

УДК 005.935.3+025.2

ББК 65.050.2-73

АГРЕГИРОВАНИЕ РЕЗУЛЬТАТОВ ЭКСПЕРТНОЙ СТРАТИФИКАЦИИ БЕЗ ИСПОЛЬЗОВАНИЯ МАТРИЦ ПАРНОГО СРАВНЕНИЯ НА ПРИМЕРЕ РЕЙТИНГА НАУЧНЫХ ЖУРНАЛОВ¹

Федорец О. В.²

*(Учреждение Российской академии наук
Всероссийский институт научной
и технической информации РАН, Москва)*

Отличительной особенностью научного журнала как объекта оценивания является его политематический характер, что требует получения оценок от экспертов, являющихся специалистами в различных научных областях. Другая особенность – каждый эксперт оценивает своё множество журналов, поэтому матрица экспертных оценок оказывается неполной. Предлагается методика построения абсолютной шкалы на множестве экспертных оценок и использование скользящих сумм для лексикографического упорядочения журналов.

Ключевые слова: организация экспертизы, рейтинг научных журналов, агрегирование экспертных оценок, неполные экспертные данные, построение количественной шкалы, скользящая сумма.

¹ Настоящая статья является ответом на критические замечания в адрес статьи [8] того же автора, опубликованной в настоящем выпуске Сборника. Как составная часть дискуссии о рейтингах научных журналов настоящая статья не прошла научного рецензирования и публикуется в авторской редакции. Редакционная коллегия Сборника может не разделять взглядов и идей автора (прим. ред.).

² Федорец Олег Владимирович, старший научный сотрудник (ovf@viniti.ru).

1. Введение

Нельзя сказать, что отказ от матриц парного сравнения, вынесенный в заглавие статьи, продиктован каким-то предубеждением против этого инструмента агрегирования экспертных оценок. Просто в ситуации, когда количество оцениваемых объектов измеряется тысячами, количество экспертов сотнями и лишь менее 2% элементов матриц парного сравнения определены, необходимо было искать другой подход.

Новый объект оценивания (научный журнал), большая размерность задачи, неполные матрицы экспертных оценок, отказ от использования матриц парного сравнения, предложенный эвристический алгоритм, не лишённый недостатков – всё это, описанное в предыдущей статье [8], вызвало оживлённую дискуссию. Отзывов поступило достаточно много (восемь), поэтому невозможно ответить на все замечания и предложения. Однако по некоторым ключевым вопросам рецензенты высказали сходное мнение, что позволяет ответить сразу многим.

Статья состоит из введения и двух разделов, сформированных следующим образом:

В первом разделе, построенном в виде диалога, дан ответ на отдельные критические замечания А. И. Орлова, так как его рецензия [7] существенно отличается от остальных. В этом же разделе идёт речь о терминах «экспертная классификация» и «кластеризованная ранжировка». В итоге делается выбор в пользу третьего термина – «экспертная стратификация».

Материал второго раздела группируется не по персонам рецензентов, а по выявленным ими проблемам. Это объясняется похожими замечаниями рецензентов по некоторым ключевым вопросам: критерий «максимальная оценка», квантификация (оцифровка) качественных оценок с помощью вербально-числовой шкалы, суммирование оцифрованных экспертных оценок. Поскольку с момента написания обсуждаемой статьи прошло более полугода, будут предложены два направления совершенствования методики проведения экспертизы и агреги-

рования оценок – организационное и математическое. Организационное направление касается расширения состава экспертной группы и разделения процедуры агрегирования оценок на две стадии. Математическое направление – это построение обоснованной количественной шкалы на множестве качественных экспертных оценок и метод скользящей суммы, основанный на критериально-частотном подходе к представлению результатов экспертной стратификации.

2. Ответы на некоторые критические замечания А. И. Орлова [7]

2.1. ЗАМЕЧАНИЕ 1

«Нелепость этого «ТОР-20» бросается в глаза. Например, в нем указаны издания второстепенных вузов (например, «Известия Томского политехнического университета»), но отсутствует «Вестник МГУ им. М. В. Ломоносова» – общепризнанный лидер среди вузов страны. Как могла появиться подобная нелепость, дискредитирующая применение экспертных методов?»

Нелепым было бы включение «Вестника МГУ» в список, поскольку такого журнала не существует. МГУ им. М. В. Ломоносова издаёт множество изданий, каждое со своей тематикой, редколлегией, подписным индексом, кодом ISSN и прочими атрибутами самостоятельного журнала. Ни одна информационная служба в мире не трактует все тематические серии «Вестника МГУ» как единое издание, и ни одна библиотека не может выдать из фондов экземпляр этого несуществующего журнала – в заказе придётся уточнить название серии.

Что касается якобы «второстепенности» Томского политехнического университета, на его сайте <http://www.tpu.ru> сообщается, что в 1997 г. он включён в «Свод особо ценных объектов культурного наследия народов России». Из других источников в Интернете можно узнать, что в этом Своде сегодня всего 14 ВУЗов, три из которых – сугубо гуманитарные.

