

УДК 005.935.3+025.2
ББК 65.050.2-73

КОЛЛЕКТИВНАЯ ЭКСПЕРТИЗА НАУЧНЫХ ЖУРНАЛОВ: МЕТОДИКА АГРЕГИРОВАНИЯ ЭКСПЕРТНЫХ ОЦЕНОК И ПОСТРОЕНИЯ РЕЙТИНГА

Федорец О. В.¹

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

Описана методика сбора и агрегирования экспертных оценок для построения экспертного рейтинга научных журналов в области естественных и технических наук. 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. Таким образом, предлагается подход к решению задачи, отличающийся от типовых задач экспертного оценивания, в которых каждый эксперт оценивает каждый объект, и все эксперты являются специалистами в одинаковой предметной области.

Ключевые слова: комплектование информационного центра, рейтинг научного журнала, коллективная экспертиза, лексикографическое упорядочение.

¹ Федорец Олег Владимирович, старший научный сотрудник (ovf@viniti.ru).

1. Введение

Объектом экспертного оценивания обычно является альтернатива в целом (проект, товар, фирма, информационный ресурс и т.п.) или критерии оценивания альтернативы. Экспертов привлекают в тех случаях, когда точно измерить объект невозможно, но можно дать ему качественную оценку.

Ответы экспертов часто носят нечисловой характер. Эксперт может сравнить два объекта, сказать, какой из двух лучше, дать им оценки типа «хороший», «приемлемый», «плохой», упорядочить несколько объектов по привлекательности, но обычно не может ответить, во сколько раз или на сколько один объект лучше другого [10].

Экспертная оценка, как правило, отображается на порядковую шкалу. При этом градации на порядковой шкале не обязательно обозначаются натуральными числами – количеством баллов или номером ранга.

Например, в экономике и финансах популярны порядковые шкалы с буквенным обозначением градаций. В рейтингах банков, корпораций, ценных бумаг и т.п. часто можно увидеть шкалу вида <AAA, AA, A, BBB, BB, BB> и т.д. При этом промежуточные градации обозначают с использованием знаков плюс и минус, например AA+ и BBB-. Тем самым подчёркивается, что результат оценки носит нечисловую природу.

Если объектом оценивания является альтернатива, то в результате агрегирования экспертных оценок создаётся упорядоченный (ранжированный) список объектов по убыванию качества. Если эксперт оценивает критерии, то на основании агрегирования экспертных данных определяется приоритет (вес) критериев, а обобщённые оценки альтернатив вычисляются с помощью свёртки критериев. Наиболее простые виды свёрток – линейная и мультипликативная.

Полученная в результате агрегированная оценка альтернативы может называться по-разному: обобщённый критерий, интегральный критерий, целевой критерий (или фокус цели)

суперкритерий, рейтинг и т.д. По убыванию обобщённого критерия строится ранжированный список альтернатив (рейтинг-лист), который затем для дополнительной проверки можно предъявить экспертам, чтобы получить качественную оценку ранжирования и определить пути совершенствования методики взвешивания критериев и вычисления рейтинга. Иногда рейтингом считается номер ранга (абсолютный или нормированный) в итоговом ранжированном списке.

Наиболее известные методы получения качественных оценок, описанные в [7], приведены в таблице 1.

Таблица 1. Методы получения качественных оценок

| Метод | Результат |
|-----------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Парное сравнение | Матрица парных сравнений объектов |
| Ранжирование | Ранжированный список объектов |
| Экспертная классификация | Оценки объектов по порядковой шкале, градациям которой соответствуют вербальные описания. |
| Метод векторов предпочтений | Вектор, первая компонента которого – число объектов, которые превосходят первый объект, вторая компонента – число объектов, которые превосходят второй объект, и т.д. |

Используя экспертную классификацию, вербальные названия для градаций порядковой шкалы желательно выбирать такими, чтобы они были привычными и примерно одинаково трактовались всеми участниками экспертизы. В сфере образования привычна шкала *<отлично, хорошо, удовлетворительно, неудовлетворительно>*. На основании сравнения с ассоциативным понятием «средний» можно построить следующие шкалы:

- *<высокий, выше среднего, средний, ниже среднего, низкий>*;
- *<очень высокий, высокий, средний, низкий, очень низкий>*;

- *<очень высокий, высокий, выше среднего, средний, ниже среднего, низкий, очень низкий>*.

Выбор метода получения качественных оценок определяется в первую очередь количеством и сложностью оцениваемых объектов.

Парные сравнения объектов имеет смысл использовать, если объектов немного (менее десяти). Количество элементов, которые эксперт должен ввести в матрицу размера $n \times n$, вычисляется по формуле $(n^2 - n)/2$. Остальные ячейки можно заполнить автоматически. Примеры матриц парного сравнения критериев по шкале отношений можно найти в [1, 11]. Таким же способом можно сравнить небольшое количество альтернатив, если эксперты в состоянии оценить их в целом, не уточняя приоритеты критериев.

Размеры ранжированных списков объектов также целесообразно ограничивать. Эксперту вряд ли стоит предлагать упорядочить более 30 объектов, даже если он будет сначала разбивать их на группы по убыванию качества, а затем ранжировать внутри групп. Причина не только в технической трудности ранжирования – эксперт может просто не обладать достаточным объёмом знаний о значительном количестве сложных объектов.

