

На правах рукописи



Родионов Алексей Владимирович

**РАЗРАБОТКА МОДЕЛЕЙ, МЕТОДОВ И ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ
ДЛЯ ОЦЕНКИ КОМПЕТЕНЦИЙ УЧАЩИХСЯ ВУЗОВ**

05.13.01 – Системный анализ, управление и обработка информации
(региональные народнохозяйственные комплексы)

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Иркутск, 2016

Работа выполнена на кафедре информатики и кибернетики ФГБОУ ВО
«Байкальский государственный университет»

- Научный руководитель: Братищенко Владимир Владимирович,
кандидат физико-математических наук, доцент,
заведующий кафедрой информатики и кибернети-
ки ФГБОУ ВО «Байкальский государственный
университет», г. Иркутск
- Официальные оппоненты: Петров Александр Васильевич, доктор
технических наук, профессор, профессор кафедры
автоматизированных систем ФГБОУ ВО «Иркут-
ский национальный исследовательский техниче-
ский университет», г. Иркутск
- Берестнева Ольга Григорьевна,
доктор технических наук, профессор, профессор
кафедры прикладной математики ФГАОУ ВО
«Национальный исследовательский Томский по-
литехнический университет», г. Томск
- Ведущая организация: ФГБОУ ВО «Иркутский государственный универ-
ситет», г. Иркутск

Защита диссертации состоится 12 апреля 2016 г. в 15-00 часов на заседании дис-
сертационного совета Д 212.070.07 на базе ФГБОУ ВПО «Байкальский государственный
университет экономики и права» по адресу: 664003, г. Иркутск, ул. Карла Маркса, д.24,
корп. 9, зал заседаний ученого совета БГУЭП

С диссертацией можно ознакомиться в библиотеке и на официальном сайте
(www.isea.ru) ФГБОУ ВО «Байкальский государственный университет» по адресу
664003, г. Иркутск, ул. Ленина, 11, БГУЭП, корпус 2, аудитория 101. (до 29.10.15 г.
ФГБОУ ВПО «Байкальский государственный университет экономики и права»).

Отзывы на автореферат направлять по адресу: 664003, г. Иркутск, ул. Ленина, 11,
БГУЭП, ученому секретарю диссертационного совета Д 212.070.07.

Автореферат разослан _____

Ученый секретарь диссертационного совета
Д 212.070.07, кандидат технических наук, доцент



Т.И. Ведерникова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Одной из проблем внедрения компетентностного подхода в образовательный процесс ВУЗов является оценка компетенций. Оценка компетенции – способности применения знаний, умений, навыков и личностных качеств для успешной деятельности – исключительно сложная задача. Во многом сложность обусловлена тем, что компетенция является латентной чертой человека, напрямую не диагностируемой, и ее оценка возможна только по «косвенным» признакам – результатам выполнения задач (индикаторов), характеризующих развитие компетенции. Различные подходы к внедрению компетенций в образовательный процесс, в т.ч. и вопросы квалиметрии рассмотрены в работах Д. Хаймса, Д.К. МакКлелланда, С. и Л. Спенсеров, В.И. Байденко, И.А. Зимней, А.В. Хуторского, А.И. Субетто, Ю.Г. Татур, А.С. Казаринова, В.А. Богословского, В. Хутмаера, Л.Л. Редько, А.В. Шумаковой и некоторых других. На основе анализа рассмотренной в работе литературы можно условно выделить два подхода к оценке компетенций.

Первый подход предполагает разработку инновационных средств оценки (индикаторов) компетенции, и включает в себя такие методы, как портфолио, деловая игра, тесты практических умений, кейс измерители и т.п. Они хорошо зарекомендовали себя в профессиональной среде (например, в рекрутинговых агентствах), однако в образовательной сфере широкого распространения не получили. Это связано со сложностью разработки, субъективизмом оценки и интерпретации результатов. Второй подход связан с использованием классических, общепринятых в образовательной сфере измерительных средств: лабораторные, практические, курсовые работы, тесты, коллоквиумы и т.д.

Согласно ФГОС третьего поколения оценка компетенций осуществляется с помощью фондов оценочных средств - комплектов методических и контрольных материалов, предназначенных для оценивания знаний, умений, навыков и компетенций на разных стадиях обучения студентов. Однако не решенной остаётся проблема выбора и обоснования математической модели обработки первичных баллов, полученных при использовании оценочных средств. Модели, основанные на понятии среднего значения (а именно они обычно и используются в сфере образования), как правило, не учитывают различий методик преподавания и оценивания заданий, а полученные результаты измерений являются нелинейными. К тому же они не обеспечивают «объективность» оценивания, т.е. результаты оценки зависят от используемого набора заданий, следствием чего является невозможность прямого сравнения оценок компетенций студентов, полученных по разным наборам заданий.

В свете вышеизложенного, разработка адекватных и объективных процедур оценки компетенции как основной описательной характеристики освоения студентом учебной программы, является актуальной задачей.

Цель исследования состоит в разработке моделей, методов, алгоритмов, методического и программного обеспечения для оценки компетенций студентов ВУЗов.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Провести исследование учебно-воспитательного процесса в университете с целью моделирования процесса формирования компетенций, предложить метод определения индикаторов сформированности компетенций.

2. Разработать математическую модель оценки компетенций и методику проведения измерения компетенций.

3. Реализовать предложенные модели, методы и методики в виде математического, алгоритмического и программного обеспечения, используемого для формирования оценки сформированности компетенций в ВУЗе.

