

МОДЕЛИ АГРЕГИРОВАНИЯ ОЦЕНОЧНЫХ ПОКАЗАТЕЛЕЙ В СИСТЕМЕ АТТЕСТАЦИИ ПЕРСОНАЛА ПРОМЫШЛЕННЫХ ПРЕДПРИЯТИЙ**Жажа Е.Ю., Суэтина Т.А., Ягудаев Г.Г.**

Эффективное управление персоналом, в том числе подготовка кадров промышленных предприятий в настоящее время невозможна без использования современных компьютерных технологий, которые интенсивно внедряются во все сферы человеческой деятельности. Одной из основных задач системы аттестации персонала является оценка соответствия кадрового состава должностным обязанностям, что требует постоянного и оперативного контроля уровня знаний и умений сотрудников [1]. Повышение эффективности организации системы аттестации достигается на основе формализованного представления процессов сбора, передачи и аналитической обработки результатов, что и составляет задачи системы мониторинга аттестационных показателей. В данной статье предлагаются методы, алгоритмы и программные компоненты сбора и агрегирования аттестационных показателей.

В общем случае механизм проведения аттестации сводится к предъявлению последовательности тестовых заданий различной сложности. При этом в случае использования адаптивных механизмов компьютерного тестового контроля последовательность формируется по следующей схеме:

$$\beta^{(n+1)} = F^{(n)}(\beta^{(1)}, \dots, \beta^{(n)}) + \xi^{(n)}(\beta^{(1)}, \dots, \beta^{(n)}), \quad (1)$$

где $\beta^{(n)}$ - сложность задания на n-ом шаге процедуры; $F^{(n)}$ – некоторое функциональное преобразование результатов ответов; $\xi^{(n)}(\beta^{(1)}, \dots, \beta^{(n)})$ – случайная величина, моделирующая ответ на n-е задание.

Приведенная схема описывает подход к построению множества алгоритмов, различающихся механизмами предъявления заданий и процедурами оценивания. Механизм предъявления дает лишь последовательность заданий различной сложности. Оценивание уровня знаний тестируемого является следующей задачей. В данном случае имеет место два подхода: классификация, когда количество значений оценок тестируемого определено заранее, и оценивание, когда численное значение оценки может быть произвольным.

В рамках IRT-теории каждому заданию приписан уровень сложности и на основе аналитических моделей вводится формализованное описание вероятности правильного выполнения задания с уровнем сложности β испытуемым с уровнем знаний Θ на основе условной вероятности. Анализ показал, что IRT-модель может быть использована совместно с регрессионными моделями и моделями классификации. Использование факторного анализа имеет свои плюсы и минусы в сравнении с IRT-теорией. С одной стороны, необходимо определение множества факторов, что

значительно проще в ИРТ-теории и, значит, дает более адекватную модель в случае гомогенного теста.

В классической теории тестов зарубежные и российские авторы исходили из идеи параллельного измерения интересующего свойства с помощью двух и большего числа тестов, имеющих общее предметное содержание и сходные статистические характеристики [2]. В рамках этой теории приняты следующие постулаты. Эмпирически получаемый результат измерения (X) представляет собой сумму истинного компонента измерения (T), обычно неизвестного, и ошибочного (E), также неизвестного: $X=T+E$. Сложность задания в классической теории определяется эмпирически и соответствует доле неправильных ответов.

Анализ показал, что необходима разработка более точных методов оценки сложности тестов, основанной на модели факторного анализа:

$$\xi_1 = \sum_{j=1}^m \lambda_{1j} F_j + \varepsilon_1, \dots, \xi_p = \sum_{j=1}^m \lambda_{pj} F_j + \varepsilon_p, \quad (2)$$

где λ_{ij} - факторные нагрузки $i, j=1..m, m < p$;

F_1, F_2, \dots, F_p - общие факторы, причем $DF_i=1; cov(F_i, F_j)=0 \quad i \neq j$;

$\varepsilon_1, \dots, \varepsilon_p$ - специфические факторы, причем $cov(\varepsilon_i, \varepsilon_j)=0, i \neq j$;

Модели латентно-структурного анализа достаточно эффективны в эмпирических исследованиях. В статье рассмотрена модель латентного профиля, которая является обобщением латентно-структурного анализа на случай количественных эмпирических переменных. Каждая подгруппа или латентный класс в последней модели должны быть однородными относительно любых латентных величин, которые необходимы для объяснения наблюдаемых взаимосвязей. Полная однородность не требуется, поскольку отклонения от среднего значения в классе случайны, то есть независимы.