Впрочем, мы строим рейтинг не университетов и не издательств, а научных журналов.

2.2. ЗАМЕЧАНИЕ 2

«Чье же мнение изучалось? Наверно, потребителей информации – читателей изданий ВИНИТИ? Отнюдь нет. Редакторов реферативных журналов. Являются ли они научными работниками? Насколько знакомы с мнениями потребителей? В статье О. В. Федорца [8] ответов на эти вопросы нет. Создается впечатление, что опрашивали не тех, от кого можно было бы получить адекватную информацию. О значимости научных изданий надо спрашивать научных работников – ведущих ученых, сотрудников НИИ, преподавателей вузов, хотя бы тех, кто работает с ВИНИТИ – внештатных референтов».

К сожалению, от автора не зависело определение состава экспертов. В данном случае необходимо чётко отделять научную проблематику статьи от организационных и финансовых возможностей, которые обычно ограничены. Разумеется, контингент экспертов напрямую влияет на результат любой экспертизы. Прежде чем привлекать к экспертизе ведущих учёных и тратить их драгоценное время, необходимо разработать и протестировать методику. А для этой цели вполне подходят и эксперты-редакторы, у которых имеется как минимум высшее образование, близкое к тематике реферативного журнала (РЖ), и которые находятся в постоянном взаимодействии с внештатными референтами. Редакторы также владеют информацией о количестве и составе подписчиков своего реферативного журнала.

2.3. ЗАМЕЧАНИЕ 3

«Свой вклад в нелепость выводов могла внести и процедура обработки данных. Вполне естественно, что свои оценки эксперты дают в порядковой шкале. Но описанная в статье О. В. Федорца процедура обработки данных полностью противоречит требованиям теории измерений. ...Выводы должны быть инвариантны относительно допустимых преобразований в порядковой шкале. Следовательно, оцифровка согласно таблицам 2 и 3 недопустима, складывать оценки экспертов нельзя, и т.д. ...»

На это утверждение можно ответить выдержкой из рецензии Б. Г. Миркина:

«Я думаю, что квантизация вербальных шкал в манере, описанной в статье – разумное дело, потому что концепция порядковой шкалы неадекватно отражает деление ее континуума человеком, тем более, экспертом. Следует, однако, позаботиться об устойчивости результатов относительно изменения интервалов между значениями шкалы, что, на мой взгляд, нетрудно, ибо достигается чисто вычислительными средствами» [5].

Одним из таких чисто вычислительных средств являются коэффициенты ранговой корреляции. Исследование ранговой корреляции позволяет определить статистическую значимость различия ранжировок, но не позволяет судить о том, в лучшую или в худшую сторону изменился ранжированный список при изменении шкалы или методики агрегирования экспертных оценок. Для ответа на этот вопрос необходимо сравнивать ранжировки с некими эталонами – эталонным множеством журналов или эталонной ранжировкой журналов.

В эталонное множество должны входить журналы, научная ценность которых не подлежит сомнению. Далее можно сравнивать верхние части (*TOP-x%*) ранжированных списков по количеству вхождений в них эталонных журналов. Чтобы сформировать эталонное множество российских журналов, достаточно взять из списка ВАК те издания, которые рекомендованы для публикации соискателям учёной степени доктора наук. Для включения в эталонный список зарубежных журналов можно отобрать издания, которые отражаются в *Journal Citation Report* и обладают импакт-фактором выше среднего среди журналов, входящих в соответствующие тематические области.

Вопрос заключается в том, какие тематические списки являются «более эталонными» – полученные в результате обработки экспертной информации или ранжированные по критерию цитирования. Этот вопрос должен решаться в зависимости от уровня компетентности экспертной группы.

2.4. ЗАМЕЧАНИЕ 4

«Неадекватные методы анализа данных сочетаются с неадекватной терминологией. «Экспертную классификацию» обычно называют кластеризованной ранжировкой, а термин «классификация» связывают с измерениями по шкале наименований, ибо термин «классификация» соответствует разбиению на неупорядоченные классы».

Искусственное приведение результата экспертной классификации по упорядоченным классам к кластеризованной ранжировке может повлечь искажение экспертных оценок и потерю важной информации, поскольку между кластеризованной ранжировкой и результатом экспертной классификации по упорядоченным классам нет взаимоднозначного соответствия.

Обратимся первоисточнику: «Каждую кластеризованную ранжировку, как и любое бинарное отношение, можно задать матрицей $\|x(a, b)\|$ из 0 и 1 порядка $k \times k$. При этом $x(a, b) = 1$ тогда и только тогда, когда $a < b$ либо $a = b$. В первом случае $x(b, a) = 0$, а во втором $x(b, a) = 1$. При этом хотя бы одно из чисел $x(a, b)$ и $x(b, a)$ равно 1» [6, стр. 399].