Экспертная классификация позволяет осуществлять экспертную оценку десятков и даже сотен объектов. Но в случае их большого количества эксперту необходимо помимо наименований объектов предоставить необходимую информацию (статистическую, описательную), на основании которой он может сделать логические выводы и оценить объект.

Во многих задачах сложность объекта оценивания требует строить иерархию критериев, которая позволяет оценить объект с различных сторон и вычислить общий рейтинг в результате иерархического синтеза частных рейтингов. В таких случаях эксперты привлекаются для определения приоритета (веса) критериев и частных рейтингов.

Традиционным результатом экспертного оценивания, который необходимо обрабатывать и агрегировать, является матрица

индивидуальных экспертных оценок (a_{ij}), элементом которой является оценка i -го объекта j -м экспертом. В случае применения экспертной классификации или ранжирования элементом матрицы является оценка в порядковой шкале (обозначение класса или номер ранга соответственно). У подобных задач можно выделить две характерные особенности:

1. полнота – каждый объект оценивается каждым экспертом, следовательно, все элементы матрицы индивидуальных экспертных оценок являются непустыми;
2. одноаспектность – эксперты являются специалистами в одинаковой предметной области и оценивают объект в одном аспекте (хотя их субъективное мнение может различаться, как может различаться и квалификация экспертов, уровень которой учитывается в некоторых моделях экспертного оценивания).

После сбора экспертных оценок в первую очередь определяется их согласованность. Если оценки несогласованные, отдельным экспертам предлагают заново оценить объекты. Если согласованные, то следующая задача – получить усреднённую оценку объекта или усреднённое ранжирование объектов, которые при этом вычисляются не как «среднее арифметическое», а более сложными методами, используемыми для «усреднения» нечисловых оценок. Описание различных методов агрегирования экспертных оценок можно найти в [1, 7, 10], обзор методов построения рейтингов – в [9].

В теории экспертного оценивания разработано немало методов агрегирования данных нечислового характера: метод Борда, медиана рангов, медиана Кемени, использование verbally-числовой шкалы для перевода качественных оценок в числовые, построение функций предпочтения и т.д.

Какой метод использовать в случае многоаспектного и неполного оценивания, когда каждый эксперт оценивает не все, а лишь некоторые известные ему объекты, причём с позиции своего аспекта, т.е. своей тематики или специальности? В таком случае в матрице (a_{ij}) допустимы пустые элементы, причём

таких элементов может быть большинство. С одной стороны, процедура агрегирования неполной матрицы оценок больше напоминает обработку результатов голосования, при котором необходимо получить некую суммарную оценку, т.е. сумму экспертных голосов. С другой стороны, наличие высокой экспертной оценки, данной объекту хотя бы одним экспертом, может оказаться более приоритетным, чем большая сумма, сложенная из средних и невысоких экспертных оценок.

Далее мы рассмотрим экспертизу научных журналов, которые в силу своей многочисленности и политематического характера приводят к неполному и многоаспектному экспертному оцениванию.

2. Постановка задачи

Рассмотрим экспертное оценивание сериальных (т.е. периодических и продолжающихся) изданий, которое проводилось во Всероссийском институте научной и технической информации Российской академии наук (ВИНИТИ) в 2008 г. для оптимизации входного потока научно-технической литературы.

Материалы из зарубежных и отечественных сериальных изданий составляют в среднем 73% всего документального потока, отражаемого в базах данных и реферативных журналах (РЖ) ВИНИТИ. При этом их доля колеблется в составе различных выпусков РЖ от 91 до 46% [3].

Одной из основных задач комплектования ВИНИТИ является совершенствование списка важнейших российских и зарубежных журналов, необходимых для обеспечения подготовки Реферативного журнала и Базы данных ВИНИТИ с требуемым уровнем отражения зарубежных публикаций и наиболее полного отражения отечественных и русскоязычных журналов [2].

Нельзя сказать, что до сих пор список журналов, обрабатываемых в ВИНИТИ, формировался стихийно. Сотрудники отдела комплектования не реже одного раза в год проводят мониторинг мирового потока научной литературы и вводят в базу

данных описания новых изданий. Ценным источником информации о периодических и продолжающихся изданиях является электронный справочник *Ulrich's Periodicals Directory*, используются также сайты издательств и подписных агентств.

Информация о новых изданиях поступает в отраслевые отделы, которые отмечают заинтересовавшие их издания. Затем комплектаторы получают образцы отмеченных изданий и направляют их в профильный отраслевой отдел на первичную экспертную оценку. В случае положительной экспертной оценки издания обеспечивается его регулярное поступление в ВИНТИ в течение года. В результате накапливается статистика отражения статей в РЖ, что позволяет впоследствии выполнять вторичную оценку издания – с учётом статистики отражения в реферативных базах данных.

В условиях неуклонного роста мирового потока сериальных изданий необходим более обоснованный механизм вторичной экспертной оценки, опирающийся на учёт мнений всех заинтересованных редакторов, в том числе, из различных отраслевых отделов.