4. Проверить валидность предложенной методики и адекватность моделей на примере оценки компетенций какого-либо направления подготовки.

Объект исследования: мониторинг образовательной деятельности студентов.

Предмет исследования: моделирование процессов формирования и оценки компетенций.

Методы исследования: для решения поставленных задач использовались методы системного анализа, математического моделирования, когнитивного моделирования, теории измерения латентных переменных и теории нечетких множеств.

Научная новизна работы заключается в следующих положениях:

1. Предложены модель процесса формирования компетенции (карта компетенции), метод ее построения и верификации.

2. Разработаны новая латентная модель экзаменационных оценок и метод оценки компетенций на основе теории измерения латентных переменных.

3. Разработана методика оценки сформированности компетенций студентов ВУЗов, статистической проверки адекватности и валидности результатов оценивания.

4. Разработано математическое и алгоритмическое обеспечения для решения задач построения карт компетенций, оценки компетенций студентов, расчетов их статистических характеристик.

Практическая значимость работы заключается в разработанных процедурах обработки данных, реализованных в виде программного продукта СПКООП, для решения поставленных в работе задач. Данные процедуры прошли практическую проверку в рамках научно-исследовательской работы «Современные информационно-телекоммуникационные технологии в управлении социально-экономическими и технологическими процессами», выполняемой на кафедре «Информатики и Кибернетики» БГУЭП в 2013-2014 гг. Разработаны и апробированы карты компетенций по направлению подготовки 230700 «Прикладная информатика», профиль подготовки «Информационные системы и технологии в управлении», квалификация (степень) бакалавр.

Апробация работы. Основные положения проведенного исследования докладывались и всесторонне обсуждались на VII международной научно-практической конференции «Теоретические и прикладные аспекты современной науки» г. Белгород, 2015 г., конференции ППС БГУЭП в 2015 г., ежегодных конференциях аспирантов кафедры «Информатики и кибернетики» (2010-2014 гг.). Получены два свидетельства о регистрации программных продуктов. Результаты

диссертации отражены в отчетах о научно-исследовательских работах «Совершенствование организации и управления научно-учебным процессом» за 2011-2015 гг. и «Современные информационно-телекоммуникационные технологии в управлении социально-экономическими и техническими процессами» за 2013-2015 гг.

Публикации и личный вклад автора. Основные результаты диссертации отражены в 8 научных работах, в том числе 3 работы в рецензируемых научных журналах, определенных ВАК Минобрнауки РФ, 2 свидетельства о регистрации программы для ЭВМ, 1 глава в коллективной монографии, написанная автором лично.

Структура и объем работы. Диссертационная работа состоит из введения, четырех глав, заключения, списка литературы из 138 наименований и 6 приложений, содержащих 66 страниц. Общий объем работы без приложений составляет 162 страницы, включая 26 таблиц, 34 рисунка, 147 формул.

СОДЕРЖАНИЕ РАБОТЫ

Во введении раскрывается актуальность выбранной темы, цели и содержания поставленных задач, выделяются объект и предмет исследования, указывается практическая и теоретическая значимость полученных результатов, формулируются результаты, выносимые на защиту.

В первой главе проводится обзор научной литературы, посвященной компетентности, компетенциям и компетентностному подходу. Рассматриваются системные закономерности формирования компетенций в рамках учебного процесса.

Проведен анализ различных видов контроля (аттестаций): текущего, промежуточного, итогового. Показано, что для оценки компетенции следует использовать оценочные средства промежуточного контроля, т.к. текущий контроль является слишком «фрагментарным», а итоговый (проводящийся всего один раз – в конце обучения) не может дать оперативную информацию о сформированности компетенций. Он может использоваться как своего рода «экспертный» контроль для подтверждения сформированности тех или иных компетенций.

Сделан обзор существующих моделей, которые могут быть применены для обработки результатов выполнения оценочных средств (первичных баллов): модели среднего значения, нечеткие модели и модели теории измерения латентных переменных (IRT). Рассмотрены их преимущества и недостатки в контексте оценки компетенций.

В завершении первой главы делается постановка задачи исследования

Вторая глава посвящена описанию предлагаемых в работе моделей, методов их построения. Для построения модели процесса формирования компетенции предлагается использование когнитивного моделирования. В этом случае модель представляется в виде когнитивной карты, получившей в работе название карты компетенции. Предлагается разработанный автором экспертно-статистический метод построения карт компетенций.

Для оценки сформированности компетенций автором предложен подход, основанный на теории латентных переменных (IRT). Тогда оценку сформированности компетенции можно определить как некоторую величину, характеризующую вероятность правильного выполнения какого-либо задания (работы) в зависимости от его сложности.

Введено основное требование, которому должна удовлетворять модель оценки компетенции – объективность измерения. Под объективностью понимается независимость полученной оценки от инструмента проведения оценки и от того, кто ее проводит.

Показано, что для «дихотомических» заданий, результаты выполнения которых можно выразить в двухбалльной шкале (например, выполнено/не выполнено), введенному требованию удовлетворяет однопараметрическая IRT-модель, а для политомических (более двух результатов) – модели частичного и рейтингового оценивания.

Исследование, проведенное на реальных эмпирических данных, выявило имеющуюся зависимость выставления оценок промежуточной аттестации от дисциплин. Для учета этой зависимости в работе предложена новая модель – латентная модель экзаменационных оценок, которая продемонстрировала лучшее соответствие эмпирическим данным, чем рассмотренные «существующие» модели.

В третьей главе рассматриваются математическое, алгоритмическое и методическое обеспечение оценки компетенций. Для оценки параметров рассмотренных в работе моделей предложено использование метода максимального правдоподобия.