Теперь предположим, что два участника экспертной группы сошлись во мнении, что три объекта ранжируются так: $[a < b < c]$. Однако они выдали различающиеся экспертные оценки: номера упорядоченных классов $\{5, 4, 1\}$ и $\{5, 2, 1\}$ соответственно. Если представить оценки двух экспертов в виде матриц $\|x(a, b)\|$ из 0 и 1 как описано выше, то матрицы будут одинаковые. Если же построить матрицы парных сравнений с кососимметрической калибровкой (где элемент равен разнице оценок), то они будут различаться.

Переход от матрицы парных сравнений, заданной в некоторой калибровке, к откалиброванной по-иному матрице возможен не всегда, но лишь при соблюдении некоторых дополнительных содержательных условий и нередко сопряжён с потерей важной информации [2].

Термин «экспертная классификация» был заимствован из учебного пособия Б. Г. Литвака, где в разделе «Методы получе-

ния количественных и качественных экспертных оценок» читаем следующее: «Экспертная классификация. Этот метод целесообразно использовать, когда необходимо определить принадлежность оцениваемых альтернативных вариантов к установленным и принятым к использованию классам, категориям, уровням, сортам (далее – классам). ... В зависимости от целей экспертизы может возникнуть необходимость отнесения альтернативных вариантов к упорядоченным классам. Скажем, необходимо отнести оцениваемые объекты к соответствующим категориям, причём так, чтобы более предпочтительные были отнесены к более предпочтительным категориям. Естественно, это отражается на процедуре экспертной классификации. Но главное, чтобы эксперт однозначно понимал поставленную перед ним задачу» [4].

Поиск более краткого и одновременно ясного названия для понятия «экспертная классификация по упорядоченным классам» привёл к термину «стратификация», который используется в социологии, геологии, медицине. Словосочетание «экспертная стратификация» обнаружено в Интернете на персональном сайте Павла Горского, где содержится подходящее определение: «Здесь важно отметить, что классы НЕУПОРЯДОЧЕНЫ друг относительно друга, т. е. нельзя сказать, что какой-то класс важнее (лучше, старше, дороже и п.т.) другого. Например, людей можно классифицировать по полу или национальности. Правильная постановка диагноза – также пример *классификации*. Второй способ структурирования называется СТРАТИФИКАЦИЯ. Это название произошло от английского термина «**страта**», (*stratum*) что означает «слой, пласт». Иными словами, *стратификация* есть разбиение множества на ряд уровней или слоев. В отличие от классов, страты упорядочены» [3].

Далее Горский пишет: «Связь страт с неким абстрактным «качеством» крайне важна для понимания идеи стратификации. «Пятизвездный» отель не просто лучше «двухзвездного», а можно говорить **на сколько** он лучше» [3].

Итак, подходящий термин найден и определён – это «экспертная стратификация».

3. Основные проблемы и пути их решения

3.1. МАКСИМАЛЬНАЯ ОЦЕНКА И НЕОБХОДИМОСТЬ ОРГАНИЗАЦИОННОГО УЛУЧШЕНИЯ

Организационное мероприятие, позволяющее улучшить качество экспертизы – это увеличение количества экспертов по каждой тематике.

В идеале экспертиза должна проводиться в два этапа. На первом этапе собираются оценки по отдельным тематикам и вычисляются медианы, на втором этапе агрегируются медианы тематических оценок. При такой организации экспертизы агрегируются не индивидуальные экспертные оценки, а усреднённые тематические оценки. В таком случае критерий МАКСИМАЛЬНАЯ_ОЦЕНКА, справедливо ставший объектом критики рецензентов, фактически превратился бы из наивысшей оценки объекта одним экспертом в наивысшую из тематических медиан, что сделало бы его более устойчивым к ошибкам и субъективизму.

3.2. ПОСТРОЕНИЕ АБСОЛЮТНОЙ ШКАЛЫ НА МНОЖЕСТВЕ ЭКСПЕРТНЫХ ОЦЕНОК

Нельзя не согласиться с замечаниями рецензентов, что произвольная оцифровка порядковых экспертных оценок недопустима. Взяв за основу вербально-числовую шкалу Харрингтона, необходимо было убедиться в соответствии её численных значений полученному множеству экспертных оценок. Но желательно было разработать методику построения собственной шкалы на множестве экспертных оценок. Этого не было сделано, что вызвало справедливую критику. Чтобы закрыть этот пробел, далее приведена математическая модель, позволяющая вычислить коэффициенты важности (весовые коэффициенты) упорядоченных классов на множестве экспертных оценок. Отличие новых коэффициентов от ранее использовавшихся оказа-

лось невелико, но теперь вербально-числовая шкала получила вероятностное обоснование.