Группа экспертов состояла из 237 редакторов реферативных журналов, работающих в 16-ти отраслевых отделах ВИНТИ. Экспертам были выданы из базы данных списки сериальных изданий, ранжированные по отражаемости в РЖ (всего 237 наименований РЖ, соответственно столько же тематических списков). По каждому изданию была предоставлена статистическая информация о тематической разметке и отражении в РЖ, а также формальные критерии, свидетельствующие о цитировании, отражении зарубежными реферативными службами, включении в список ВАК и т.д. Тематические списки были выданы экспертам в виде файлов в формате *MS Excel*, в которых они поставили оценку напротив каждого сериального издания [5].

При определении методики экспертизы учитывалось, что экспертная оценка носит субъективный характер, а экспертные суждения обычно выражаются в качественном, а не в количест-

венном виде [10]. Поэтому редакторам РЖ была поставлена задача экспертной классификации – присвоение каждому сериальному изданию порядкового номера, означающего желательность его обработки в ВИНТИ и отражения в РЖ. Порядковый номер может принимать одно из шести значений: высокий (5), выше среднего (4), средний (3), ниже среднего (2), низкий (1), нулевой (0).

Отрицательные оценки в данной шкале отсутствуют. Нулевую оценку нельзя трактовать как отрицательную, скорее она означает отказ эксперта присвоить оценку по одной из двух причин:

- журнал выдан эксперту случайно и не соответствует его тематике;
- журнал соответствует тематике эксперта, но статьи по его тематике в этом журнале носят научно-популярный характер.

В результате 11067 сериальных изданий получили ненулевую оценку хотя бы одного эксперта. Большинство изданий (9384) периодические, остальные – продолжающиеся. Количество изданий в тематических списках варьировалось от 18 до 831, в среднем 180 изданий в одном списке.

Существенная особенность экспертной оценки научного журнала – невозможность чётко разделить её на оценку качества и оценку использования (спроса), так как она аккумулирует в себе обе составляющие. Эксперт принимает во внимание не только среднее качество статей, но и примерное количество статей в год, поэтому баллы, выставленные экспертом, можно трактовать как субъективную оценку величины КАЧЕСТВО × СПРОС.

После сбора тематических экспертных оценок задача состояла в том, чтобы агрегировать их и построить политематический ранжированный список сериальных изданий. Ранг, полученный сериальным изданием в итоговом ранжированном списке, считался его экспертным рейтингом.

3. Зарубежный и отечественный опыт оценки журналов

В 1970 г. R. L. Zwemer предложил следующий список характеристик «хорошего журнала» [15]:

1. Высокие стандарты принятия рукописей.
2. Широкая представительная редакционная коллегия.
3. Редактор использует систему рецензирования.
4. Быстрота публикации.
5. Покрытие большинством реферативных и индексных служб.
6. Высокий уровень доверия содержимому журнала со стороны учёных, использующих его статьи.
7. Высокая частота цитирования другими журналами.

Основные принципы отбора журналов были определены ещё тогда, когда индекс цитирования и ежегодный отчёт о цитировании журналов (*Journal Citation Report*) издавались Институтом научной информации (*Institute of Scientific Information, ISI*), базирующемся в Филадельфии. В дополнение к перечисленным критерием в *ISI* были добавлены следующие требования [13]:

8. Включение рефератов или резюме на английском языке.
9. Включение адресной информации об авторах.
10. Предоставление полной библиографической информации об используемых источниках.

В 2005 г. персонал *Thomson ISI* (подразделение корпорации *Thomson*, которая с 2008 г. называется *Thomson Reuters*) просматривал около 2000 новых журналов, но только 10-12% из них отбирались для включения в базу данных цитирования *Web of Science*. Помимо уже перечисленных критериев, уделяется внимание своевременному выходу номеров издания, географической распространённости авторов статей, а также авторов цитируемых ими статей (т.е. приоритетом пользуются издания международного уровня) [16].

Обзор критериев цитирования и зарубежной литературы по данной тематике можно найти в [15], где приведены формулы для вычисления различных импакт-факторов на основе статистики цитирования, в том числе формула для вычисления стандартного импакт-фактора *ISI*, ежегодно публикуемого в *Journal Citation Reports*.

Импакт-факторы – это всегда коэффициенты в форме «количество полученных ссылок, делённое на количество публикаций». Импакт-фактор рассчитывается в отношении определённого множества журналов, покрываемых базой данных цитирования [15]. Помимо *Thomson*, отдающей предпочтение журналам на английском языке (свыше 85%), в некоторых странах (Китай, Япония, Россия) созданы свои национальные базы данных научного цитирования и публикуются импакт-факторы.

Обзор критериев оценки качества журналов, не основанных на статистике цитирования, опубликован в [12]. Там же можно найти ссылки на публикации, в которых исследуется ограниченность подхода, при котором импакт-фактор используется в качестве единственного измерителя качества журнала.

Обзор статей, посвящённых оценке и ранжированию журналов по библиотековедению и информатике, опубликовал Т. Е. Nisonger [14]: 178 статей за 50 лет разделяются на 5 категорий по используемым критериям ранжирования журналов: цитирование (94); продуктивность, т.е. количество статей в год (33); субъективное суждение, т.е. экспертная оценка (25); читательский спрос (18); смешанные критерии (8).

