы с ее последующим переводом в 5-и балльную номинальную шкалу.

В результате оценка, полученная в суммарной балльной шкале, будет равна

$$OЦ_{\Sigma} = K_1 * OЦ_1 + K_2 * OЦ_2 + K_3 * OЦ_3 + K_4 * OЦ_4 + K_5 * OЦ_5. (1)$$

Из условия согласованности наивысших оценок, т.е.

$$OЦ_1 = OЦ_2 = OЦ_3 = OЦ_4 = OЦ_5 = 5 \text{ и } OЦ_{\Sigma} = 100$$

следует, что (2)

$$K_1 + K_2 + K_3 + K_4 + K_5 = 20. (1)$$

Числа K_t , описывающие количество различных правил оценивания качества функционирования системы на «t»-ом временном уровне являются, по существу, весами, учитываемыми как сложность соответствующих дидактических модулей, так и значимость соответствующих временных уровней знаний. Это естественно, т.к. чем сложнее система, тем должно быть больше соответствующих правил.

Заметим, что получаемые с помощью модульно-рейтинговой технологии оценки $OЦ_{\Sigma}$ качества знаний имеют вероятность ошибки оценивания $P_{ош1} = (1/T)^2 = 0,0025$ или 0,25%, т.к. основаны на рассмотренной выше рейтинговой технологии при $K_{\Sigma} = 20$ различных правил оценивания качества знаний.

Однако здесь возникает другая серьезная проблема — как выбрать количество K_t различных правил оценивания качества функционирования системы на «t»-ых временных уровнях? Например, при равнозначных для дисциплины дидактических модулях и равномерном распределении знаний по временным уровням (конечно, это всего лишь начальное приближение) можно считать, что $K_1 = K_2 = K_3 = K_4 = K_5 = 4$. Но что делать в более реалистичных случаях, когда дидактические и экспертные модули имеют разную степень сложности. Как оценить эти степени и выбрать величины K_t ?

Оценивание степени сложности дидактических и экспертных модулей

Естественно, что для дисциплины «Математика», впрочем, как и других дисциплин, для различных специальностей будут различны и степени сложности дидактических модулей, а, следовательно, и модулей экспертного оценивания. Например, для экономических специальностей более важно, по сравнению с другими модулями, изучить такие модули, как: «Алгебра» и «Матанализ». Для инженерных специальностей, несомненно, важны также дидактические модули «Геометрия» и «Тригонометрия».

Возможны различные методы оценивания степеней сложности дидактических и экспертных модулей. Остановимся лишь на некоторых из них.

Метод эвристических суждений. Это самый простой метод, основанный на генерировании и утверждении специалистами кафедры некоторых балльных оценок B_t ($t = 1, 2, \dots, T$), например в 10-и балльной шкале, степеней сложности K_{Σ} дидактических и экспертных модулей. Пусть данные модули дисциплины «Математика» для некоторой специальности получили оценки $B_1 = 5, B_2 = 8, B_3 = 6, B_4 = 9, B_5 = 10$. Тогда возможно оценить степени сложности модулей или приоритеты