Пусть результаты отнесения объектов к упорядоченным классам представлены в виде n экспертных списков. Количество элементов в списках различно. Элемент списка – это пара значений (a, x) , где a – уникальный идентификатор объекта (код, наименование, обозначение и т.п.), x – порядковый номер класса (категории, уровня, сорта и т.п.). Помимо порядкового номера, каждому классу соответствует вербальное описание (например «высокий», «средний», «низкий» и т.п.), которое должно быть одинаково понятным всем экспертам. Требуется определить числовые значения, соответствующие упорядоченным классам. В результате будет построена абсолютная шкала на множестве экспертных оценок, соответствующая порядковым экспертным оценкам.

Рассмотрим случайное событие «объект с оценкой x предпочитается при сравнении со случайно выбранным объектом из i -го экспертного списка». Точечную оценку вероятности этого события можно вычислить следующим образом:

$$(1) \quad P_{x,i} = \frac{m_{x,i} + k_{x,i}/2}{s_i},$$

где $m_{x,i}$ – количество объектов в i -м списке с оценкой, менее предпочтительной, чем x ; $k_{x,i}$ – количество объектов в i -м списке с оценкой, равной x ; s_i – общее количество объектов в i -м списке.

Следует пояснить смысл выражения $k_{x,i}/2$ в формуле (1). В среднем на большом количестве сравнений случайно выбранный объект с оценкой x предпочтительнее половины объектов и, соответственно, менее предпочтителен, чем половина объектов с такой же оценкой x .

В данной модели считается, что у всех объектов есть неповторяющиеся «точные» оценки, огрубляемые номерами классов. Учитывается и то, что случайно выбранный объект сравнивается не только с другими объектами того же класса, но и с самим собой. Тогда при нестрогой трактовке предпочтения («более

предпочтителен или равноценен») среднее количество менее предпочтительных объектов равно

$$g_1 = \frac{1}{k_{x,i}} \sum_{i=1}^{k_{x,i}} i.$$

При строгой трактовке предпочтения («более предпочтителен») среднее количество менее предпочтительных объектов равно

$$g_2 = \frac{1}{k_{x,i}} \sum_{i=0}^{k_{x,i}-1} i.$$

Используя формулу суммы первых n членов арифметической прогрессии, нетрудно доказать, что величина $k_{x,i}/2$ является средним арифметическим величин g_1 и g_2 .

Теперь убедимся, что сумма оценок вероятностей противоположных событий равна единице:

$$P_{x,i} + \overline{P_{x,i}} = \frac{m_{x,i} + k_{x,i}/2}{s_i} + \frac{s_i - m_{x,i} - k_{x,i} + k_{x,i}/2}{s_i} = 1.$$

Размеры экспертных списков могут существенно различаться, поэтому в качестве числовой интерпретации качественной оценки x предлагается использовать взвешенное среднее арифметическое вероятностей $P_{x,i}$. Принимая в качестве весовых коэффициентов размеры экспертных списков и затем подставляя вместо $P_{x,i}$ формулу (1), получаем следующее выражение:

$$(2) \quad P_x = \frac{\sum_{i=1}^n s_i \cdot P_{x,i}}{\sum_{i=1}^n s_i} = \frac{\sum_{i=1}^n (m_{x,i} + k_{x,i}/2) s_i}{\sum_{i=1}^n s_i},$$

где n – количество экспертных списков, в которых есть хотя бы один объект с оценкой x .

Продемонстрируем построение числовой шкалы на примере оценивания восьми объектов, которые получили порядковые оценки 6-ти экспертов по пятибалльной шкале, обозначенных как $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_6$ (см. таблицу 1). Объектам для наглядности

присвоены имена – латинские буквы от *A* до *H*. Данные для примера взяты без изменения из предыдущей статьи [8].

Таблица 1. Индивидуальные экспертные оценки

Объект	Э ₁	Э ₂	Э ₃	Э ₄	Э ₅	Э ₆
<i>A</i>	5		4		3	1
<i>B</i>	5	2		4		3
<i>C</i>		3	5		5	
<i>D</i>		4	3	4		
<i>E</i>	4	4			3	
<i>F</i>	3		4	3		5
<i>G</i>	1			3	2	
<i>H</i>		5			4	

Для примера в таблице 1 применение (2) позволяет вычислить следующие ненормированные коэффициенты важности для градаций порядковой шкалы: $P_5 = 0,86$; $P_4 = 0,61$; $P_3 = 0,31$; $P_2 = 0,10$; $P_1 = 0,13$.

Необычное неравенство $P_2 < P_1$, объясняется очень малым количеством экспертов и оцениваемых объектов в примере, представленном в таблице 1. На выборках значительного объема подобное событие представляется маловероятным.

Рассмотрим результаты экспертной стратификации 11067 научных журналов. Генеральная совокупность экспертных оценок состояла из 43336 ненулевых оценок, хранящихся в 240 экспертных списках. Средний размер экспертного списка – 180 журналов, каждый журнал получил в среднем 4 оценки. Нетрудно подсчитать, что в случае построения матриц парного сравнения журналов в среднем лишь менее 2% элементов матриц были бы определены.