Nisonger пишет о двух известных подходах к экспертному оцениванию журналов:

- оценка по шкале, подобной шкале Лайкерта («*Likert-like scale, such as 1-5 or 0-10*»);
- подход «*top-n*», при котором эксперту предлагается составить неупорядоченный список из *n* «наиболее важных» журналов в его предметной области (3, 5, 10 и т.д.)

В первом случае журналы ранжируются по убыванию суммы экспертных оценок, во втором – по убыванию количества вхождений в экспертные списки.

В [4] описана математическая модель формирования оптимального списка журналов, приобретаемых Библиотекой по естественным наукам Российской академии наук (БЕН РАН). Вводится понятие «информативность журнала», которая вычисляется с помощью линейной свёртки критериев читательского спроса, экспертной оценки и цитирования.

Система формирования заказа на журналы в БЕН РАН предполагает реализацию следующих условий [4]:

- обеспечение максимальной информативности репертуара периодических изданий в рамках имеющихся финансовых средств (решается как задача линейного программирования);
- поддержка сбалансированного обеспечения литературой различных научных направлений, определённых тематическим планом комплектования библиотеки в соответствии с их значимостью (разработана математическая модель распределения денежных средств с учётом значимости направлений и коэффициентов стоимости изданий по каждому направлению).

БЕН РАН собирает экспертные оценки через свой сайт <http://www.ben.ran>. Авторизованные эксперты имеют возможность знакомиться с информацией об изданиях и оценивать каждое из них с точки зрения целесообразности приобретения его в фонды библиотеки своего института или Центральной библиотеки. В 2007 г. база данных экспертов включала 410 учёных из 88 научно-исследовательских институтов РАН [6].

Изучение отечественного и зарубежного опыта оценки научных журналов по материалам открытой печати привело к выводу, что математические методы активно используются для обработки статистических данных цитирования и читательского спроса, которые накапливаются в базах данных информационных служб и библиотек. Однако не удалось найти описаний

математических моделей, применяемых для экспертного оценивания и агрегирования экспертных оценок в многоотраслевых информационных службах или библиотеках.

4. Алгоритм ранжирования

От экспертов были получены и загружены в базу данных экспертные оценки в соответствии с порядковой шкалой. Для вычисления обобщённой оценки порядковые номера было решено заменить числовыми значениями в соответствии с вербально-числовой шкалой.

За основу была взята вербально-числовая шкала Харрингтона, которая достаточно широко применяется в экспертных оценках для характеристики степени выраженности критериального свойства. Шкала (см. таблицу 2) имеет универсальный характер и может в соответствующих модификациях использоваться для оценки показателей качественного характера [7].

Таблица 2. Шкала Харрингтона

| № п/п | Содержательное описание градаций | Численное значение |
|-------|----------------------------------|--------------------|
| 1 | Очень высокая | 1,0 – 0,8 |
| 2 | Высокая | 0,8 – 0,63 |
| 3 | Средняя | 0,63 – 0,37 |
| 4 | Низкая | 0,37 – 0,2 |
| 5 | Очень низкая | 0,2 – 0,0 |

В разработанной модификации шкалы (таблица 3) интервалы заменены числами, каждое из которых является средним арифметическим значением границ интервала, а также добавлен нулевой уровень.

Таблица 3. Вербально-числовая шкала для агрегирования экспертных оценок сериальных изданий.

| № уровня | Содержательное описание уровня | Численное значение | Вычисление численного значения |
|----------|--------------------------------|--------------------|--------------------------------|
| 5 | Высокий | 0,9 | $= (1,0 + 0,8) / 2$ |
| 4 | Выше среднего | 0,715 | $= (0,8 + 0,63) / 2$ |
| 3 | Средний | 0,5 | $= (0,63 + 0,37) / 2$ |
| 2 | Ниже среднего | 0,285 | $= (0,37 + 0,2) / 2$ |
| 1 | Низкий | 0,1 | $= (0,2 + 0,0) / 2$ |
| 0 | Нулевой | 0 | $= 0$ |

Результатом агрегирования экспертных оценок является «Обобщённая оценка», для получения которой суммируются численные оценки уровней, полученные от всех экспертов. Например, журнал получил четыре оценки от редакторов различных наименований РЖ:

«Список оценок» = 5; 5; 4; 3

В результате вычисляются два критерия:

«Максимальная оценка» = $\max(5; 5; 4; 3) = 5$

«Обобщённая оценка» = $(0,9+0,9+0,715+0,5) = 3,015$

Алгоритм агрегирования оценок и ранжирования журналов состоит из семи последовательных шагов.

Шаг 1. Собрать индивидуальные экспертные оценки в матрицу (a_{ij}) размера $n \times m$, каждый элемент которой является качественной экспертной оценкой i -го журнала j -м экспертом.

Шаг 2. Построить матрицу (b_{ij}) , которая получается путём замены каждого элемента матрицы (a_{ij}) числом в соответствии с вербально-числовой шкалой (см. таблицу 3).