$$P_1 = B_1 / (B_1 + B_2 + B_3 + B_4 + B_5) = 5 / 38; (4)$$

$$P_2 = B_2 / (B_1 + B_2 + B_3 + B_4 + B_5) = 8 / 38;$$

$$P_3 = B_3 / (B_1 + B_2 + B_3 + B_4 + B_5) = 6 / 38;$$

$$P_4 = B_4 / (B_1 + B_2 + B_3 + B_4 + B_5) = 9 / 38;$$

$$P_5 = B_5 / (B_1 + B_2 + B_3 + B_4 + B_5) = 10/38;$$

$$P_1 + P_2 + P_3 + P_4 + P_5 = 1.$$

Перевод приоритетов (4) в числа K_t на основании (3) очевиден

$$K_1 = N * P_1 = 20 * 5 / 38 = 50 / 19 \cong 2,632; (5)$$

$$K_2 = N * P_2 = 20 * 8 / 38 = 80 / 19 \cong 4,211;$$

$$K_3 = N * P_3 = 20 * 6 / 38 = 60 / 19 \cong 3,158;$$

$$K_4 = N * P_4 = 20 * 9 / 38 = 90 / 19 \cong 4,737;$$

$$K_5 = N * P_5 = 20 * 10/38 = 100/19 \cong 5,263;$$

$$K_1 + K_2 + K_3 + K_4 + K_5 = 20,001$$

Как видно из (5) возникает некоторая ошибка в силу того обстоятельства, что все K_t ($t = 1, 2, 3, 4, 5$) должны быть целыми числами — число заданий экзаменуемым. Эту ошибку возможно отнести к неопределенности назначения балльных оценок B_t дидактических и экспертных модулей. Рассмотрим самый лучший — оптимистический случай, когда балльные оценки дают $N = 20$ специалистов кафедры, используя рассмотренную выше рейтинговую технологию. Тогда вероятность ошибки равна $P_{ош2} = (1/T)^2 = 0,0025$. Неопределенность можно поправить, округляя числа в (5) до целых, например: $K_1 = 3; K_2 = 4; K_3 = 3; K_4 = 5; K_5 = 5$. Тем самым достигается согласие между эвристическим оцениванием степеней сложности дидактических и экспертных модулей и технологией модульно-рейтингового оценивания качества знаний по дисциплине.

Оптимистическую результирующую вероятность ошибки оценивания знания по дисциплине можно оценить выражением

$$P_{ош\Sigma} = P_{ош1} + P_{ош2} - P_{ош1} * P_{ош2} \cong 2 (1 / K_{\Sigma})^2 = 0,0050. (6)$$

При этом, обучающемуся последовательно во времени дают 3, 4, 3, 5 и 5 разных заданий (вопросов, тестов, задач) со степенями сложности: $P_1 = 3/20$; $P_2 = 4/20$; $P_3 = 3/20$; $P_4 = 5/20$ и $P_5 = 5/20$ соответственно.

Метод согласованных суждений. Возможен более строгий способ оценивания степеней сложности дидактических и экспертных модулей. Для этого строят таблицу сравнений (см. таблицу 1).

Таблица 1

	A	B	C	D	...
A	1				
B		1			
C			1		
D				1	
...					1

Пусть даны модули A, B, C, D и т. д. Таблица сравнений строится по следующим правилам:

- расставляют модули A, B, C, D в порядке убывания степеней сложности (используя, например, рассмотренный выше метод эвристических суждений);
- если A и B одинаково важны, заносят в позицию (A,B) таблицы сравнений число 1;
- если A незначительно важнее B — число 3;
- если A значительно важнее B — число 5;
- если A явно важнее B — число 7;
- если A по своей значимости абсолютно превосходит B — число 9.
- в позицию (B,A) таблицы сравнений заносят числа, обратные основным числам.

В результате формируют обратно симметричную матрицу.

Числа 2, 4, 6 и 8 используют для облегчения компромиссов между оценками, слегка отличающимися от основных чисел.

Пусть, например, сравнивая модули дисциплины «Математика», мы получили следующую таблицу 2 сравнения.

Таблица 2

	«Экзамен»	«Матанализ»	«Алгебра»	«Арифметика»	«Геометрия»
«Экзамен»	1	5	6	7	9
«Матанализ»	1/5	1	3	5	8
«Алгебра»	1/6	1/3	1	3	4
«Арифметика»	1/7	1/5	1/3	1	2
«Геометрия»	1/9	1/8	1/4	1/2	1

Далее находят собственный столбец полученной матрицы (в данном случае матрицы $M = 5$ -го порядка) или столбец (0,27; 0,23; 0,20; 0,15; 0,15) приоритетов модулей, максимальное собственное значение λ_{\max} и коэффициент или индекс согласованности (ИС) суждений.