В результате применения (2) получены следующие числовые значения (ненормированные коэффициенты важности) порядковых балльных оценок для построения вербально-числовой шкалы, которые можно увидеть в табл. 2 в колонке P_x .

Для сравнения приводится шкала Харрингтона и числовые значения, использовавшиеся в предыдущей версии шкалы [8].

Из таблицы 2 видно, что все полученные значения P_x укладываются в интервалы шкалы Харрингтона, а отклонения P_x от Q_x незначительны, при этом относительная погрешность уменьшается по мере продвижения от «низкой» оценки к «высокой».

Для разработанного алгоритма не обязательно нормирование коэффициентов важности таким образом, чтобы их сумма была равна единице, так как упорядочение объектов при этом не изменится.

Таблица 2. Статистическая вербально-числовая шкала

Название уровня	x	Кол-во объектов	Новое значение, P_x	Шкала Харрингтона	Пред. Значение, Q_x
Высокий	5	8618	0,894	1,0 – 0,8	0,900
Выше среднего	4	6674	0,731	0,8 – 0,63	0,715
Средний	3	8739	0,550	0,63 – 0,37	0,500
Ниже среднего	2	8585	0,351	0,37 – 0,2	0,285
Низкий	1	10720	0,133	0,2 – 0,0	0,100

3.3. МЕТОД СТРОЧНЫХ СУММ И СУММИРОВАНИЕ ОЦЕНОК

Одним из наиболее известных алгоритмов упорядочения объектов по результатам парных сравнений является метод строчных сумм. В частности, рекомендация ознакомиться с этим методом содержится в рецензии Л. Ч. Абаева [1].

Допустим, матрицы парных сравнений служат для представления данных более простого типа – числовых оценок. Матрицы парных сравнений определены так, что каждый элемент представляет собой разницу оценок объектов. В этом случае метод строчных сумм приводит к упорядочению объектов по сумме полученных ими оценок [9].

К сумме экспертных оценок можно прийти и другим путём. Далее предлагается критериально-частотный подход к представлению результатов экспертной стратификации, который также приводит к суммированию числовых оценок.

3.4. КРИТЕРИАЛЬНО-ЧАСТОТНЫЙ ПОДХОД И ВЗВЕШЕННАЯ СУММА ЧАСТОТ ОЦЕНОК

Пусть m – количество упорядоченных классов или качественных уровней порядковой шкалы. Каждому объекту сопоставлен вектор частотного распределения оценок $K(z) = (k_1(z), \dots, k_i(z), \dots, k_m(z))$, где $i = 1, 2, \dots, m$, $k_i(z)$ – количество экспертов, отнёсших объект z к i -му качественному уровню (т. е. частота оценки).

В нашем случае эксперты – специалисты в различных научных областях – независимо друг от друга оценивают различные тематические аспекты политематических журналов, поэтому замена абсолютной частоты на относительную некорректна. Действительно, пусть объекты A и B оценены по пятибалльной порядковой шкале, получив при этом векторы частот $(0, 0, 0, 0, 1)$ и $(0, 0, 1, 0, 1)$. После замены на относительные частоты получим $(0, 0, 0, 0, 1)$ и $(0, 0, 1/2, 0, 1/2)$, т. е. частота оценки объекта B занижается в 2 раза. Например, журнал B публикует статьи в смежных областях – биологии и химии. Биолог присвоил оценку 5 («высокая»), а химик оценку 3 («средняя»). Означает ли это, что мы должны при этом искусственно занижать частоту оценок у журнала B ? Очевидно, что нет.

Будем считать, что $k_i(z)$ – это значение i -го критерия оценивания объекта. Исходя из гипотезы о независимости критериев по предпочтению, зададим аддитивную функцию ценности $F(z)$ в виде линейной свёртки критериев:

$$(3) \quad F(z) = \sum_{i=1}^m w_i k_i(z),$$

где w_i – весовой коэффициент i -го критерия, $k_i(z)$ – значение критерия, равное количеству экспертных оценок i -го уровня, присвоенных объекту z .

В случае целочисленных значений критериев линейная свёртка равносильна суммированию весовых коэффициентов. Таким образом, мы вновь приходим к суммированию «оцифрованных» экспертных оценок, что и было сделано в [8] при вычислении одного из критериев, применявшихся для лексикографического упорядочения журналов.

3.5. МЕТОД СКОЛЬЗЯЩЕЙ СУММЫ ЧИСЛОВЫХ ОЦЕНОК

Нужно ли было вычислять взвешенную сумму **всех** оценок объекта? Здесь следует согласиться с рецензентами, что не нужно. Более правильным представляется отдельно учитывать высокие и низкие оценки, например, суммировать оценки по группам с последующим лексикографическим упорядочением по суммам. Например, сначала вычислить взвешенную сумму оценок, близких к «высокой», затем к «выше средней» и т.д., так как сравнимы между собой только близкие уровни порядковой шкалы, а «высокий» уровень может настолько качественно отличаться от «низкого», что это отличие невозможно измерить количественно. Приведём простой пример. Вряд ли можно измерить знания учащихся таким образом, чтобы определить, сколько «двоечников» могут компенсировать объём знаний одного отличника (например, совместными усилиями написать контрольную работу или пройти тестирование на «отлично»). Зато возможно сравнить «отличника» с «хорошистом» или «троечника» с «двоечником».