Шаг 3. Для каждого i -го журнала вычислить обобщённую оценку S_i по формуле:

$$S_i = \sum_{j=1}^m b_{ij} .$$

Шаг 4. Для каждого i -го журнала вычислить максимальную оценку M_i по формуле:

$$M_i = \max_j a_{ij} .$$

Шаг 5. Построить матрицу (r_{ij}) размера $n \times m$, в которой каждый элемент представляет собой номер ранга i -го журнала в ранжированном списке j -го эксперта. Ранжирование осуществляется по убыванию двух значений: a_{ij} (экспертная оценка) и S_i (обобщённая оценка). Если в ранжированном списке j -го эксперта оказываются группы совпадающих (a_{ij}, S_i) , их ранги заменяются средним арифметическим рангов группы.

Шаг 6. Для каждого i -го журнала вычислить высший нормированный ранг R_i по формуле:

$$R_i = 1 - \min_j (r_{ij} / \max_i r_{ij}) .$$

Шаг 7. Ранжировать журналы методом лексикографического упорядочения по трём критериям: M_i , S_i и R_i

Лексикографическое упорядочение (*lexicographic ordering*) – упорядочение объектов (в многокритериальной задаче, задаче выявления предпочтений) таким образом, что, например, объект a' предпочитается объекту a'' , если он имеет большую оценку по наиболее важному критерию x_1 , невзирая на то, насколько он является хорошим или же плохим по другим менее важным критериям. Но если значения x_1 для них совпадут, вводится в рассмотрение следующий по важности критерий x_2 и по нему выбирается предпочитаемый объект. Соответственно, в случае совпадения оценок по критериям x_1, x_2 вводится критерий x_3 и т. д. Определение «лексикографическое» объясняется тем, что эта процедура напоминает построение словаря [8].

5. Разработка алгоритма ранжирования

Описанный выше алгоритм был разработан в результате экспериментальных компьютерных сортировок по различным критериям и добавления новых критериев по мере необходимо-

сти, т.е. имитационного моделирования и последовательного улучшения промежуточных результатов.

В первоначальном ранжированном списке журналы сортировались одному критерию – по убыванию обобщённой оценки. Однако анализ списка журналов показал неадекватность такого ранжирования. Преимущество получили журналы, относящиеся к отраслям науки, по которым издаётся значительное количество выпусков РЖ, имеющих смежную тематику. В первую очередь это такие крупные отрасли, как Биология, Химия и Физика. В то же время журналы, относящиеся к менее крупным и более изолированным отраслям (Транспорт, Экономика промышленности, Информатика, Астрономия) подверглись дискриминации. Несмотря на высокие оценки, присвоенные отдельными экспертами, в итоговом списке такие журналы оказались значительно ниже журналов, получивших множество средних оценок от экспертов смежных научных тематик. Подобный отбор «популярных» журналов в ущерб узкоспециализированным журналам, высоко оценённым по отдельным научным тематикам, не обеспечивал сбалансированное покрытие всех тематик качественными сериальными изданиями.

В результате был добавлен второй критерий. Использование лексикографического упорядочения журналов по двум критериям – «максимальная оценка» и «обобщённая оценка» – значительно улучшило результат, обеспечив более равномерное распределение ранжированного списка по тематикам. Теперь ранжированный список представляет собой пять упорядоченных кластеров: на первом месте кластер журналов, оцененных как «высокий» хотя бы одним экспертом, затем кластер журналов, оцененных «выше среднего» хотя бы одним экспертом, и т.д. Внутри каждого кластера журналы сортируются по убыванию обобщённой оценки, которая агрегирует в себе все экспертные оценки журнала.

Конечно, существует риск, что эксперт может поставить высокую оценку журналу случайно. Присвоена хотя бы одна высокая оценка – и журнал автоматически оказывается в лиди-

рующем кластере. В этом случае журнал может быть ошибочно включён в подписку, запущен в технологию обработки, и ошибка будет исправлена не раньше, чем через год. Однако многоотраслевой реферативной службе, стремящейся отражать наиболее передовые научные издания в каждом тематическом РЖ, гораздо важнее не упустить ценный монотематический журнал. Подписчики РЖ быстрее заметят, что реферативная служба не отражает какой-либо авторитетный журнал в их области, чем обратят внимание на то, что в список реферируемых изданий включён малоизвестный журнал.

Несомненно, что при выделении достаточных средств на организацию экспертизы необходимо расширять группу экспертов – привлекать несколько специалистов по каждой тематике (например, внештатных референтов) и усреднять тематическую оценку. Тогда влияние одного эксперта было бы сведено к минимуму. Впрочем, статистика экспертных оценок показала, что и сейчас влияние одного эксперта невелико – в лидирующем кластере оказалось 45,5% журналов, получивших хотя бы одну высокую оценку. В следующие четыре кластера включены соответственно 21,1%, 16,3%, 10,6%, и 6,4% журналов.

Экспериментальное ранжирование показало, что двух критериев недостаточно, так как для некоторых пар значений $(S_i; M_i)$ образовались значительные кластеры журналов с повторяющимися значениями. Количество повторяющихся значений для некоторых пар $(S_i; M_i)$ исчислялось сотнями, поэтому возник вопрос, в каком порядке сортировать журналы внутри этих кластеров.

