

УДК 021.8 + 025.1
ББК 78.34

ТРУДНОСТИ ПОСТРОЕНИЯ СИСТЕМЫ РЕЙТИНГОВ НАУЧНЫХ ЖУРНАЛОВ

Сидельников Ю. В.¹

*(Учреждение Российской академии наук
Институт проблем управления РАН, Москва)*

Рассматриваются трудности, возникающие при строго формальном построении системы рейтингов научных журналов.

Ключевые слова: построение системы рейтингов научных журналов, трудности и проблемы.

1. Актуальность задачи

Поставленная в статье [9] О. В. Федорца задача – создать эффективно работающую процедуру, позволяющую достаточно строго ранжировать произвольное множество российских научных журналов – безусловно, актуальна и чрезвычайно сложна. Мне представляется, что решение ее невозможно до тех пор, пока не будут осознаны, по крайней мере, нижеуказанные трудности и не будут найдены пути их преодоления. Но, прежде чем рассматривать эти трудности и соответствующие проблемы, укажем на основные недочеты, часто встречающиеся в других работах на данную тему.

2. Типичные недостатки материалов по экспертному оцениванию на примере [9]

Мне представляется, что концепция построения рейтинга научных журналов, излагаемая в статье О. В. Федорца, подобна концепции флогистона, широко распространенной в свое время

¹ Сидельников Юрий Валентинович, г.н.с., доктор технических наук, профессор (Sidelnikovy@mail.ru).

в физике. Хотя используемая им концепция сегодня весьма привычна и объективно является одной из лучших в своем роде, она не может претендовать на строгую научность.

Один из основных недостатков исследований, проводимых в рамках этой концепции – использование свертки критериев, чаще всего линейной или мультипликативной (см. стр. 13 [9]). Критика такого подхода широко известна и подробно изложена в работах В. В. Подиновского [3, 4], В. Д. Ногина [2] и других авторов [5].

В обсуждаемом материале, статьях и даже в монографиях известных авторов [1] часто встречается другой недостаток: парное сравнение и экспертная классификация называются методами. Таким образом, вид оценки – «парное сравнение» – и процедура, на основе которой эта оценка осуществляется, имеют одинаковое наименование (см. стр. 14 [9]). Нами предлагается использовать термин: «процедура попарных сравнений», результатом которой и является парное сравнение. Именно так сделано в [7]. Аналогично замечание и по названию другого метода: экспертная классификация (см. стр. 14 [9]). Конечно, необходимо различать экспертный метод и аналогичную процедуру. Для этих целей, возможно, будет полезно формальное определение понятия экспертного метода, впервые предложенное в [8].

Нам представляется, что использовать понятие «коллективная экспертиза» имеет смысл лишь тогда, когда существует именно коллектив экспертов, а не просто некая совокупность невзаимодействующих экспертов (см. стр. 12 [9]). Но в рассматриваемом материале отмечена следующая особенность: «каждый эксперт оценивает свое множество журналов» (см. стр. 12 [9]).

Следующий общий недочет – использование словосочетания «убывание качества» (оно используется и в рассматриваемой работе, см. стр. 13 [9]). Возможно, лучше использовать словосочетание: «убывание уровня качества». Кроме того, в обсуждаемом материале говорить о качестве или его уровне не всегда корректно.

На наш взгляд, вся процедура вычисления обобщенной оценки при замене порядковых номеров числовыми значениями плохо обоснована (см. стр. 23-24 [9]). В следующем разделе статьи мы остановимся на возможных причинах этого явления.

3. Трудности построения системы рейтинга российских научных журналов

Рейтинги теннисистов и шахматистов не вызывают таких нареканий со стороны теоретиков экспертизы, как рейтинги журналов. Перечислим ряд трудностей, которые возникают при решении задачи, заявленной в статье О. В. Федорца. Хотелось бы подчеркнуть, что эти трудности стоят не перед экспертами, а перед экспертологами, т. е. перед теми, кто занимается теорией экспертизы [6].