Проведен анализ шкал для представления оценки компетенции. Помимо шкалы логитов предложено использование рейтинговой стобалльной и уровневой лингвистической шкал. Показаны их сильные и слабые стороны, сферы применения, разработаны алгоритмы работы с ними.

Для статистического анализа результатов измерения латентных параметров предложено использование ряда статистик. В англоязычной литературе они называется «goodness of fit», или «fit statistics». Данные статистики позволяют проверить не только соответствие заданий модели (item fit), но и соответствие ответов испытуемых модели (person fit).

Для организации процесса оценки компетенций в ВУЗе предложена новая методика оценки сформированности компетенций. Описан разработанный программный комплекс, автоматизирующий основные операции согласно предложенной методике оценки сформированности компетенций.

В четвертой главе рассматриваются практические задачи оценки компетенций на примере БГУЭП по направлению 230700 «Прикладная информатика», профиль подготовки «Информационные системы и технологии в управлении» с помощью разработанного программного средства и предложенной методики.

Были разработанные карты компетенций, проведено их исследование, определены конструкторы, получены оценки компетенций для студентов. Разработанные карты могут использоваться в дальнейшем для оценки компетенций новых студентов. Результаты статистического анализа показали хорошую согласованность исходных данных разработанной в диссертации латентной модели экзаменационных оценок, и как следствие, хорошую применимость методики.

Проведено сравнительное исследование точности оценки параметров латентной модели экзаменационных оценок. Показано, что ошибки, возникающие при измерении компетенции, минимальны и в полной мере соотносятся с рассматриваемыми в работе теоретическими исследованиями.

На реальных данных проведено исследование возможного влияния многомерности исходных данных на оценку компетенции. Результаты показали, что, хотя введение в модель дополнительного параметра, учитывающего влияние многомерности данных, в ряде случаев обеспечивает лучшую точность измерения, однако приводит к существенному увеличению сложности расчетов и усложняет интерпретацию результатов, а выигрыш в точности с практической точки зрения незначителен.

В заключении описаны основные результаты и решенные в диссертации задачи.

ОСНОВНЫЕ ПОЛОЖЕНИЯ ДИССЕРТАЦИОННОГО ИССЛЕДОВАНИЯ

1. Предложены модель процесса формирования компетенции (карта компетенции), метод ее построения и верификации.

Базовой задачей для проведения оценки компетенций студентов является построение модели процесса ее формирования. Перспективным подходом к решению данной задачи является использование когнитивного моделирования и представление модели в виде когнитивной карты – ориентированного графа, вершинами которого являются факторы, влияющие на формирование компетенции, а ребрами – связи между ними. Однако подавляющее большинство методов, используемых для когнитивного моделирования, являются экспертными, что вносит определенную субъективность, а достоверность и надежность результатов моделирования зависят от компетентности эксперта(ов). Для устранения этих недостатков в диссертации предлагается новый метод, включающей в себя элементы когнитивного моделирования и статистического анализа. Суть метода состоит в следующем: моделирование начинается с определения учебных дисциплин, влияющих на формирование исследуемой компетенции. После этого определяется последовательность изучения предметов. Для верификации модели применяется статистический анализ, который позволяет подтвердить (или опровергнуть) наличие связей дисциплина – компетенция при условии наличия эмпирических данных (имеется база оценок студентов по заданиям, которые они выполнили при изучении дисциплин). Такой анализ проводится с использованием теории латентных переменных (IRT).

Совершенно очевидно, что на формирование компетенций влияет множество факторов, но в случае обучения в ВУЗе «основными» факторами будут учебные дисциплины, учебная и дипломная практика, и т.п., определенные в учебном плане. Назовем модель картой компетенции и выразим ее в следующем виде:

$$g = \{D^{(g)}, \Omega\}, \quad (1)$$

где $D^{(g)}$ – элементы (дисциплины), влияющие на компетенцию g ; Ω – бинарное отношение на множестве $D^{(g)}$, которое задает набор связей между его элементами.

Следует отметить, что такое представление модели позволяет формализовать суждение об оценке компетенции следующим образом: если студент владеет компетенцией на некотором уровне, то он с некоторой вероятностью, зависящей от уровня освоения компетенции, выполнит учебные задания по дисциплинам, входящим в карту компетенций. Этапы построения карты компетенции:

1. Определяется набор дисциплин D , входящих в учебный план. Экспертная группа, посредством коллективного опроса, определяет подмножество

дисциплин $D^{(g)}$ из множества D , изучение которых, по их мнению, служит формированию исследуемой компетенции g .

2. Строится последовательность изучения предметов. Для этого предлагается использовать систему нестрогого ранжирования, по причине того, что некоторые предметы могут изучаться параллельно, не зависимо друг от друга. Объекты ранжирования – дисциплины, формирующие компетенцию. Основание ранжирования – порядок следования изучения дисциплины формирования компетенции. В качестве показателя согласованности ранжировок, полученных от различных экспертов, используется коэффициент конкордации Кендела (для нестрогого ранжирования). Результирующее (итоговое) ранжирование должно удовлетворять условиям, сформулированным К. Эрроу, и одним из немногих способов получения итогового ранжирования, практически полностью удовлетворяющий этим условиям (четырёх из пяти условий), является медиана Камени. Именно она используется для вычисления итогового ранжирования в случае согласованности ранжировок экспертов.