В итоге было принято решение добавить третий критерий, который был назван «высшим рангом».

Сначала рассмотрим методику вычисления ненормированного «высшего ранга» и покажем его ограниченность. Для вычисления «высшего ранга» были построены отсортированные тематические списки сериальных изданий, получивших экспертные оценки от 1 до 5. Количество тематических списков равнялось количеству экспертов. Сортировка списков произво-

дилась согласно лексикографическому упорядочению по двум критериям: «экспертная оценка» и «обобщённая оценка». Сериальные издания в каждом тематическом списке были пронумерованы порядковыми номерами (рангами), начиная с единицы. Если группа журналов имеет совпадающие значения по обоим критериям, то ранги заменяются средним арифметическим значением рангов данной группы.

В результате значение «высший ранг» вычисляется как минимальное значение ранга сериального издания по всем ранжированным тематическим спискам. Например, если некое сериальное издание присутствует в трёх ранжированных списках на позициях 12, 15 и 30, то «высший ранг» = 12.

Очевидно, что абсолютные значения рангов не учитывают размеры тематических ранжированных списков, которые, как было указано выше, варьировались в широком диапазоне (от 18 до 831). Поэтому вместо рангов необходимо использовать нормированные ранги. К сожалению, идея нормировать тематические ранги пришла уже после того, как экспертные оценки были агрегированы, а ранжированный список окончательно проверен и утверждён. Однако в дальнейшем для вычисления «высшего ранга» планируется использовать нормированные тематические ранги. Нормированный ранг предлагается вычислять как отношение тематического ранга к максимальному рангу тематического списка. Расширим предыдущий пример: пусть некое издание получило в трёх ранжированных списках ранги 12, 15 и 30, при этом максимальные ранги списков равны соответственно 40, 60 и 300. Тогда «высший нормированный ранг» равен $1 - \min(12/40; 15/60; 30/300) = 0,9$.

Таким образом, в результате вычисления новых критериев и экспериментальных сортировок был разработан алгоритм, описанный в предыдущем разделе.

6. Пример

Продemonстрируем алгоритм ранжирования на простом примере оценивания 8-ми объектов, которые получили оценки 6-ти экспертов, обозначенных как $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_6$ (см. таблицу 4). Объектам для наглядности присвоены имена – латинские буквы от A до H (хотя подразумевается, что каждый объект имеет порядковый номер – индекс i). Если упорядочить объекты по убыванию суммы экспертных оценок, то получится следующая последовательность: $F > B > A = C > D = E > H > G$. Далее мы увидим, что результатом выполнения предложенного алгоритма лексикографического упорядочения является другая последовательность: $F > B > C > A > H > D > E > G$.

Таблица 4. Индивидуальные экспертные оценки (шаг 1)

| i | Объект | \mathcal{E}_1 | \mathcal{E}_2 | \mathcal{E}_3 | \mathcal{E}_4 | \mathcal{E}_5 | \mathcal{E}_6 | M_i |
|-----|--------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-------|
| 1 | A | 5 | | 4 | | 3 | 1 | 5 |
| 2 | B | 5 | 2 | | 4 | | 3 | 5 |
| 3 | C | | 3 | 5 | | 5 | | 5 |
| 4 | D | | 4 | 3 | 4 | | | 4 |
| 5 | E | 4 | 4 | | | 3 | | 4 |
| 6 | F | 3 | | 4 | 3 | | 5 | 5 |
| 7 | G | 1 | | | 3 | 2 | | 3 |
| 8 | H | | 5 | | | 4 | | 5 |

После замены экспертных оценок числами по вербально-числовой шкале получаем таблицу 5, в которой в последней колонке вычислена обобщённая оценка S_i .

Построенная на шаге 5 матрица рангов и вычисленные на шаге 6 высшие нормированные ранги можно увидеть в таблице 6.

Таблица 5. Числовая интерпретация экспертных оценок (шаг 2)

| i | Объект | \mathcal{E}_1 | \mathcal{E}_2 | \mathcal{E}_3 | \mathcal{E}_4 | \mathcal{E}_5 | \mathcal{E}_6 | S_i |
|---|--------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-------|
| 1 | A | 0,9 | | 0,715 | | 0,5 | 0,1 | 2,215 |
| 2 | B | 0,9 | 0,285 | | 0,715 | | 0,5 | 2,4 |
| 3 | C | | 0,5 | 0,9 | | 0,9 | | 2,3 |
| 4 | D | | 0,715 | 0,5 | 0,715 | | | 1,93 |
| 5 | E | 0,715 | 0,715 | | | 0,5 | | 1,93 |
| 6 | F | 0,5 | | 0,715 | 0,5 | | 0,9 | 2,615 |
| 7 | G | 0,1 | | | 0,5 | 0,285 | | 0,885 |
| 8 | H | | 0,9 | | | 0,715 | | 1,615 |