Следовательно, перед суммированием необходимо определить отрезки порядковой шкалы, в пределах которых оценки допустимо суммировать. Если применить скользящий отрезок шкалы, как это делается в методе скользящего среднего при сглаживании временных рядов, то предлагаемый метод агрегирования экспертных оценок можно назвать методом скользящей суммы. Для пятибалльной порядковой шкалы и длины отрезка, равной трём, имеем три обобщённых критерия на основе модификации формулы (3):

$$F_5(z) = \sum_{i \in \{5,4,3\}} w_i k_i(z), \quad F_4(z) = \sum_{i \in \{4,3,2\}} w_i k_i(z), \quad F_3(z) = \sum_{i \in \{3,2,1\}} w_i k_i(z).$$

Далее в таблице 3 приведён результат вычисления скользящих сумм для примера из таблицы 1. В случае уменьшения длины скользящего отрезка до двух точек мы получим четыре обобщённых критерия для лексикографического упорядочения объектов.

3.6. ПРИМЕР

Заменяем номера классов из таблицы 1 числовыми значениями, полученными ранее по формуле (2): $P_5 = 0,86$; $P_4 = 0,61$; $P_3 = 0,31$; $P_2 = 0,10$; $P_1 = 0,13$. Результат замены – в таблице 3, там же – значения трёх скользящих сумм, по которым производится упорядочение объектов.

Лексикографическое упорядочение (по убыванию) по трём колонкам F_x приводит к следующей кластеризованной ранжировке: $F \succ C \succ A \succ B \succ \{D, E\} \succ H \succ G$.

Таблица 3. Агрегирование оценок методом скользящей суммы по трём точкам

x	Θ_1	Θ_2	Θ_3	Θ_4	Θ_5	Θ_6	F_5	F_4	F_3
A	0,86		0,61		0,31	0,13	1,78	1,02	0,44
B	0,86	0,10		0,61		0,31	1,78	1,02	0,41
C		0,31	0,86		0,86		2,03	0,31	0,31
D		0,61	0,31	0,61			1,53	1,53	0,31
E	0,61	0,61			0,31		1,53	1,53	0,31
F	0,31		0,61	0,31		0,86	2,09	1,23	0,62
G	0,13			0,31	0,1		0,31	0,41	0,54
H		0,86			0,61		1,47	0,61	0

Остался кластер $\{D, E\}$, в котором объекты не упорядочены. Для дополнительного упорядочения внутри кластеров можно применять различные подходы в зависимости от характера задачи и оцениваемых объектов. Переход от оценок к рангам не всегда подходит, так как является огрублением результатов экспертной стратификации. Предлагается два способа вычисления дополнительного критерия для упорядочения внутри кластеров:

1. Вычислить сумму оценок объектов, которые менее предпочтительны, чем данный объект, и упорядочить объекты внутри кластеров по убыванию этой суммы.

2. Вычислить сумму оценок объектов, которые более предпочтительны, чем данный объект, и упорядочить объекты внутри кластеров по возрастанию этой суммы.

Первый способ задаёт следующее отношения предпочтения внутри кластера: $D \succ E$, так как $1,03 > 0,95$.

Второй способ приводит к аналогичному результату $D \succ E$, так как $2,33 < 4,05$.

Окончательный ответ: $F \succ C \succ A \succ B \succ D \succ E \succ H \succ G$.

В предыдущем варианте алгоритма [8] данные этого же примера привели к другому отношению предпочтения: $F \succ B \succ C \succ A \succ H \succ D \succ E \succ G$.

Главное отличие нового результата ранжирования – существенное изменение позиций объектов B и H . Позиции остальных объектов изменились не более чем на один ранг.

Для научных журналов второй способ вычисления дополнительного критерия для упорядочения внутри кластеров представляется более логичным, так как факт вхождения журнала «в число первых» более важен, чем его превосходство над значительным количеством других журналов. Здесь можно провести аналогию с различными видами спорта. Есть виды спорта, в которых важно количество побеждённых соперников (турниры), а есть виды, в которых важно прийти к финишу в числе первых (гонки).