3. Верификация карты компетенции осуществляется следующим образом: берётся набор заданий j , которые выполняют студенты при изучении дисциплин, формирующих компетенцию. Пусть y_{ij} – ответ i -го студента на j -ое задание по проверяемым дисциплинам, $i = \overline{1, N}; j = \overline{1, M}$, N – количество студентов, M – количество заданий. Используя рассмотренные в диссертации модели теории латентных переменных, можно определить E_{ij} – «ожидаемый» ответ i -го студента на j -ое задание. Остатки между наблюдаемым баллом и ожидаемым ответом согласно модели можно вычислить по формуле: $\xi_{ij} = y_{ij} - E_{ij}$, а формулы вычисляемых статистических критериев (item fit – анализ) представлены в таблице 1.

Таблица 1

Применяемые статистики

Формула	Обозначение и комментарии	Рекомендуемые / допустимые значения
$Q3_{jj'} = \frac{\sum_{i=1}^N (\xi_{ij} - \bar{\xi}_j)(\xi_{ij'} - \bar{\xi}_{j'})}{\sqrt{\sum_{i=1}^N (\xi_{ij} - \bar{\xi}_j)^2} \sqrt{\sum_{i=1}^N (\xi_{ij'} - \bar{\xi}_{j'})^2}}$ <p>где $\bar{\xi}_j$ и $\bar{\xi}_{j'}$ есть средние значения ξ_{ij} и $\xi_{ij'}$ соответственно</p>	<p>Q3 статистика. Анализ на локальную независимость</p>	<p>$\pm 0.2 \div \pm 0.25$ от ожидаемого значения</p>
$OM_j = \frac{1}{N} \sum_{i=1}^N \frac{\xi_{ij}^2}{D[\xi_{ij}]} = \sum_{i=1}^N \frac{z_{ij}^2}{N},$ $z_{ij} = \frac{y_{ij} - E_{ij}}{\sqrt{D[y_{ij}]}}, \quad \frac{\xi_{ij}}{\sqrt{D[y_{ij}]}}$	<p>OUTFIT MEANSQ Базовый анализ на соответствие эмпирических данных модели</p>	<p>0,4 ÷ 1,4 / 0,3 ÷ 1,7</p>
$OZ_j = (\sqrt[3]{OM_j} - 1) \left(\frac{3}{\sigma_j} \right) - \left(\frac{\sigma_j}{3} \right),$ <p>где $\sigma_j = \frac{\sqrt{\sum_{i=1}^N \frac{1}{D[y_{ij}]}}}{N}$</p>	<p>OUTFIT ZSTD. Чувствительность к аномальным ответам, например, таким, как неправильные ответы на простые задания испытуемыми с высокими уровнями способностей.</p>	<p>0,4 ÷ 1,4 / 0,3 ÷ 1,7</p>
$IM_j = \frac{\sum_{i=1}^N \xi_{ij}^2}{\sum_{i=1}^N D[y_{ij}]}$	<p>INFIT MEANSQ. Меньшее влияние многомерности, количества заданий и размерности выборки, чем OUTFIT</p>	<p>-2,5 ÷ +2,5 / -3 ÷ +2,5</p>

$IZ_j = \left(\sqrt[3]{IM_j} - 1 \right) \left(\frac{3}{\sigma_j} \right) - \left(\frac{\sigma_j}{3} \right),$ $\sigma_j = \frac{\sqrt{\sum_{i=1}^N D[y_{ij}] - 4 \sum_{i=1}^N (D[y_{ij}])^2}}{\sum_{i=1}^N D[y_{ij}]}$	INFIT ZSTD. Меньшее влияние многомерности, количества заданий и размерности выборки, чем OUTFIT	-2,5 ÷ +2,5/ -3 ÷ +2,5
--	---	---------------------------

Если значения статистик для всех j -ых заданий не превышают рекомендуемых, то делается вывод о том, что предложенные экспертами предметы действительно влияют на формирование компетенции, в связи с тем, то эти задания «диагностируют» компетенцию.

Результатом применения данного алгоритма является карта компетенции, которая может быть выражена в виде графа или таблицы. В связи с тем, что процесс освоения каждой дисциплины завершается зачетом или экзаменом, карта компетенции, в случае использования для оценки компетенции оценочных средств промежуточной аттестации, однозначно определяет конструкт оценки компетенции. Также карты компетенций могут использоваться для построения траектории изучения дисциплин – компетентностно-ориентированного учебного плана.

2. Разработаны новая латентная модель экзаменационных оценок и метод оценки компетенций на основе теории измерения латентных переменных.

В работе предлагается метод оценки компетенций, основанный на теории измерения латентных переменных, в англоязычных работах известной как Item Response Theory (IRT). Основной предмет применения математических моделей IRT – оценивание вероятности правильного ответа испытуемых разной подготовленности на задания различной трудности. Само задание рассматривается как некоторая самостоятельная единица, обладающая определенным набором параметров. Предполагается, что ответ на задание есть реализация некой случайной величины, распределение которой зависит от параметров как самого испытуемого, так и задания.

В общем случае ответы на выполненные задания (первичные баллы) можно выразить в виде матрицы:

$$Y = \{y_{ij_d}\}, \quad (2)$$

где y_{ij_d} – оценка, полученная i -ым студентом за выполненное j -ое задание по d -ой дисциплине; $i = \overline{1, N}$, N – общее количество студентов, обучающихся по какому-либо направлению подготовки; $j = \overline{1, M}$, M – количество заданий в фонде оценочных средств.