Таблица 6. Ранги объектов (шаги 5 и 6)

| i | Объект | \mathcal{E}_1 | \mathcal{E}_2 | \mathcal{E}_3 | \mathcal{E}_4 | \mathcal{E}_5 | \mathcal{E}_6 | R_i |
|---|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-------|
| 1 | A | 1,5 | | 2,5 | | 3,5 | 3 | 0,7 |
| 2 | B | 1,5 | 5 | | 1,5 | | 2 | 0,7 |
| 3 | C | | 4 | 1 | | 1 | | 0,8 |
| 4 | D | | 2,5 | 4 | 1,5 | | | 0,57 |
| 5 | E | 3 | 2,5 | | | 3,5 | | 0,5 |
| 6 | F | 4 | | 2,5 | 3,5 | | 1 | 0,67 |
| 7 | G | 5 | | | 3,5 | 5 | | 0 |
| 8 | H | | 1 | | | 2 | | 0,8 |
| | $\max_i(r_{ij})$ | 5 | 5 | 4 | 3,5 | 5 | 3 | |

Высшие нормированные ранги объектов вычисляются следующим образом:

$$R_1 = 1 - \min(1,5/5; 2,5/4; 3,5/5; 3/3) = 0,7 ;$$

$$R_2 = 1 - \min(1,5/5; 5/5; 1,5/3,5; 2/3) = 0,7 ;$$

$$R_3 = 1 - \min(4/5; 1/4; 1/5) = 0,8 ;$$

$$R_4 = 1 - \min(2,5/5; 4/4; 1,5/3,5) = 0,57 ;$$

$$R_5 = 1 - \min(3/5; 2,5/5; 3,5/5) = 0,5 ;$$

$$R_6 = 1 - \min(4/5; 2,5/4; 3,5/3,5; 1/3) = 0,67 ;$$

$$R_7 = 1 - \min(5/5; 3,5/3,5; 5/5) = 0 ;$$

$$R_8 = 1 - \min(1/5; 2/5) = 0,8.$$

В таблице 7 представлен итоговый ранжированный список объектов, полученный в результате лексикографического упорядочения по трём критериям – M_i , S_i , R_i .

Результат упорядочения: $F > B > C > A > H > D > E > G$.

Таблица 7 Результат выполнения алгоритма

| Ранг | Объект | M_i | S_i | R_i |
|------|----------|-------|-------|-------|
| 1 | <i>F</i> | 5 | 2,615 | 0,7 |
| 2 | <i>B</i> | 5 | 2,4 | 0,7 |
| 3 | <i>C</i> | 5 | 2,3 | 0,8 |
| 4 | <i>A</i> | 5 | 2,215 | 0,57 |
| 5 | <i>H</i> | 5 | 1,615 | 0,5 |
| 6 | <i>D</i> | 4 | 1,93 | 0,67 |
| 7 | <i>E</i> | 4 | 1,93 | 0 |
| 8 | <i>G</i> | 3 | 0,885 | 0,8 |

7. Заключение

В качестве инструментального средства для хранения данных, вычисления критериев и ранжирования журналов использовалась реляционная СУБД *Microsoft SQL Server* и её диалект языка *SQL*, который называется *Transact-SQL*. Итоговый ранжированный список сериальных изданий, выданный из базы данных, был окончательно уточнён отраслевыми экспертами и комплектаторами ВИНТИ. В результате был утверждён перечень журналов, которые обязательно должны отражаться в реферативных базах данных ВИНТИ в 2009 г.

Для практических целей комплектования были созданы отдельные ранжированные списки по видам изданий (периодические и продолжающиеся) и по географическому признаку (зарубежные и российские). В таблице 9 представлен фрагмент общего ранжированного списка – первые двадцать изданий, получившие наивысший экспертный рейтинг. В этом списке только первое сериальное издание является продолжающимся, остальные – периодические издания.

Очевидно, что на основании предложенного алгоритма можно создавать его различные модификации в зависимости от предметной области, целей экспертизы, состава экспертной группы. Модификации могут производиться двумя способами:

- 1) изменение вербально-числовой шкалы;
- 2) изменение порядка критериев при лексикографическом упорядочении.

Таблица 9. TOP-20 научных сериальных изданий в 2008 г. по результатам обработки экспертных оценок

| Ранг | Код ISSN | Страна | Название |
|------|-----------|----------------|----------------------------------------------------------------------------------------|
| 1 | 1046-6770 | США | Proceedings of SPIE |
| 2 | 0869-5652 | Россия | Доклады Российской академии наук |
| 3 | 0028-0836 | Великобритания | Nature: International Weekly Journal of Science (Gr. Brit.) |
| 4 | 0341-5775 | Германия | Maschinenmarkt (MM) |
| 5 | 0027-8424 | США | Proceedings of the National Academy of Sciences of the United States of America |
| 6 | 0036-8075 | США | Science |
| 7 | 0042-1758 | Германия | VDI-Nachrichten |
| 8 | 0344-6166 | Германия | Produktion: Die Wirtschaftszeitung fur die deutsche Industrie |
| 9 | 0321-2653 | Россия | Известия высших учебных заведений (вузов). Северо-Кавказский регион. Технические науки |
| 10 | 1025-6415 | Украина | Доповіді Національної академії наук України |
| 11 | 0022-2461 | США | Journal of Materials Science |
| 12 | 0003-6951 | США | Applied Physics Letters |

| | | | |
|----|-----------|----------------|-------------------------------------------------------------------------------------------------------|
| 13 | 0007-7135 | США | Business Week. European Edition |
| 14 | 1684-8519 | Россия | Известия Томского политехнического университета |
| 15 | 0921-5093 | Швейцария | Materials Science and Engineering. A. Structural Materials: Properties, Microstructure and Processing |
| 16 | 0236-1493 | Россия | Горный информационно-аналитический бюллетень |
| 17 | 0367-6765 | Россия | Известия Российской академии наук (РАН). Сер. физическая |
| 18 | 0040-6090 | Швейцария | Thin Solid Films |
| 19 | 0021-8979 | США | Journal of Applied Physics |
| 20 | 0013-5194 | Великобритания | Electronics Letters |

