



МЕДИАЦИЯ В НАУКОВЕДЕНИИ: ЭКСПЕРТНО-НАУКОМЕТРИЧЕСКИЙ ПОДХОД

С.А. Салтыков, Е.Ю. Русяева

Предложен медиационный подход в науковедении, объединяющий достоинства экспертного и наукометрического подходов, указаны концептуальные основания этого подхода. Описан алгоритм экспертно-наукометрического оценивания.

Ключевые слова: медиационный, экспертно-наукометрический подход, факторы оценивания, страсти ученых.

ВВЕДЕНИЕ

В современном науковедении существуют, как минимум, два основных подхода к оценке научной деятельности. Один из них, давно известный и широко применяемый — экспертный, когда некие авторитеты в науке оценивают качество научной работы своих коллег. Другой, возникший позже, но активно применяемый в последнее время — формальный, наукометрический. Оба эти подхода обладают своими достоинствами и недостатками, которые заключаются в следующем.

В экспертном подходе в идеале высокое качество оценок обусловлено, по-видимому, тем обстоятельством, что нейросеть эксперта принимает на вход и обрабатывает очень большое число факторов, что является свойством человеческого когнитивного аппарата. Но при этом, в итоге, та же человеческая экзистенция (конкретное бытие, существование) приводит к когнитивным искажениям, поэтому итоговые экспертные оценки все же слишком легко манипулируемы. Эксперты, как многие люди, по своей природе могут быть «алчны и лживы» [1], а поскольку логику их рассуждений эксплицировать (выявлять, делать явной, открытой) никто не заставляет, поэтому есть великое искушение исказить свою экспертную оценку себе во благо.

В наукометрическом, формализованном подходе логика использования факторов эксплицирована, дана в явном виде, в этом его преимущество. Но, главным образом, из-за того, что здесь обычно используется гораздо меньше факторов, чем в экспертном подходе, возникает проблема качества,

адекватности оценивания. Дело в том, что существующие на данный момент перечни формальных показателей не могут в полной мере отразить ни нетривиальности, ни новизны, ни всей многогранности научного исследования. Более того, подчас стремление к более высоким наукометрическим показателям идет вразрез с истинным научным поиском, на который требуется время и серьезные интеллектуальные усилия. «Правила игры», задаваемые современной наукометрией для ученых таковы, что сейчас им выгоднее многократно дублировать свои тексты, чем каждый раз публиковать лишь оригинальные разработки. Именно это и приводит в итоге к манипулированию формальными показателями, что неоднократно критиковалось научной общественностью.

Итак, раз у каждого — экспертного и наукометрического — подходов как у полюсов бинарных оппозиций есть свои плюсы и минусы, возникает вопрос: как использовать достоинства этих подходов и при этом избежать указанных недостатков? Ответ, на наш взгляд, кроется в разработке медиационного¹ подхода, в котором экспертное качество не противоречит формализованному количеству, а, напротив, дополняет достоинства, диалектически возникает из него (по закону перехода количества в качество).

Поясним, каким образом авторы пришли к такому решению. Дело в том, что современные возросшие информационно-технологические возмож-

¹ Медиация — от лат. *mediare* — посредничать. Здесь в значении предложить средний путь, нейтральный и беспристрастный.

ности, в том числе в плане ранжирования данных в вебе [2], дают интересные подсказки и для наукометрии. Так, к примеру, самые известные серверы-поисковики, такие как Яндекс и Гугл, давно пользуются некоторыми приемами, позволяющими им все более адекватно оценивать и ранжировать контент веб-сайтов. В Яндексе успешно применяется формула поискового ранжирования сайтов² [3].

У нас возникла гипотеза: предположим, что аналогичные наработки (типа формулы Яндекса) можно применить для ранжирования научных работ. В этом случае у нас нет необходимости вступать в очередную дискуссию и пытаться доказать, какой именно из указанных выше подходов лучше и важнее. Стоит подойти диалектически: соединить плюсы наукометрического подхода, где имеет место формальная обработка на малом объеме данных, и преимущества экспертного подхода, где много данных, но нет их формальной (т. е. эксплицированной) обработки. Сейчас технические возможности вполне позволяют осуществить подобного рода медиационный подход. Таким образом, современная медиация в науковедении может представлять собой результат диалектического снятия: большой объем данных и их формальная обработка.

Итак, основная цель данного исследования заключается в том, что мы выявляем концептуальные основания нового, медиационного подхода, создаваемого на базе ныне существующих экспертного и наукометрического подходов. Также покажем на примере разработанного алгоритма экспертно-наукометрического оценивания, как можно по-новому, комплексно и многоаспектно ранжировать исследователей, определять их научный рейтинг.

1. МЕДИАЦИОННЫЙ ПОДХОД В НАУКОВЕДЕНИИ: ОСНОВНЫЕ ПОЛОЖЕНИЯ

Метод диалектического снятия, на наш взгляд, вполне применим для поиска нового, нетривиального решения при оценке всех научных построений в каждой из групп наук, как естественных, так и номенклатурных, гуманитарных и математических.