Предлагается модель кластеризации уровня подготовки персонала на основе методов латентно-структурного анализа, в которой предполагается, что каждый латентный класс должен быть однородным относительно любых исследуемых величин. Требуется, чтобы каждый латентный класс был достаточно однородным по отношению к любой латентной величине, так чтобы все единичные высказывания внутри класса были статистически независимы. Эта независимость внутри классов выражается следующими уравнениями:

$$p_{ijk} = p_{ik} \cdot p_{lj}, p_{2jk} = p_{2k} \cdot p_{2j}, \dots, p_{qjk} = p_{qk} \cdot p_{qj}, \quad (3)$$

$$p_{ijkl} = p_{ik} \cdot p_{lj} \cdot p_{il}, p_{2jkl} = p_{2k} \cdot p_{2j} \cdot p_{2l}, \dots, p_{qjkl} = p_{qk} \cdot p_{qj} \cdot p_{ql},$$

Преобразование уравнений в соответствии с требованиями однородности групп приводит к системе уравнений:

$$\begin{aligned} n &= n_1 + n_2 + \dots + n_q \\ n_j &= n_1 \cdot p_{1j} + n_2 \cdot p_{2j} + \dots + n_q \cdot p_{qj} \\ n_{jk} &= n_1 \cdot p_{1k} \cdot p_{1j} + n_2 \cdot p_{2k} \cdot p_{2j} + \dots + n_q \cdot p_{qk} \cdot p_{qj} \\ n_{jkl} &= n_1 \cdot p_{2k} \cdot p_{2j} \cdot p_{2l} + n_2 \cdot p_{2k} \cdot p_{2j} \cdot p_{2l} + \dots + n_q \cdot p_{qk} \cdot p_{qj} \cdot p_{ql} \end{aligned} \quad (4)$$

и т.д.

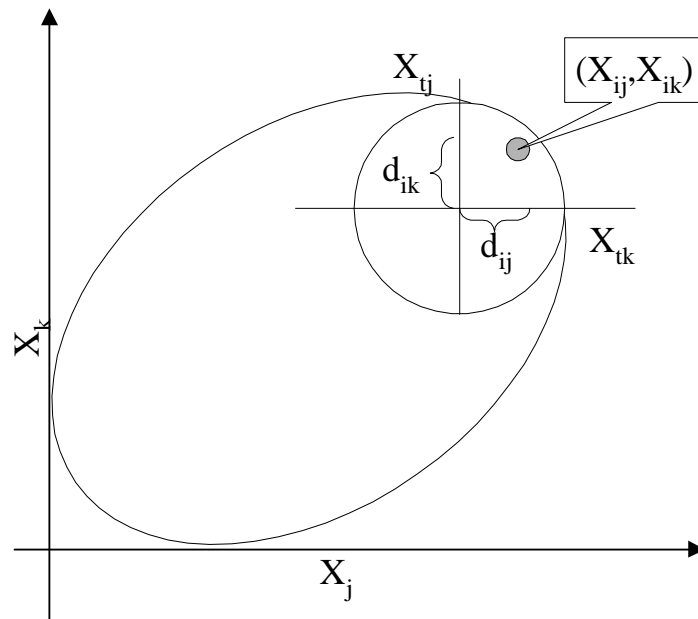


Рисунок 1 - Гипотетическая диаграмма рассеяния

Все наблюдаемые совместные частоты выражаются через $(q+sq)$ латентных параметров, q объемов классов и q латентных вероятностей $(p_{1j}, p_{2j}, \dots, p_{qj})$ для каждого из s признаков теста. Последовательные ступени эмпирических частот насчитывают соответственно $1, s, s(s-1)/2$ и т.д. членов, являющихся коэффициентами бинома $(a+b)^s$. Складывая их, получаем 2^s уравнений, связывающих наблюдаемые и латентные величины в этой модели.

Задача, как и в факторном анализе, заключается в решении основных уравнений относительно неизвестных латентных параметров. Большинство из известных решений не используют совместные частоты с повторяющимися индексами $(n_{jj}, n_{jjk}, n_{jjj}, n_{jjkl}$ и т. д.). В анализе латентной структуры они рассматриваются как аналоги общих факторных дисперсий факторного анализа, которые нам неизвестны. Показано, что представление их в виде эквивалентов, соответствующих смешанным частотам более низкой ступени без повторяющихся индексов (то есть $n_{jj}=n_j, n_{jjk}=n_{jk}$, и т. д.) дает аналог использования равных единиц корреляций в факторном анализе.