$$ИС = (\lambda_{\max} - M) / (M - 1) = (5,01 - 5) / (5 - 1) = 0,0025. (7)$$

Индекс согласованности дает вероятность ошибки определения приоритетов модулей $Р_{ош2} = ИС = 0,0025$. Следовательно, результирующую вероятность ошибки оценивания знаний по дисциплине можно оценить по аналогии с (6) выражением

$$Р_{ош\Sigma} = (1/T)^2 + 0,0025 - 0,0025 * (1/T)^2 \cong (1/T)^2 + 0,0025. (8)$$

Зависимость вероятности ошибок оценивания знаний по дисциплине от числа заданий и рассмотренных методов оценивания степеней сложности дидактических и экспертных моделей приведена в таблице 3.

Из таблицы 3 следует, что для значительного снижения вероятности ошибок следует использовать не менее 5-и модульно-рейтинговых заданий (соответственно две и одна ошибка на 25 экзаменуемых). Наибольшая точность (одна ошибка на 200 экзаменуемых) достигается, когда $K\Sigma = 20!$

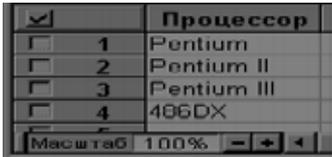
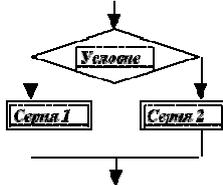
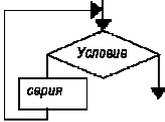
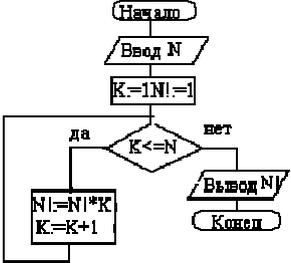
Таблица 3

Число заданий $K\Sigma$ (тестов, задач)	Вероятность $Р_{ош\Sigma}$	
	Метод эвристических суждений	Метод согласованных суждений
20	0,5% (1 случай из 200)	0,5% (1 случай из 200)
10	2,2% (2 случая из 100)	1,25% (1 случай из 100)
5	8% (2 случая из 25)	4,25% (1 случай из 25)
4	12,5% (1 случай из 10)	6,5% (1 случай из 20)
3	22% (1 случай из 5)	10% (1 случай из 10)

Пример реализация оценивания рейтинга при тестировании.

Пример фрагментов специального листа (задания 10 и 11), соответствующего интерфейсу программы для оценивания рейтинга на основе тестовых заданий приведен в таблице 4. Как видно из таблицы, для оценивания рейтинга каждое задание включает три подзадания (вопросы и задачи), с предполагаемыми тремя решениями.

Таблица 4

Фрагмент заданий для проведения компетентностного модульно-рейтингового контроля				
10	Определить истинность составного высказывания: “(2 x 2 = 4 или 3 x 3 = 10) и (2 x 2 = 5 или 3 x 3 = 9)”	ложно		
		истинно		
		не ложно и не истинно		
	Какой вид примет содержащая абсолютную и относительную ссылку формула, записанная в ячейке C1, после ее копирования в ячейку C2?		=A\$1*B2	
			=A\$1*B1	
			=A\$2*B2	
В какой последовательности расположатся записи в базе данных после сортировки по возрастанию в поле Память?		1, 2, 3, 4		
		4, 3, 2, 1		
		4, 1, 2, 3		
11	Алгоритмическая структура какого типа изображена на блок-схеме?		цикл	
			ветвление	
			подпрограмма	
	Алгоритмическая структура какого типа изображена на блок-схеме?		цикл	
			ветвление	
			линейная	
При заданных исходных данных (N=3) определите результат выполнения алгоритма вычисления факториала, изображенного в виде блок-схемы.		9		
		6		
		12		

© NOTA BENE (ООО «НБ-Медиа») www.nbpublish.com

При выполнении заданий в программной оболочке расчет результата (рейтинга) выполняется автоматически. При ручном оценивании используется трафарет.