Сразу отметим, что нас интересуют не объективные характеристики российского научного журнала или их совокупности, как-то: год издания или их вес в граммах, а именно предпочтения экспертов относительно предъявляемых журналов, которые (предпочтения), конечно же, имеют объективную составляющую. Рассмотрим специфику этой задачи. На наш взгляд, трудности, возникающие с обоснованием процедур вычисления обобщенной оценки, связаны, помимо всего прочего, со спецификой объекта оценивания и типом задачи. Укажем на два основных типа задач. Задачи первого типа – это задачи оценивания параметра или характеристики некой объективно существующей реальности. Например, необходимо оценить давление атмосферы на поверхности планеты Венера². И хотя эксперты, которых пригласили для решения этой задачи перед разработкой спускаемого аппарата, существенно ошиблись, в данном случае важно то, что на поверхности планеты существует атмосфера, значение давления в ней конкретно и оно единственно³. В

² Вряд ли этот оцениваемый параметр можно назвать «альтернативой в целом», как это делает автор статьи, рассматривая объекты экспертного оценивания (см. стр. 13 [9]).

³ Конечно, для фиксированного места на Венере и времени оценивания.

качестве примера другого типа задач рассмотрим ранжирование блюд гастрономического пристрастия участников опроса. Возможна ситуация, когда одному из участников дыня нравится больше, чем арбуз, а другому – наоборот, но оба правы! И эту задачу, по-видимому, лучше представить как выявление предпочтений участников опроса относительно предъявляемых продуктов, а не как задачу оценивания продуктов. И для решения такой задачи требуются не эксперты, а просто респонденты.

Рассматриваемый нами случай, когда участники опроса оценивают совокупность известных научных журналов, является неким промежуточным вариантом. В общем случае, нет единственного и объективного значения истины. Опрашиваемые, каждый раз рассматривая множество журналов, используют свое представление относительно их приоритетности. Более того, им не всегда понятно, о какой приоритетности идет речь и к каким критериям им необходимо обращаться. И каждый из опрашиваемых людей является не только экспертом, но и респондентом. Но это разные ролевые функции.

Рассмотрим другие трудности задачи ранжирования. Что общего между формированием прайс-листа бабочек (что исключительно важно для энтомологов) и построением рейтинга научных журналов? На этом пути исследователи встретятся с общими трудностями. Для разных этапов развития некоторых видов насекомых придуманы разные названия. Например: куколка и бабочка. И для тех, кто формирует «прайс-лист» бабочек, не встает вопрос, что есть объект оценивания – они не смешивают понятия «куколка» и «бабочка». Но ведь и журнал, который лишь зарегистрирован соответствующим ведомством и только-только получил лицензию, еще не может являться объектом оценивания при построении соответствующих рейтингов. Лицензия и другие соответствующие документы на журнал – еще не объект оценивания. И даже тот первый номер, который только что вышел из типографии, хотя и является журналом для тех, кто платит за его изготовление или работает над его формированием, не является объектом оценивания. Таким образом, необходимо учитывать стадийность развития журнала и не включать в рассматриваемую совокупность псевдо журналы. Дихо-

томическая модель представления журнала – журнал-«куколка» и журнал-«бабочка» – нас не устроит. Учет наличия стадийности развития здесь явно не достаточен, и необходимо прояснить, в какой мере объект можно включить в список для ранжирования, и это уже другая трудность.

Рассмотрим следующий вопрос: что общего между построением системы рейтингов научных журналов и апориями Зенона? Зерно – еще не куча. Малоизвестный журнал не является объектом ранжировки. Но малоизвестный – для кого? Для его редактора или для его читателей (потребителей)? Этот уточняющий вопрос подталкивает нас к осознанию наличия и необходимости учета позиции такого коллективного субъекта как научное сообщество. Потому что лишь его потребности будут формировать объективные основы для построения рейтинга научных журналов. Лишь высокий уровень потребности научного сообщества или его части, например, англоязычной, в данном журнале (конечно, не в конкретном номере) позволяет эксперту включить его в список для ранжирования.