Литература

1. АНДРЕЙЧИКОВ А. В., АНДРЕЙЧИКОВА О. Н. *Анализ, синтез, планирование решений в экономике: учебник.* – 2-е изд., доп. и перераб. – М.: Финансы и статистика, 2004. – 464 с.
2. АРСКИЙ Ю. М. *ВИНИТИ в решении проблем современной информатики // НТИ-2007. Информационное общество, интеллектуальная обработка информации, информационные технологии. Материалы международной конференции.* – М.: ВИНТИ, 2007. – С. 3-5.
3. АРСКИЙ Ю. М., ЛЕОНТЬЕВА Т. М., НИКОЛЬСКАЯ И. Ю., ШОГИН А. Н. *Банк данных ВИНТИ: состояние и перспективы развития.* – М.: ВИНТИ, 2006. – 242 с.
4. КАЛЕНОВ Н. Е., КОЗЛОВА Е. И., ГИАЦИНТОВ О. М. *Математическая модель оптимизации подписки на журналы в научной библиотеке // Научно-техническая информа-*

- ция. Сер. 1. Организация и методика информационной работы. – М: ВИНТИ. – 1999. – №12. – С. 9-12.
5. КИРИЛЛОВА О. В., ФЕДОРЕЦ О. В. *Сбор и агрегирование экспертных оценок для ранжирования научных журналов* // Двенадцатая международная конференция LIBCOM-2008, п. Ершово, Одинцовский район, Московская область, 17-21 ноября 2008 г. [Электронный ресурс]. – Режим доступа: <http://www.gpntb.ru/libcom8/disk/22.pdf>
 6. КОЧУКОВА Е. В., НАУМОВА В. С. *Современные информационные технологии в практике комплектования ЦБС БЕН РАН* // НТИ-2007. Информационное общество, интеллектуальная обработка информации, информационные технологии. Материалы международной конференции. – М.: ВИНТИ, 2007. – С. 158-159.
 7. ЛИТВАК Б. Г. *Экспертные технологии в управлении: учеб. пособие.* – 2-е изд., испр. и доп. – М.: Дело, 2004. – 400 с.
 8. ЛОПАТНИКОВ Л. И. *Экономико-математический словарь: Словарь современной экономической науки.* — 5-е изд., перераб. и доп. — М.: Дело, 2003. – 520 с.
 9. НИКОЛАЕВА М. А., ЮНЦЕВИЧ О. Ф. *Методы и алгоритмы построения рейтингов* // Информационные технологии. – 2003. – №12. – С. 7-18.
 10. ОРЛОВ А. И. *Теория принятия решений: учеб. для вузов.* – М.: Экзамен, 2006. – 576 с.
 11. СААТИ Т. *Принятие решений. Метод анализ иерархий.* – М.: Радио и связь, 1989. – 316 с.
 12. COLEMAN A. *Assessing the Value of a Journal Beyond the Impact Factor* // Journal of the American Society for Information Science and Technology.– 2007.– №58(8).– P. 1148-1161.
 13. GARFIELD E. *How ISI selects journals for coverage. Quantitative and qualitative considerations* // Current Contents. – 1990. – №22. – P. 5-13.
 14. NISONGER T. E. *JASIS and library and information science journal rankings: A review and analysis of the last half-*

- century* // Journal of the American Society for Information Science.– 1999.– №50(11).– P. 1004-1019.
15. ROUSSEAU R. *Journal Evaluation: Technical and Practical Issues* // Library Trends. – 2002. – Vol. 50, №3. – P. 418-439.
16. TESTA J. *The Thompson Reuters Journal Selection Process.* – 2008. [Электронный ресурс]. – Режим доступа: http://www.thomsonreuters.com/business_units/scientific/free/essays/journalselection/

COLLECTIVE EXPERT EXAMINATION OF SCIENTIFIC JOURNALS: PROCEDURE OF EXPERT JUDGMENTS AGGREGATION AND RATING CONSTRUCTION

Oleg V. Fedorets, All-Russian Institute of Scientific and Technical Information, Senior research officer (ovf@viniti.ru).

Abstract: The technique of gathering and aggregation of expert estimations for the expert rating construction for scientific journals in natural and technical sciences is described. One of the distinctive features of 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. Thus, an approach is offered to this problem that differs from the typical process of expert estimation where each expert estimates each object, and all experts have expertise in the same subject area.

Keywords: acquisition of information centre, scientific journal rating, collective expert examination, lexicographic ordering.

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