Говоря кратко, медиационность предлагаемого нами подхода заключается в том, что для формирования рейтинга исследователей в расчет принимается много факторов, проводится машинное

² Формула ранжирования Яндекса — это алгоритм, определяющий порядок страниц веб-сайтов в поисковой выдаче.

обучение³ на больших экспертно размеченных выборках⁴, в результате которого определяется формула для расчета рейтинга, например, аналогичная формуле Яндекса для поискового ранжирования сайтов [3].

В медиационном подходе, предлагаемом нами для наукометрии, предлагается использовать весь арсенал факторов, задействовать все показатели. В этом случае не стоит бояться переизбытка информации, напротив, стоит воспользоваться всеми известными подходами при анализе имеющегося текстового научного контента для оценки наиболее содержательного из них. В итоге, чем больше будет предоставлено материала для анализа, тем точнее будут определяться показатели для оценки. Тем самым постепенно можно будет и в автоматическом режиме отсеивать, «отбраковывать» не ценный в системе науковедения, не научный материал.

Итак, резюмируем: снятие антиномий в медиационном подходе в наукометрии происходит благодаря использованию в одном машинном обучении как всех формальных показателей, так и обучающей выборки, полученной путем экспертного ранжирования.

Теперь опишем алгоритм экспертно-наукометрического оценивания, применяемого в медиационном подходе, и укажем некоторые способы сбора материала для его работы.

2. АЛГОРИТМ ЭКСПЕРТНО-НАУКОМЕТРИЧЕСКОГО ОЦЕНИВАНИЯ

Рассмотрим алгоритм, позволяющий определять научный рейтинг (см. рисунок). Его работа

³ Машинное обучение — это методы построения алгоритмов, способных обучаться. Обучение бывает двух типов: 1) обучение по прецедентам, или индуктивное обучение, основанное на выявлении общих закономерностей по частным эмпирическим данным; 2) дедуктивное обучение предполагает формализацию знаний экспертов и их перенос в компьютер в виде базы знаний. Дедуктивное обучение принято относить к области экспертных систем, поэтому термины *машинное обучение* и *обучение по прецедентам* можно считать синонимами.

⁴ Обучающая выборка — это набор размеченных данных, т. е. тех данных, которым эксперт приписал некие признаки. Эти признаки зависят от решаемой задачи. Если решается задача ранжирования или регрессии, то данным приписываются количественные или балльные значения. Если решается задача классификации или мультиклассификации (разбиение на множества, которые могут пересекаться), то данные относятся экспертом к каким-либо классам или мультиклассам. Кроме того, эксперт может указать, что некоторый признак является категориальным, т. е. не может быть в большей или меньшей степени (время года, например, или др.). С точки зрения теории измерений [4] категориальные признаки — это признаки в шкале наименований, а все остальные признаки (некатегориальные) в шкале, не слабее шкалы порядка.



осуществляется в несколько циклов, что позволяет определять итоговые оценки и приоритеты с наибольшей точностью. В перспективе, вероятно, можно будет определять и приоритетность финансирования исследовательских групп, базирываясь на данных медиационного подхода в науковедении.

Опишем последовательность работы алгоритма. На первом этапе в блок 1 поступает исходный набор факторов. Отметим, что мы собираем их из всех ныне имеющихся. Достаточное условие для включения в этот перечень заключается в том, что какой-либо эксперт высказал соображение, что этот фактор может быть полезен для оценки рейтинга ученого. Сюда же включаем, к примеру, весь перечень факторов из РИНЦ и других источников.

Далее, на следующем шаге (в блоке 2), собираем экспертные оценки для обучающей выборки, другими словами, это будет разметкой обучающей выборки. Сюда же возвращается цикл, если нужно добавить оценки новых экспертов. Этот цикл не завершится до тех пор, пока у нас не будет стабильного топ-10 ученых — только данный топ будет показателем необходимости выхода из первого цикла.