Освоение каждой компетенции, определённых в ФГОС, в ВУЗе происходит в результате изучения некоторого конкретного набора дисциплин, который может быть определён при построении карты компетенции. Тогда из матрицы Y можно выразить подматрицы первичных баллов по каждой компетенции $Y^{(g)} = \{y_{ij_d'}\} \in Y; d' \in D^{(g)}$. В случае использования заданий промежуточного контроля (когда каждой дисциплине соответствует только одно индикаторное задание – экзамен (зачёт)), матрицы первичных баллов для g -ой компетенции можно упростить и записать в виде $Y^{(g)} = \{y_{ij}\}$.

Пусть определен конструктор (набор индикаторных заданий) для оценки некоторой латентной черты человека, а y_{ij} – ответ i -го испытуемого на j -ое задание конструктора, θ_i – параметр, характеризующий исследуемую латентную черту i -го человека, δ_j – параметр, характеризующий j -ое задание. Тогда обобщенно IRT-модель можно записать в следующем виде:

$$P(y_{ij}|\theta_i, \delta_j) = f(\theta_i, \delta_j, y_{ij}), \quad (3)$$

Для оценки параметров моделей вида (3) в работе предлагается использовать метод максимального правдоподобия, который позволяет найти эффективные и состоятельные оценки неизвестных параметров.

Тогда суть предлагаемого метода оценки компетенций состоит в следующем: имея набор данных – первичные баллы студентов за выполнение заданий промежуточной аттестации y_{ij} и вероятностную модель вида (3), используя метод максимального правдоподобия, можно получить значения (оценки) параметров модели: оценки сформированности компетенций студентов θ_i и параметров заданий δ_j .

На основе анализа различных видов контроля, определенных во ФГОС, для оценки сформированности компетенции предлагается использовать оценочные средства промежуточного контроля. Проведенное во второй главе исследование существующих IRT-моделей показало, что они могут использоваться только в случае, если все задания конструктора оцениваются одинаково, по единым «правилам». В связи с этим их применение для оценки компетенций ограничено, т.к. обычно все дисциплины ведут разные преподаватели, а правила оценки заданий промежуточной аттестации у них разные. Данный вывод подтверждается проведенными в работе статистическими исследованиями. Поэтому в диссертации разработана новая модель, названная латентной моделью экзаменационных оценок:

$$P(y_{ij} = k | \theta_i, \delta_j, \tau_{ju}) = \frac{e^{k\theta_i - k\delta_j - \sum_{u=0}^k \tau_{ju}}}{1 + \sum_{v=1}^K e^{\sum_{u=1}^v (\theta_i - \delta_j - \tau_{ju})}}, \quad (4)$$

где θ_i – оценка сформированности компетенции i -го студента; $y_{ij} = k$ – оценка, полученная i -ым студентом по j -ому заданию; δ_j – параметр задания; τ_{ju} – параметр «шага» шкалы оценки по j -му заданию; K – максимальная оценка.

Разности $\delta_j - \tau_{ju}$ можно интерпретировать как требования к уровню сформированности компетенции студента, при достижении которых вероятность получить u -оценку j -го задания становится равна 50%. Данные разности называются «пороговыми» (threshold) – параметрами задания.

Оценка компетенции осуществляется в единице, называющейся логит (log-odds) – единице измерения логарифмических шансов. Пусть θ_1 и θ_2 уровни сформированности компетенции двух студентов. Эти параметры являются латентными, но они проявляются на практике в том, что вероятность выполнить какое-либо задание выше у того студента, у кого выше значение θ . Успех выполнения задания носит вероятностный характер, иными словами успешно ответить на задания мо-

жет как первый, так и второй студент, но при выполнении нескольких заданий эффект более высокого уровня освоения компетенции обязательно проявится.

Пусть $P_{k,k-1}$ – вероятность того, что первый студент получит оценку k , тогда как второй студент получить оценку $k-1$, а $P_{k-1,k}$ – вероятность того, что второй студент получит оценку k , тогда как первый студент получить оценку $k-1$. Тогда, согласно теореме умножения вероятности:

$$\frac{P_{k,k-1}}{P_{k-1,k}} = \frac{P(k | \theta_1, \delta_j, \tau_{uj})P(k-1 | \theta_2, \delta_j, \tau_{uj})}{P(k-1 | \theta_1, \delta_j, \tau_{uj})P(k | \theta_2, \delta_j, \tau_{uj})}, \quad (5)$$

следовательно,

$$\ln \left(\frac{P_{k,k-1}}{P_{k-1,k}} \right) = \ln \left(\frac{e^{k\theta_1 - k\delta_j - \sum_{u=1}^k \tau_{uj}}}{e^{k\theta_2 - k\delta_j - \sum_{u=1}^k \tau_{uj}}} \right) = \theta_1 - \theta_2 \quad (6)$$

Это позволяет представить уровни освоения компетенции точками метрической оси на определенном расстоянии друг от друга. Отношение шансов одного студента получить оценку k к аналогичным шансам другого студента зависит только от уровней их подготовленности, и не зависит от каких-либо других параметров, в т.ч. и от используемых заданий, что является свидетельством объективности полученных оценок компетенций. Шкала логитов обладает следующими свойствами: является интервальной; имеет единицу измерения (один логит); дает возможность перейти от ранжирования студентов к измерению уровней освоения компетенций; легко переводится в любую другую, удобную для восприятия шкалу. Экспериментальное апробирование показало, что предложенные в работе модели и методы позволяют давать обоснованные суждения об уровне сформированности компетенций.