Одна из основных трудностей, стоящих перед экспертами, заключается в том, что в обсуждаемой задаче принимаемое решение оценивается по нескольким критериям одновременно. Для ее разрешения можно использовать многокритериальный анализ. Но успешное решение таких задач невозможно без учета данных о предпочтениях лиц, принимающих решения. (Одним из важнейших источников таких сведений является информация об относительной важности критериев). По-видимому, в этом случае тот, кто создает ранжированный список альтернатив, одновременно является и ЛПР. То есть опять происходит смешение ролевых функций.

Но кроме смешения ролевых функций здесь существует и другая проблема: какой из имеющихся подходов выбрать? Метод взвешенной суммы, теорию полезности или теорию важности критериев? О некорректности метода взвешенной суммы мы уже писали, сейчас остановимся на этом более подробно. Как справедливо отмечает в своих работах В. В. Подиновский [3, 4], при характеристике метода взвешенной суммы надо обязательно отметить и учесть следующее:

- согласно методу обобщенного критерия, в общем случае, когда у критериев K_i «свои» шкалы, они вначале приводятся к сопоставимому виду («нормализуются»), т. е. заменяются на $K_i^o = h_i(K_i)$. И лишь затем строится обобщенный критерий $F(a_1, K_1^o; \dots; a_m, K_m^o)$. Обоснованный выбор видов функций h_i и вида функции свертки F – очень сложная проблема, которой обычно уделяется мало внимания. Чаще всего попросту берут

$$h_i(K_i) = \frac{K_i - K_{i*}}{K_i^* - K_{i*}} \text{ или } h_i(K_i) = \frac{K_i^* - K_i}{K_i^* - K_{i*}}, \quad F = \sum_{i=1}^m a_i K_i^o,$$

где K_i^* и K_{i*} – соответственно наибольшие и наименьшие значения критериев на множестве вариантов (первая формула для h_i применяется, если критерий желательно максимизировать, вторая – минимизировать);

- с другой стороны, открытой остается проблема обоснованного назначения величин коэффициентов важности (относительных весов⁴) критериев a_i , так как строгого определения самого понятия важности критериев в методе обобщенного критерия нет;
- в том случае, когда все критерии K_i изначально имеют общую шкалу и в роли обобщенного критерия используется взвешенная сумма $F = \sum_{i=1}^m a_i K_i$, фактически принимается ничем не подкрепленное допущение о том, что рост предпочтений вдоль шкалы равномерен, а проблема обоснованного оценивания коэффициентов важности остается открытой.

Можно сконструировать контрпримеры использования метода взвешенной суммы: составить такие множества вариантов, для которых лучший вариант с принятием этого ничем не подкрепленного допущения и без этого допущения будут разными.

⁴ Конечно, понятия относительного веса и коэффициента важности следует различать.

Такие контрпримеры для случая однородных и неоднородных критериев, например, приведены в [4]. Если существуют контрпримеры, то использование метода взвешенной суммы нельзя признать корректным (правомерным).