Предполагается, что имеется набор s количественных измерений, таких, как баллы тестов в выборке из n человек. По некоторому правилу каждый член этой выборки приписывается одной, и только одной, из q подгрупп. Тогда размер выборки, суммы баллов и суммы произведений баллов для всей выборки выражаются через соответствующие статистики для подгрупп следующим образом:

$$n = n_1 + n_2 + \dots + n_q$$

$$\sum^n X_{ij} = \sum^{n_1} X_{ij} + \sum^{n_2} X_{ij} + \dots + \sum^{n_q} X_{ij} \quad (5)$$

$$\sum^n X_{ij} X_{ik} = \sum^{n_1} X_{ij} X_{ik} + \sum^{n_2} X_{ij} X_{ik} + \dots + \sum^{n_q} X_{ij} X_{ik}$$

$$\sum^n X_{ij} X_{ik} X_{il} = \sum^{n_1} X_{ij} X_{ik} X_{il} + \sum^{n_2} X_{ij} X_{ik} X_{il} + \dots + \sum^{n_q} X_{ij} X_{ik} X_{il} \text{ и т.д.}$$

Все суммирования в (5) проводятся по индивидам. Суммирования слева проводятся по всей выборке, а справа — по членам различных подгрупп; величина X_{ij} есть балл индивида i по тесту j , и она может быть дана в единицах стандартного отклонения или в каких-либо других единицах. То же самое относится и к X_{ik} , X_{il} и т. д.

Таблица 1 - Решение латентного профиля для гипотетического примера трех классов

	Номер теста	Латентный класс		
		I	II	III
Средние по классам	1	-1,50	0,50	0,50
	2	-1,50	0,50	0,50
	3	-1,00	0,00	1,00
	4	-0,50	-0,50	1,50
	5	-0,50	-0,50	1,50
Размеры класса		0,25	0,50	0,25

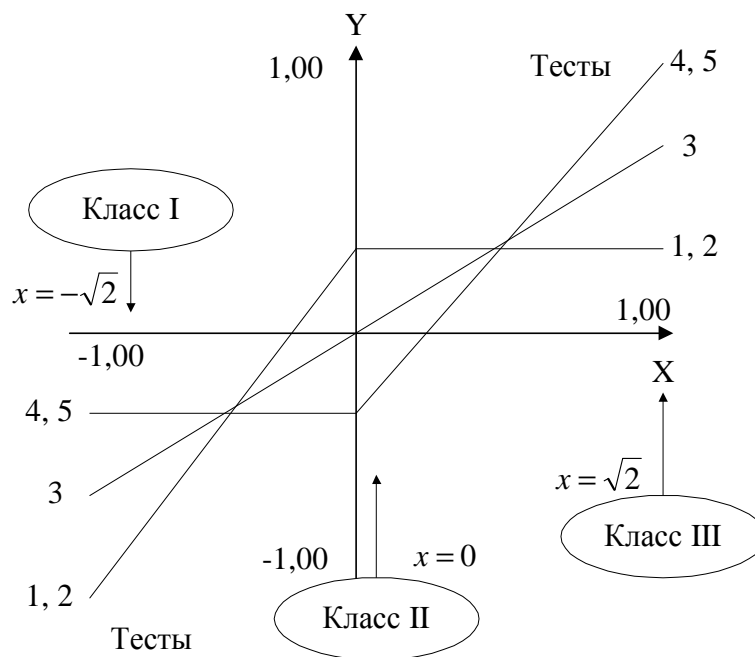


Рисунок 2 - Линия регрессии тестов на латентном континууме для гипотетического случая двух классов

Вид различных регрессий на рисунке 2 объясняются предположением об относительной трудности тестов. Простые тесты (1 и 2) являются различающимися на нижнем конце латентного континуума, а сложные тесты (4 и 5) – на верхнем. Тест (3) средней трудности – на всем континууме.

База данных содержит интегрирующие тесты [3] с привязкой к методическим материалам и результатам выполнения каждого задания каждым сотрудником. Система реализована в виде отдельного программного приложения, включенного в интегрированную адаптивную систему подготовки и аттестации персонала.

Программный комплекс разработан по открытому принципу, что позволяет наращивать его функциональные возможности, добавляя новые механизмы агрегирования аттестационных показателей. Схема подготовки отчетов по квалификационным показателям представлена на рисунке **Ошибка! Источник ссылки не найден.**

При разработке данной технологии подготовки отчетных материалов учтена возможность организации обратной связи. Аналитический отдел реализует обработку и формирует отчет для руководства своего подразделения, после просмотра которого, и внесения соответствующих правок руководством он редактируется и затем выставляется на сервер.