3.7. РАНЖИРОВАНИЕ ЖУРНАЛОВ МЕТОДОМ СКОЛЬЗЯЩЕЙ СУММЫ

Вернёмся к агрегированию результатов экспертной стратификации научных журналов и их ранжированию. Фрагмент списка (*TOP-20*), полученного методом скользящих сумм, представлен в таблице 4. Первые двадцать журналов, полученные в этот раз, почти полностью совпадают с первой двадцаткой в предыдущей статье. Учитывая большие размеры ранжированных списков (11067 журналов), различие по первым двадцати элементам списков можно считать незначительным: в новый список *TOP-20* вошли журналы *Technica (Suisse)* и *Journal of*

Crystal Growth, вышли журналы *Thin Solid Films* и *Известия Томского политехнического университета*. В старом списке журналы *Technica (Suisse)* и *Journal of Crystal Growth* занимали соответственно 21 и 24 места, в то время как покинувшие *TOP-20* журналы *Thin Solid Films* и *Известия Томского политехнического университета* заняли в новом списке соответственно 21 и 22 места.

В таблице 5 представлен фрагмент списка *TOP-20*, полученный методом скользящих сумм при длине скользящего отрезка, равной двум точкам данных. Лексикографическое упорядочение производилось по четырём скользящим суммам. Изменения стали более заметными. По сравнению с таблицей 4 в *TOP-20* сменились 6 журналов.

Таблица 4. *TOP-20 ранжированного списка научных журналов, полученного методом скользящей суммы по трём точкам*

№	Страна	Код ISSN	Название
1	США	1046-6770	Proceedings of SPIE
2	Россия	0869-5652	Доклады Российской академии наук
3	Германия	0341-5775	Maschinenmarkt (MM)
4	Великобритания	0028-0836	Nature: International Weekly Journal of Science (Gr. Brit.)
5	Германия	0042-1758	VDI-Nachrichten
6	США	0036-8075	Science
7	США	0027-8424	Proceedings of the National Academy of Sciences of the United States of America
8	Германия	0344-6166	Produktion: Die Wirtschaftszeitung für die deutsche Industrie
9	Украина	1025-6415	Доповіді Національної академії наук України
10	США	0007-7135	Business Week. European Edition
11	США	0022-2461	Journal of Materials Science
12	США	0003-6951	Applied Physics Letters
13	Россия	0236-1493	Горный информационно-аналитический бюллетень

14	Швейцария	0040-0866	Technica (Suisse)
15	Великобритания	0013-5194	Electronics Letters
16	Россия	0321-2653	Известия высших учебных заведений (вузов). Северо-Кавказский регион. Технические науки
17	Россия	0367-6765	Известия Российской академии наук (РАН). Сер. физическая
18	Швейцария	0921-5093	Materials Science and Engineering. A. Structural Materials: Properties, Microstructure and Processing
19	США	0021-8979	Journal of Applied Physics
20	Нидерланды	0022-0248	Journal of Crystal Growth

Таблица 5. TOP-20 ранжированного списка научных журналов, полученного методом скользящей суммы по двум точкам.

№	Страна	Код ISSN	Название
1	Россия	0869-5652	Доклады Российской академии наук
2	США	1046-6770	Proceedings of SPIE
3	Германия	0341-5775	Maschinenmarkt (MM)
4	Великобритания	0028-0836	Nature: International Weekly Journal of Science (Gr. Brit.)
5	США	0007-7135	Business Week. European Edition
6	Германия	0344-6166	Produktion: Die Wirtschaftszeitung für die deutsche Industrie
7	Германия	0042-1758	VDI-Nachrichten
8	США	0036-8075	Science
9	США	0003-6951	Applied Physics Letters
10	Украина	1025-6415	Доповіді Національної академії наук України
11	США	0027-8424	Proceedings of the National Academy of Sciences of the United States of America
12	Россия	1998-8621	Вестник Томского государственного университета (ТГУ)

13	Германия	0341-2636	F + H: Fördern und Heben: Zeitschrift für Materialfluss und Warenwirtschaft
14	Россия	0367-6765	Известия Российской академии наук (РАН). Сер. физическая
15	Россия	1729-3642	Экономика и финансы
16	Россия	0236-1493	Горный информационно- аналитический бюллетень
17	США	0031-9007	Physical Review Letters
18	США	0021-8979	Journal of Applied Physics
19	Германия	0012-0901	DVZ: Deutsche Logistik-Zeitung
20	США	0018-9464	IEEE Transactions on Magnetics

Второй вариант скользящего отрезка (по двум точкам) представляется более адекватным решаемой задаче, так как повышается ранг журналов, получивших значительное количество оценок «высокая» и «выше средней», при этом «средние» оценки не влияют на сумму оценок по первому отрезку шкалы.