При решении всех трех подзадач задания обучающемуся выставляется оценка 5, при решении двух — 4, одной 3, ни одной — 2.

Максимальный балл по 20 заданиям — 100, минимальный зачетный — 60.

Реальные системы балльно-рейтингового оценивания

В Московском экономико-статистическом институте (МЭСИ) принята следующая шкала оценивания как академической активности обучающихся, так и их компетентностей:

- академическая активность — 14 баллов; (посещение лекций, лабораторных занятий)

- выполнение практических задач — 36 баллов; (лабораторных работ и домашних заданий)
- рубежное тестирование — 40 баллов;
- итоговое тестирование — 10 баллов.

Карта рейтинга по дисциплине представлена в таблице 5.

Карта рейтинга по дисциплине представлена в таблице 5.

Таблица 5

Карта балльно-модульного рейтинга по дисциплине

Виды учебной работы	Максимальный балл	Зачётный балл
Посещение лекций	7	5
Посещение лабораторных занятий	7	5
Выполнение лабораторных работ, включая допуск и отчет по работе	31	20
Промежуточная и рубежная аттестация, включая активность обучающегося (рубежное и зачетное тестирование, олимпиады, участие в научно-исследовательской работе)	40	30
Экзамен:		
– итоговое тестирование	10	5
– решение задачи	5	5
Всего по дисциплине:	100	65

Окончательные итоговые оценки по дисциплине формируются на основании правил:

- 3 — 65 ≤ рейтинги < 75 баллов;
- 4 — 76 ≤ рейтинги < 85 баллов;
- 5 — 86 ≤ рейтинги ≤ 100 баллов.

В Московском государственном университете технологий и управления им. К.Г. Разумовского приняты следующие правила перевода рейтингов в номинальные оценки:

- 3 — 60 ≤ рейтинги < 70 баллов;
- 4 — 70 ≤ рейтинги < 90 баллов;
- 5 — 90 ≤ рейтинги ≤ 100 баллов.

Если по результатам работы по дисциплине в семестре обучающийся не набрал до итогового контроля минимально допустимого количества баллов — 65/60, ему выставляется итоговая оценка «неудовлетворительно». В этом случае обучающемуся предлагается изучить дисциплину повторно. В случае успешной пересдачи обучающийся может получить только оценку «удовлетворительно».

Максимальное количество баллов, которое обучающийся может получить на экзамене, равно 15 (итоговое тестирование — 10 баллов и задача, решаемая на компьютере — 5 баллов).

Если в ходе выполнения всех заданий обучающийся набрал до итогового контроля 86 баллов, то ведомость и зачетную книжку обучающегося выставляется оценка «отлично».

Мониторинг качества обучения проводится в форме выставления преподавателями кредитных и компетентностных баллов в рейтинг-листке обучающегося после освоения им каждой темы (модуля) дисциплины.

Рекомендуется распределять баллы по модулям курса, в соответствии с их сложностью.

Модульная карта рейтинга по дисциплине

Модульная карта рейтинга по дисциплине «Информационная технология управления» кафедры информационных технологий МГУТУ представлена в таблице 6.

Заключение

1. Рейтинговая технология оценивания качества функционирования сложных систем, в том числе качества знаний в системе образования, путем выставления, так называемого, рейтинга в многобалльной шкале (например, 100 балльной) и последующего его пересчета в оценку по номинальной шкале (например, 5-и балльной) значительно снижает вероятность ошибки оценивания.