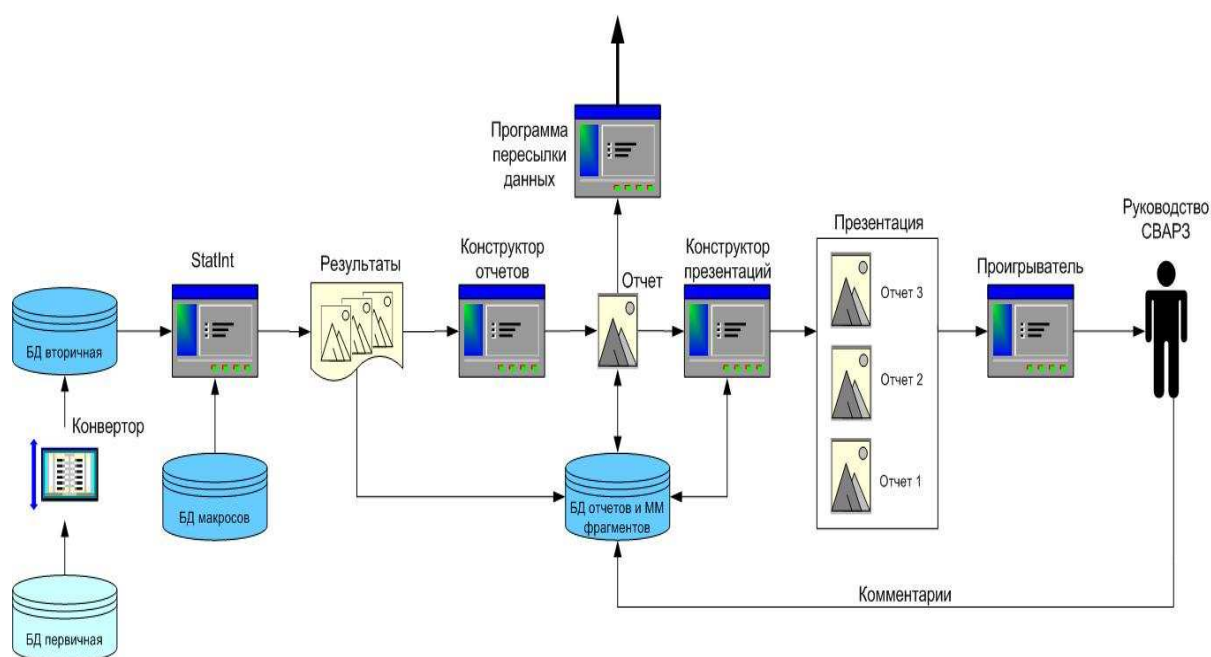


Рисунок 3 - Технология подготовки отчетов по аттестационным показателям

При разработке системы передачи и сбора отчетных материалов были выделены структуры данных, обработка и визуализация которых представляет особый интерес. При этом учитывается, что на предприятиях сложилась определенная технология сбора и хранения данных. Для реализации механизмов обработки данные приводятся к определенному виду, согласованному со структурами данных пакета Statistica. При организации оперативной обработки и агрегирования данных используется система

программных конверторов, которые фильтруют данные и приводят их к выбранному формату.

Список информационных источников

- [1] Николаев А.Б. Использование терм-множеств для оценки связности учебных модулей в системе повышения квалификации административных кадров для предприятий / Николаев А.Б., Строганов В.Ю., Мазуренко С.В., Рогова О.Б. // Моделирование технологических процессов в промышленности и образовании: Сб. науч. тр. М., 2004, МАДИ (ГТУ). - С.19-23.
- [2] Карташев М.И. "Интеграция приложений в системе ""СОТА"" / Ягудаев Г.Г., Карташев М.И. // Методы описания и моделирования бизнес-процессов и технологий в промышленности, строительстве и образовании: сб. науч. тр. МАДИ № 3/47. – М.: МАДИ, 2010. – С. 97-104.
- [3] Свободин В.Ю. Анализ базовых моделей связности учебного материала / Строганов В.Ю., Карташов М.И., Ульянова А.И., Свободин В.Ю. // Оптимизация решений в промышленности, строительстве и образовании: сб. науч. тр. МАДИ № 1/45. – М.: МАДИ, 2010. – С. 49-52.