Какие же корректные многокритериальные подходы можно предложить? Таких подходов существует, как минимум два – использование теории полезности и теории важности критериев. Операция построения функции полезности чрезвычайно трудоемка и сложна для эксперта, так как приходится определять, как изменяются предпочтения ЛПР вдоль шкалы, и в ряде случаев она не имеет смысла. Так, С. А. Салтыковым на основе проведенного вычислительного эксперимента было показано, что в доле от 17,8% до 37,4% случаев ее вообще можно было бы не строить [5]. Конечно, это показано лишь для случая однородных критериев, но они довольно часто встречаются на практике. Для этого случая один лучший (недоминируемый) вариант (с точностью до эквивалентности) можно определить, вообще не задаваясь никакими предположениями относительно того, как изменяются предпочтения ЛПР вдоль шкалы. Попросту говоря, в этом проценте случаев при использовании теории полезности делается лишняя работа. Нам не известны аналогичные исследования для случая неоднородных критериев, но даже вышеприведенные соображения позволяют сделать вывод о том, что из двух многокритериальных подходов (теория полезности и теория важности критериев) следует выбрать именно последний. Таким образом, вместо метода взвешенной суммы мы предлагаем использовать корректную и экономящую время (по сравнению с теорией полезности) теорию важности критериев. Конечно, в обсуждаемой проблеме существуют и другие трудности. Например, отделить реальный уровень журнала от «бренда» этого журнала, который позволяет его узнавать, и наличие психологического фактора – стремления экспертов получить от теоретиков экспертизы некий волшебный индикатор, благодаря которому они сразу и легко смогут определить значимость того или иного научного журнала. Любителям такой волшебной палочки необходимо помнить, что средняя температура по больнице, включая морг, близка к 36 градусам по Цельсию.

4. Заключение

Анализ задачи создания эффективно работающей процедуры строгого ранжирования произвольного множества российских научных журналов показывает: решение этой задачи невозможно до тех пор, пока научное сообщество не осознает, по крайней мере, рассмотренные выше трудности и не найдет пути их решения.

Литература

1. ЛИТВАК Б. Г. *Управленческие решения*. – М.: Ассоциация авторов и издателей «ТАНДЕМ», Издательство ЭКМОС, 1998 г. – 246 с.
2. НОГИН В. Д. *Принятие решений в многокритериальной среде: количественный подход*. – М.: ФИЗМАТЛИТ: 2002. – 176 с.
3. ПОДИНОВСКИЙ В. В., НОГИН В. Д. *Парето – оптимальные решения многокритериальных задач*. – М.: Наука, 1982.
4. ПОДИНОВСКИЙ В. В. *Введение в теорию важности критериев в многокритериальных задачах принятия решений*. – М.: Физматлит, 2007. – 64 с.
5. САЛТЫКОВ С. А. *Экспериментальное сопоставление различных многокритериальных подходов* // Приложение к журналу «Открытое образование», Материалы XXXVI Международной конференции «Информационные технологии в науке, социологии, экономике и бизнесе». – Ялта, Гурзуф, май 2008. – С. 315-317.
6. СИДЕЛЬНИКОВ Ю. В. *Экспертология – новая научная дисциплина* // *АиТ*. – 2000. – №2. – С. 107-126.
7. СИДЕЛЬНИКОВ Ю. В. *Системный анализ технологии экспертного прогнозирования*. – М.: Изд-во МАИ, 2007. – 348 с.
8. СИДЕЛЬНИКОВ Ю. В., САЛТЫКОВ С. А. *Отбор эффективного экспертного метода для задачи III уровня сложности и II уровня обширности* // Приложение к журналу «Открытое образование», Материалы XXXVI Международной

конференции «Информационные технологии в науке, социологии, экономике и бизнесе». – Осенняя сессия, Украина, Крым, Ялта-Гурзуф, 2009. (в печати).

9. ФЕДОРЕЦ О. В. *Коллективная экспертиза научных журналов: методика агрегирования экспертных оценок и построения рейтинга* // *Управление большими системами.* – 2009. – Выпуск 27. – С. 12 – 35.

DIFFICULTIES OF SCIENTIFIC JOURNALS RATING SYSTEM CONSTRUCTION

Yuri Sidelnikov, Institute of Control Sciences RAS, Moscow, chief research officer, Dr.Sc., professor (Sidelnikovy@mail.ru).

Abstract: Difficulties are discussed that arise in formal construction of scientific journals rating system.

Keywords: scientific journals rating system construction, difficulties and problems.