2. Для выбранного случая 5-и и 100 балльных шкал оценивания на уровне вероятности ошибки $P_{0ш} = (5/100)^2 = 0,0025$ (не более одной ошибки на четыреста случаев) обоснованы следующие правила перевода рейтингов в оценки:

- 2 — рейтинги до 48 баллов;
- 3 — рейтинги лежат в пределах 52 ÷ 68 баллов;
- 4 — рейтинги лежат в пределах 72 ÷ 88 баллов;
- 5 — рейтинги лежат в пределах 92 ÷ 100 баллов.

3. В диапазонах 48 ÷ 52, 68 ÷ 72 и 88 ÷ 92 баллов обоснованные оценки качества функционирования системы (например, знаний обучаемых) на уровне вероятности ошибки $P_{0ш} = 0,0025$ отсутствуют. Здесь необходимо проводить дополнительное оценивание, например, в случае образования — переэкзаменовку, если студент пожелает повысить свою оценку, или принимать решения: «в пользу студента», «в пользу преподавателя».

4. Модульно-рейтинговая технология оценивания качества функционирования сложных систем в процессе их развития, что типично для системы образования, при использовании 5-и и 100 балль-

Таблица 6

Модульная карта рейтинга по дисциплине «Информационные технологии управления»						
Название темы (модуля)	Лекционные занятия	Акад балл	Лабораторные занятия (решение задач)	Акад балл	Пром. и рубежн. аттестация (балл)	Максимальный/минимальный балл модуля
Модуль 1						
ИТ автоматизированного управления.	Автоматизированное управление.	0,5	6	1,5	10	18,9/12
	Подготовка принятия управленческих решений.	0,5				
	Вероятностное и нечеткое описания состояний ОУ.	0,4				
Модуль 2						
ИТ автоматического управления.	Автоматическое управление.	0,5	6	1,5	10	18,9/12,5
	Цифровые системы управления.	0,5				
	Экспертные корреляционно-экстремальные системы управления.	0,4				
Модуль 3						
ИТ управления финансами, производством и бизнесом.	Основные задачи управления финансовыми операциями.	0,4	7	2	10	19,4/12,5
	Управление производственными и коммерческими процессами.	0,5				
	Решение оптимизационных задач управления производственными и коммерческими процессами.	0,5				
Модуль 4						
Информационные технологии прогнозирования состояний объектов управления.	Оперативное и стратегическое планирование и прогнозирование.	0,4	6	1	5	13,4/12,5
	Условные математические ожидания. Оптимальный стохастический прогноз.	0,5				
	Синтез предикторов.	0,5				
Модуль 5						
Информационные технологии управления рисками.	Управление рисками на основе их вероятностных и нечетких параметрических характеристик.	0,5	6	1	5	13,4/12,5
	Управление рисками при инвестировании.	0,5				
	Статистическая теория принятия управленческих решений.	0,4				
Экзамен:						
– итоговое тестирование					10	10
– решение задачи			5			5
Всего:		7	36	7	50	100/65

© NOTA BENE (ООО «НБ-Медиа») www.nbpublish.com

ных шкал предполагает декомпозицию систем на $M \leq 20$ последовательно связанных системных (дидактических в случае системы образования) модулей и использование до $K_{\Sigma} \leq 20$ различной степени сложности правил (экспертных модулей) оценивания качества функционирования систем на различных временных интервалах ($M \leq K_{\Sigma} \leq 20$).

5. При $K_{\Sigma} = 20$ вероятность ошибки оценивания $P_{0ш} = 2(1/K_{\Sigma})^2 = 0,005$ (не более одной ошиб-

ки на двести случаев), а при $K_{\Sigma} < 20$ вероятность ошибки оценивания заключена в пределах $(1/K_{\Sigma})^2 + 0,0025 \leq P_{0ш} \leq 2(1/K_{\Sigma})^2$, обусловленных обоснованностью выбора степеней сложности экспертных модулей. Данные степени сложности, в свою очередь, зависят от сложности соответствующих системных (дидактических в случае системы образования) модулей. Так, например, для $K_{\Sigma} = 5$ получим: «1 ошибка на 25 случаев» $\leq P_{0ш} \leq$ «2 ошибки

на 25 случаев»; для $K_{\Sigma} = 10$ получим: «1 ошибка на 100 случаев» $\leq P_{ош} \leq$ «2 ошибки на 100 случаев».