Вверх поднялись следующие издания:

№	Название журнала	пред. №	Вектор частот
12	Вестник Томского государственного университета (ТГУ)	35	(9, 3, 1, 4, 9)
13	F + H: Fördern und Heben: Zeitschrift für Materialfluss und Warenwirtschaft	28	(2, 4, 2, 10, 4)
15	Экономика и финансы	27	(1, 0, 3, 6, 7)
17	Physical Review Letters	25	(7, 5, 4, 3, 9)
19	DVZ: Deutsche Logistik-Zeitung	40	(3, 1, 1, 4, 8)
20	IEEE Transactions on Magnetics	33	(4, 3, 3, 5, 7)

Из первой двадцатки спустились вниз следующие издания:

№	Название журнала	пред. №	Вектор частот
23	Electronics Letters	15	(2, 5, 7, 7, 5)
26	Journal of Materials Science	11	(7, 8, 12, 5, 6)
41	Technica (Suisse)	14	(3, 5, 10, 6, 4)

54	Journal of Crystal Growth	20	(5, 3, 10, 4, 5)
63	Известия высших учебных заведений (вузов). Северо-Кавказский регион. Технические науки	16	(21, 15, 11, 5, 4)
138	Materials Science and Engineering. A. Structural Materials: Properties, Microstructure and Processing	18	(10, 8, 13, 3, 4)

4. Заключение

Несмотря на существенную доработку, в основе процедуры агрегирования неполных результатов многоаспектной экспертной стратификации по-прежнему лежит вербально-числовая шкала, суммирование оценок и лексикографическое упорядочение, но теперь числовые оценки получили вероятностное обоснование, а суммирование выполняется внутри скользящих отрезков шкалы.

Агрегирование результатов экспертной стратификации состоит из следующих этапов:

1. сбор тематических экспертных списков, элементами которых являются пары значений (a, x) , где a – уникальный идентификатор объекта, x – порядковый номер класса;

2. построение абсолютной шкалы на множестве порядковых экспертных оценок, задающей соответствие между номером упорядоченного класса и числовой оценкой класса, соответствующей его коэффициенту важности (весовому коэффициенту);

3. замена порядковых экспертных оценок на весовые коэффициенты;

4. выбор длины скользящего отрезка на шкале;

5. вычисление скользящих сумм весовых коэффициентов для каждого объекта;

6. лексикографическое упорядочение объектов по убыванию полученных сумм;

7. дополнительное упорядочение объектов внутри образовавшихся кластеров на основании результатов суммирования

весовых коэффициентов более (или менее) предпочтительных объектов в экспертных списках.

В заключение благодарю редакцию за предоставленную мне возможность опубликовать результаты работы и организацию обсуждения, а также всех участников дискуссии, пожертвовавших своё время на прочтение моей статьи и написание отзывов.

Литература

1. АБАЕВ Л. Ч. *Об одном подходе к использованию экспертных методов для оценки научных журналов* // Управление большими системами. – 2009. – Выпуск 27. – С. 36 – 46.
2. БЛЮМИН С. Л., ШУЙКОВА И. А., САРАЕВ П. В., ЧЕРПАКОВ И. В. *Нечеткая логика: алгебраические основы и приложения: Монография.* – Липецк: ЛЭГИ, 2002. – 113 с.
3. ГОРСКИЙ П. *Введение в дисциплину «Поддержка принятия решений».*– [Электронный ресурс].– Режим доступа: <http://www.gorskiy.ru/Articles/Dmss/d0.html>
4. ЛИТВАК Б. Г. *Экспертные технологии в управлении: учебн. пособие.* – 2-е изд., испр. и доп. – М.: Дело, 2004. – 400 с.
5. МИРКИН Б. Г. *О статье «Коллективная экспертиза научных журналов: методика агрегирования экспертных оценок и построения рейтинга»* // Управление большими системами. – 2009. – Выпуск 27. – С. 53 – 58.
6. ОРЛОВ А. И. *Теория принятия решений: учеб. для вузов.* – М.: Экзамен, 2006. – 576 с.
7. ОРЛОВ А. И. *Методологические ошибки ведут к неправильным управленческим решениям* // Управление большими системами. – 2009. – Выпуск 27. – С. 59 – 65.
8. ФЕДОРЕЦ О. В. *Коллективная экспертиза научных журналов: методика агрегирования экспертных оценок и построения рейтинга* // Управление большими системами. – 2009. – Выпуск 27. – С. 12 – 35.

9. ЧЕБОТАРЕВ П. Ю. *Метод строчных сумм и приводящие к нему модели* // Сборник трудов ВНИИ системных исследований.– 1989.– №3.– С. 94 – 110.

AGGREGATING RESULTS OF EXPERT STRATIFICATION BYPASSING MATRICES OF PAIRED COMPARISON: CASE OF SCIENTIFIC JOURNALS RATING

Oleg V. Fedorets, All-Russian institute of scientific and technical information, senior research officer (ovf@viniti.ru).

Abstract: One of the distinctive features of a scientific journal is its multidisciplinary character that demands estimations from the experts in various scientific fields. Another feature is that each expert estimates the set of journals; therefore a matrix of expert estimations turns out to be incomplete. The routine is offered for absolute scale construction based on an expert estimations set; moving totals method is suggested for journals lexicographic ordering.

Keywords: expertise organization, scientific journals rating, aggregation of expert estimations, incomplete expert data, numerical scale construction, moving total.

Статья представлена к публикации членом редакционной коллегии Д. А. Новиковым