3. Разработана методика оценки сформированности компетенций, статистической проверки адекватности и валидности результатов

Представленная в работе методика позволяет получить объективную информацию как об оценках компетенции, так и об индикаторных заданиях, следствием чего является потенциальная возможность принятия управленческих решений, направленных на повышение качества образовательного процесса. Основные этапы методики:

1. Строятся карты компетенций по направлению подготовки с помощью разработанного в диссертации экспертно-статистического метода.

2. По картам компетенции определяются индикаторные задания, осуществляется выбор модели оценивания, проводится расчет значений параметров модели и статистических критериев (при наличии эмпирических данных).

3. По мере обучения студентов формируются (или дополняются) матрицы первичных баллов, по которым оцениваются компетенции студентов и параметры моделей. Проводится статистическая проверка (верификация) полученных результатов.

4. Анализируются результаты, при необходимости проводится корректировка карт компетенций или принимаются иные управленческие решения: изменение заданий, учебного плана и пр.

Для некоторых этапов методики целесообразно создание экспертной группы в целях снижения субъективности и повышения адекватности анализа в связи с тем, что один человек вряд ли может с уверенностью знать специфику всех читаемых дисциплин. Рассмотрим подробнее некоторые из этапов методики.

На этапе построения карт компетенций экспертная группа определяет для каждой компетенции список дисциплин (из общего списка) и последовательность их изучения. На основе анализа сформированных карт компетенций могут быть проведены корректировки учебного плана и заданий промежуточной аттестации.

После построения карт компетенций определяются индикаторные задания. Выбор индикаторных заданий тесно связан с выбором модели оценивания. В работе показано, что наиболее удобными и в тоже время достаточными является задания промежуточной аттестации, которые однозначно определяются картой компетенции. Тогда для оценки компетенции следует использовать предложенную в работе латентную модель экзаменационных оценок. В противном случае могут быть использованы другие, рассмотренные в работе модели, в частности, 1PL, RSM, РСМ. Для выбора модели можно использовать следующий алгоритм:

1. Рассчитываются значения ряда критериев, связанных с понятием информационной энтропии и расстоянием Кульбака – Лейблера (см. таблица 2). Лучшей считается модель, в достаточной мере полно описывающая данные с наименьшим количеством параметров

Таблица 2

Информационные критерии

Информационный критерий	Формула
Акаике (<i>AIC</i>)	$AIC = 2p - 2\ln(L)$
Скорректированный Акаике (<i>AICc</i>)	$AICc = AIC + \frac{2p(p-1)}{N-p-1}$
Байесовский (<i>BIC</i>)	$BIC = p\ln(N) - 2\ln(L)$
Скорректированный Байесовский (<i>aBIC</i>)	$aBIC = \ln\left(\frac{N-2}{24}\right)p - 2\ln(L)$
Состоятельный Акаике (<i>CAIC</i>)	$CAIC = (1 + \ln(N))p - 2\ln(L)$

где p – количество оцениваемых параметров модели; L – максимизированное значение функции правдоподобия, N – размер выборки.

2. Вычисляется информационная функция, которая определяет количество информации, обеспеченное j -ым заданием (величину, обратно пропорциональную стандартной ошибке измерения θ с помощью задания j):

$$I_j(\theta) = -M \left[\frac{d^2}{d\theta^2} \ln L_j(\theta | y_j) \right], \quad (8)$$

где $L_j(\theta | y_j)$ есть функция правдоподобия j -го задания. Информация всего конструкта, складывается из информационных функций, построенных для каждого задания:

$$I(\theta) = \sum_{j=1}^M I_j(\theta) \quad (9)$$

Значения этой функции являются характеристикой эффективности оценки, чем больше количество информации обеспечивает модель, тем лучше.

3. На основе анализа получившихся результатов определяется модель, которая наилучшим образом описывает эмпирические данные.

Для проведения статистического анализа заданий используются критерии item fit (см. таблицу 1). Если остаются задания с неудовлетворяющими значениями статистик, то их следует проанализировать с целью доработки или удаления из конструкта. Вычисляются «пороговые» параметры заданий, которые определяют уровень требования освоения компетенции, при котором у студента есть 50% вероятность получить данную оценку («сложность» получения оценки). Их анализ позволяет определить «нарушения» в процедурах оценки индикаторных заданий, их излишнюю легкость или сложность.

После прохождения анализа набор индикаторных заданий можно использовать для оценки компетенций обучающихся студентов. Оценки компетенций студентов проходят статистическую проверку (person fit). Формулы и рекомендации по интерпретации всех статистических критериев даются в третьей главе диссертации. При необходимости оценки в логитах могут быть преобразованы в уровневую лингвистическую или рейтинговую шкалы.

4. Разработано математическое и алгоритмическое обеспечения для решения задач построения карт компетенций, оценки компетенций студентов, расчетов их статистических характеристик.

Предложенные в работе математическое и алгоритмическое обеспечение формализуется в виде программного продукта «Система поддержки компетентностно-ориентированной образовательной программы» – СПКООП. Продукт доступен в двух версиях – непосредственно сам продукт отдельно (свидетельство о регистрации программы для ЭВМ № 2015616399) и в виде подсистемы автоматизированной информационной системы «Внеучебная деятельность» – АИС ВД (свидетельство о регистрации программы для ЭВМ № 2014614914).