Вполне естественно, что увеличение количества K_{Σ} правил оценивания качества функционирования систем на различных временных интервалах оправдано лишь в том случае, если цена ошибки велика. Так, например, в развитых

странах чрезвычайно велики цены ошибок в системах правозащиты, здравоохранения, образования (огромные штрафы страховых компаний), что обуславливает высокие цены соответственно судопроизводства, лечения, обучения. Эти высокие цены отражают не только погоню за прибылью, но и качество уровня жизни!

Список литературы:

1. Буромский И.В., Романенко Г.Х. Методическая основа применения балльно-рейтинговой системы для контроля и оценки освоения материала учебной дисциплины // *ВВ: Российское полицейское право*. 2014. № 1. С. 108–124. (http://www.e-notabene.ru/pm/article_9913.html).
2. Грязнова Е.В. Идентификация человека в информационной реальности // *Психология и психотехника*. 2013. № 4. С. 371–379.
3. Евсевичева И.В., Любимова Г.А. Самооценивание в профессиональной сфере // *Психология и психотехника*. 2011. № 9. С. 49–59.
4. Краснов А.Е., Красников С.А., Валентинов В.А., Валентинова Н.И. и др. Отчет по НИР: «Разработка принципов метрологического обеспечения менеджмента качества вузов на основе совокупности базовых показателей учебной, научной и инновационной деятельности». М.: МГУТУ, 2009–2010. 60 с. (Номер государственной регистрации НИР: 3.4.1/6057).
5. Краснов А.Е., Кузнецова Ю.Г. и др. Проблемы количественного оценивания качества знаний в высшем образовании // *Стандарты и качество*. 2007. № 11.
6. Краснов А.Е., Кузнецова Ю.Г. Информационные технологии оценивания компетенций студентов на основе сложного рейтинга // *Вестник РУДН. Серия «Информатизация образования»*. 2009. № 2.

References (transliteration):

1. Buromskii I.V., Romanenko G.Kh. Metodicheskaya osnova primeneniya ball'no-reitingovoi sistemy dlya kontrolya i otsenki osvoeniya materiala uchebnoi distsipliny // *NB: Rossiiskoe politseiskoe pravo*. 2014. № 1. S. 108–124. (http://www.e-notabene.ru/pm/article_9913.html).
2. Gryaznova E.V. Identifikatsiya cheloveka v informatsionnoi real'nosti // *Psikhologiya i psikhotehnika*. 2013. № 4. S. 371–379.
3. Evsevicheva I.V., Lyubimova G.A. Samootsenivanie v professional'noi sfere // *Psikhologiya i psikhotehnika*. 2011. № 9. S. 49–59.
4. Krasnov A.E., Krasnikov S.A., Valentinov V.A., Valentinova N.I. i dr. Otchet po NIR: «Razrabotka printsipov metrologicheskogo obespecheniya menedzhmenta kachestva vuzov na osnove sovokupnosti bazovykh pokazatelei uchebnoi, nauchnoi i innovatsionnoi deyatel'nosti». M.: MGUTU, 2009–2010. 60 s. (Nomer gosudarstvennoi registratsii NIR: 3.4.1/6057).
5. Krasnov A.E., Kuznetsova Yu.G. i dr. Problemy kolichestvennogo otsenivaniya kachestva znaniy v vysshem obrazovanii // *Standarty i kachestvo*. 2007. № 11.
6. Krasnov A.E., Kuznetsova Yu.G. Informatsionnye tekhnologii otsenivaniya kompetentsii studentov na osnove slozhnogo reitinga // *Vestnik RUDN. Seriya «Informatizatsiya obrazovaniya»*. 2009. № 2.