Основным требованием при разработке программного продукта являлось условие интеграции с существующими информационными системами управления ВУЗа. Это достигнуто за счет разработки соответствующего интерфейса обмена данными, который позволяет напрямую подключаться к большинству используемых баз данных, а в случае невозможности этого обмен данными может быть организован через файлы определенной структуры. Модульная организация позволяет легко дополнять и модернизировать программу. Разработанное программное обеспечение структурно подразделяется на четыре блока (модуля). Первый модуль («экспертная процедура») предназначен для обработки данных экспертов и служит для: 1) построения карт компетенций; 2) построения лингвистической шкалы. Разработанное средство автоматизирует самые трудоёмкие операции: формирует экспертные листы, проводит обработку матриц парных сравнений, строит ранжировки по каждому эксперту и итоговое ранжирование и т.п.

Второй модуль «оценка параметров» реализует оценку параметров предложенных в работе IRT-моделей с помощью метода максимального правдоподобия. Для поиска максимума функции правдоподобия используется численный метод Ньютона.

Третий модуль «статистический анализ» предназначен для статистической обработки данных – проверки применимости выбранной модели оценивания к имеющимся эмпирическим данным, адекватности и валидности результатов.

Четвертый модуль «отчетность» включает в себя возможности обработки и просмотра отчетов непосредственно в приложении; сохранения отчетов в форматах *.docx, *.xlsx, *.pdf; распечатки отчетов непосредственно из интерфейса. Основные формируемые отчеты – карты компетенции, оценки компетенций студентов. Структурная схема программного продукта приведена на рисунке 1.

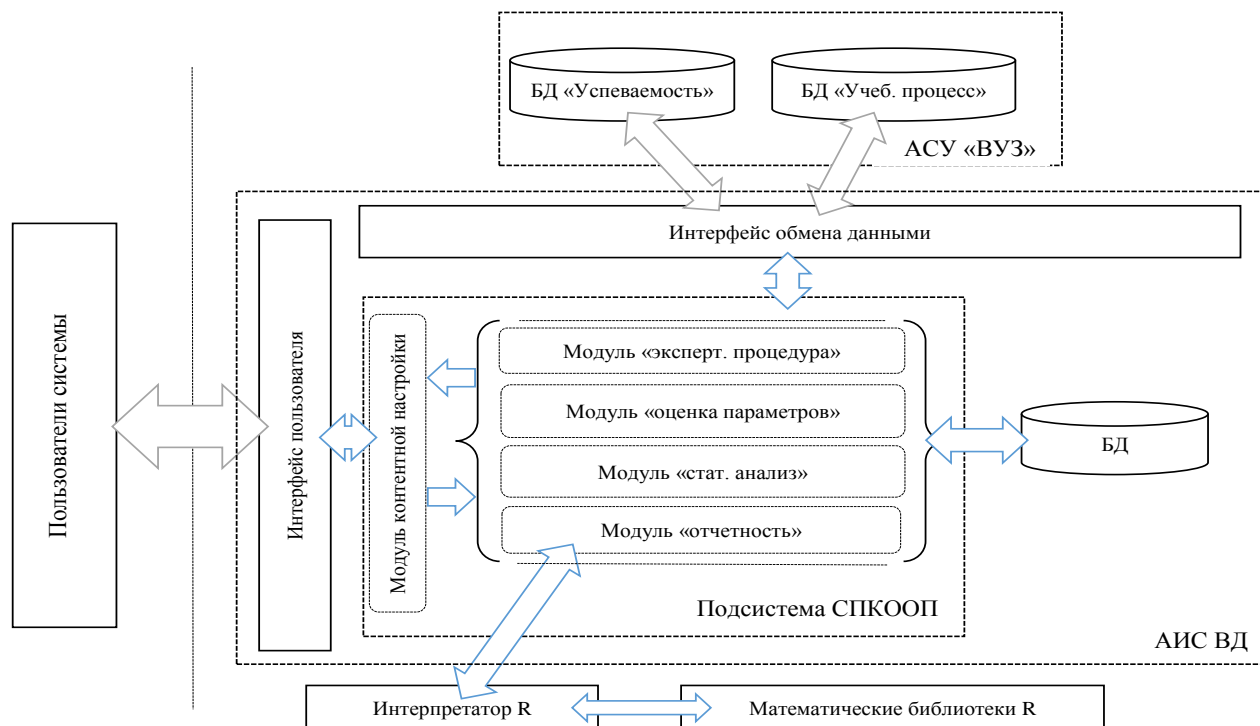


Рисунок 1. Структурная схема программного продукта

Интерфейс программного продукта ориентирован на применение специалистами – методистами, работниками кафедр, поэтому многие технические и математические вопросы скрыты от пользователя, а большинство операции осуществляется с помощью т.н. мастеров. Информационное обеспечение содержит базу данных «типовых» ситуаций и в случае необходимости программный комплекс дает необходимые рекомендации по интерпретации значений параметров и статистик.

Для разработки программного продукта (базовой оболочки) использовалась среда разработки MS Visual Studio, язык VB.NET. Для математических расчетов использовались библиотеки, написанные на языке R. В качестве базы данных для хранения информации используется MS SQL Server. Программный продукт предназначен для эксплуатации на персональном компьютере типа IBM PC. Для работы в диалоговом режиме используется экран дисплея, клавиатура и манипулятор типа «мышь». Программный продукт тестировался на работоспособность на системе Microsoft Windows 7, с установленным NET. Framework v. 4.0, программой – интерпретатором скриптов R v.3.0., и системой работы с отчетами Microsoft Report Viewer 2010.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ РАБОТЫ

1. Предложена модель процесса формирования компетенций в ВУЗе, которая формализуется в виде когнитивной карты – карты компетенций. Разработан новый метод построения карты компетенции на основе системного анализа, когнитивного моделирования и статистического анализа.

2. Предложен конструкт для оценки компетенций, состоящий из индикаторов – оценочных средств промежуточного контроля. Проведенное в диссертации исследование позволяет сделать вывод, что оценочные средства промежуточной аттестации являются достаточными для проведения оценки компетенции.

3. Разработана новая параметрическая латентная модель экзаменационных оценок для оценки компетенции по первичным баллам – результатам выполнения заданий промежуточного контроля.

4. Предложено использование ряда статистик ($Q3$, $OUTFIT\ MEANSQ$, $OUTFIT\ ZSTD$, $INFIT\ MEANSQ$, $INFIT\ ZSTD$) для проверки соответствия латентной модели экзаменационных заданий эмпирическим данным, а также сформулированы рекомендации по их применению.

5. Разработаны несколько видов представления оценки компетенции: в шкале логитов, ранговой рейтинговой шкале и уровневой лингвистической, а также процедуры построения и преобразования шкал.

6. Создан программный продукт для построения карт компетенций, оценки сформированности компетенций, проведения анализа полученных оценок, используемых индикаторных заданий и т.д., на основе разработанного математического и алгоритмического обеспечения.

7. Предложена методика проведения оценки компетенций, включающая в себя процедуры построения карт компетенций, оценки компетенций, оценки параметров заданий, их статистическую проверку и выработку на основе полученных оценок рекомендаций по улучшению учебных программ.

8. Предложенные в работе модели, методы, математическое и алгоритмическое обеспечение прошло апробацию в Байкальском Государственном Университет Экономике и Права. Для апробации было взято направление подготовки 230700 «Прикладная информатика». В процессе апробации успешно решены следующие задачи:

а. Построены модели процесса освоения компетенций в рамках выбранного направления подготовки, определены параметры зачетных, экзаменационных, курсовых заданий.

б. Проведен анализ заданий с использованием статистического анализа на реальных эмпирических данных. Результаты показали хорошую согласованность исходных данных с предложенной в работе латентной моделью экзаменационных заданий. Выявлены несколько заданий с «плохими» характеристиками, выработаны рекомендации по улучшению учебного плана. Определены оценки компетенций студентов.

с. Проведено исследование точности измерения компетенций с использованием предложенных латентной модели экзаменационных оценок и метода максимального правдоподобия. Результаты показывают, что предложенная методика оценки компетенции является эффективной и может быть использована на практике.

д. Проведено исследование влияния многомерности исходных данных. Введение в модель ряда параметров для учета влияния многомерности в ряде случаев позволяет увеличить точность измерения, однако увеличение количества параметров существенно усложняет расчеты и процедуру проведения оценки. Выигрыш в точности с практической точки зрения незначителен.

Таким образом, можно отметить, что все поставленные в исследовании задачи в целом выполнены успешно.

СПИСОК ОСНОВНЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

а) статьи в рецензируемых научных журналах, определенных ВАК Минобрнауки РФ:

1. Родионов А.В. Исследование влияния многомерности данных на оценку компетенций с использованием IRT-моделей / А.В. Родионов // *Фундаментальные исследования*. – 2015. – №10-2.– С. 299-304.

2. Родионов А.В. Модификация рейтинговой параметрической модели оценки латентных факторов для измерения уровня сформированности компетенций / А.В. Родионов // *Известия Иркутской государственной экономической академии*. – 2014. – № 6. – С. 168-174.

3. Родионов А.В. Применение IRT-моделей для анализа результатов обучения в рамках компетентностного подхода / А.В. Родионов, В.В. Братищенко // *Современные проблемы науки и образования*. – 2014. – № 4. – URL: www.science-education.ru/118-13858

б) статьи в иных научных изданиях:

4. Родионов А.В. Решение задачи оценки результатов образовательного процесса с помощью балльно-рейтинговой системы / А.В. Родионов // *Современные информационно-телекоммуникационные технологии в управлении социально-экономическими системами* / под общ. ред. А. П. Суходолова. – Иркутск: Изд-во БГУЭП. – 2013. – С.62-87.

5. Родионов А.В. Использование методов когнитивного моделирования для построения компетентностно-ориентированного образовательного процесса / А.В. Родионов // *Теоретические и прикладные аспекты современной науки*. – 2015. – № 7-1. – С.19-24.

6. Родионов А.В. Использование латентно-структурного анализа при проверке пригодности индикаторных переменных для оценки общекультурных компетенций / А.В. Родионов // *Применение математических методов и информационных технологий в науке, образовании и экономике: сб. науч. тр.* – Иркутск: Издательство БГУЭП. – 2013. – №11 – С.103-109.

в) свидетельства о государственной регистрации программ для ЭВМ:

7. Свидетельство о государственной регистрации программы для ЭВМ № 2014614914. Автоматизированная информационная система «Внеучебная деятельность» / А.В. Родионов, Т.Ю. Новгородцева, В.В. Братищенко // Федеральная служба по интеллектуальной собственности, патентам и товарным знакам. – 2014.

8. Свидетельство о государственной регистрации программы для ЭВМ № 2015616399. СПКООП / А.В. Родионов // Федеральная служба по интеллектуальной собственности, патентам и товарным знакам. – 2015.

Подписано в печать 11.02.2016 г.

Формат бумаги 60 × 84 1/16. Бумага офсетная.

Печать трафаретная. Печ. л. 1. Заказ _____. Тираж 100 экз.

Отпечатано с готового оригинал-макета в ИПО ФГБОУ ВО

«Байкальский государственный университет»

664003, г. Иркутск, ул. Ленина, 11.