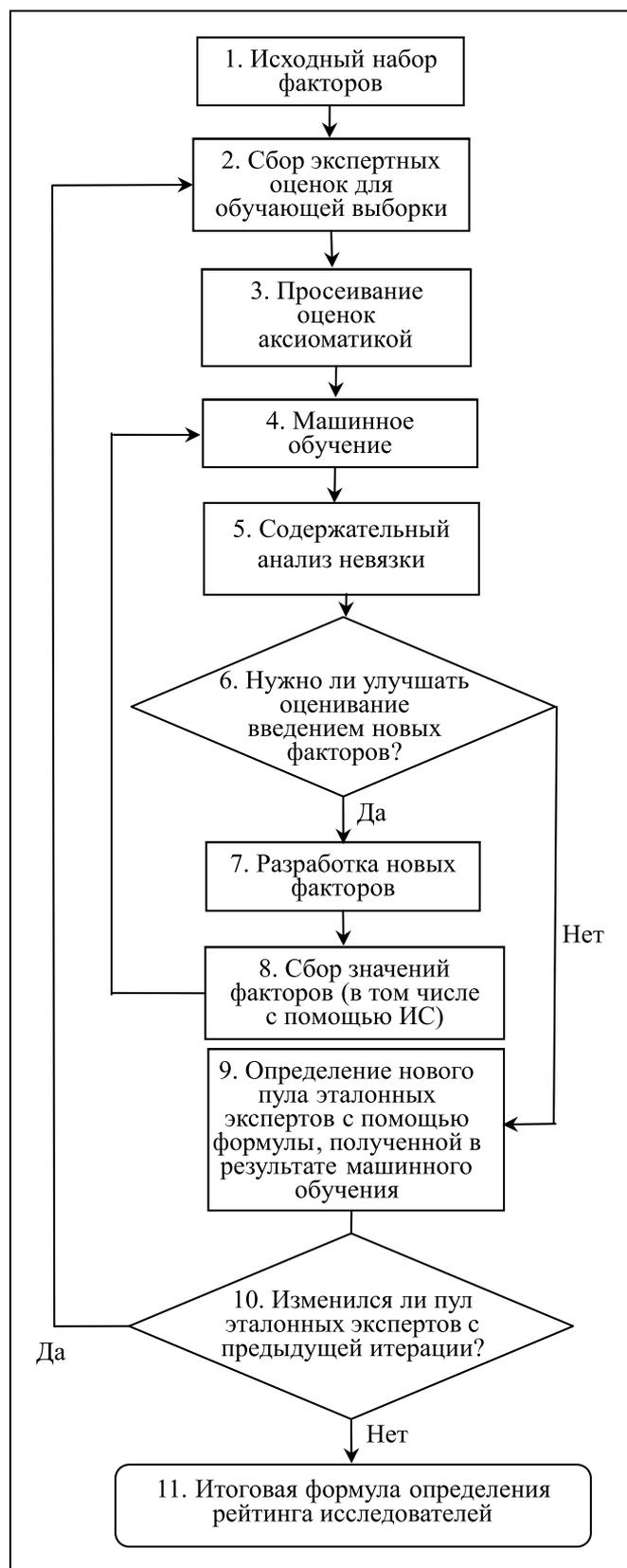
Затем (блок 3) просеиваем имеющиеся оценки аксиоматикой. Этот процесс сводится к тому, что отсеиваются или отбраковываются оценки, которые не соответствуют какому-либо консенсусному суждению относительно соотношения страты исследователей из различных наук. Поясним далее, в § 3, какое именно допущение нами будет принято в этом аспекте.

В блоке 4 осуществляется машинное обучение: автоматическая настройка параметров формулы ранжирования, содержащей вектор уже имеющихся факторов. Сюда же, на этот этап, возвращается цикл, если у нас обновляется набор значений факторов, собранных, в том числе и с помощью информационной системы (ИС).

На следующем шаге (блок 5) нам необходимо провести содержательный анализ невязки. Здесь мы анализируем наибольшие отклонения, которые имеются между экспертными оценками и результатом, который дает машинно обученная формула.

Далее мы попадаем в блок 6, в котором задаем вопрос: «Нужно или можно ли улучшить оценивание введением новых факторов?» Если ответ «Да», то мы переходим на следующий этап, блок 7, когда необходимо разрабатывать новые факторы. После этого, в блоке 8 нам нужно собрать значения новых факторов, возможно, и с привлечением ИС, и после этого выходим на новый цикл машинного обучения (блок 4) и проходим вновь по всем этапам, что следуют по ходу этого цикла.

Если же ответ на вопрос блока 6 «Нет», то мы переходим к этапу определения нового пула эта-



Структурная схема алгоритма экспертно-наукометрического формирования рейтинга исследователей

лонных экспертов с помощью формулы, полученной в результате машинного обучения — блок 9. По завершении этого этапа мы вновь попадаем в еще одну развилку с вопросом (блок 10): «Изменился ли пул эталонных экспертов с предыдущей итерации?» Если ответ «Да», то мы проводим очередную итерацию, возвращаясь в блок 2, где заново собираем экспертные оценки вновь выявленных экспертов и, теперь уже новых экспертов, просим дать оценки для обучающей выборки, проводя весь цикл заново. Напомним, что только стабильный топ-10 ученых служит показателем необходимости выхода из цикла. Если ответ на вопрос блока 10 «Нет», то мы получили итоговую формулу (блок 11) определения рейтинга исследователей.

В целом же экспертно-наукометрический подход, скорее всего, потребует создания целого каскада алгоритмов для того, чтобы ввести, учесть и согласовать значения всех вновь выявляемых факторов. Эту работу необходимо провести для того, чтобы сделать итоговую формулу определения рейтинга ученых наиболее адекватной.

3. О СПОСОБАХ СБОРА МАТЕРИАЛА ДЛЯ МЕДИАЦИОННОГО ПОДХОДА В НАУКОВЕДЕНИИ

В данном параграфе мы лишь кратко затронем некоторые способы сбора материала для разрабатываемого нами медиационного подхода. Тема эта сама по себе обширна и нестандартна, требует отдельного детального исследования.

Отбор факторов машинного обучения может осуществляться с двумя целями. Одна из них — вычислительная: отбор факторов предназначен для повышения точности и снижения трудоемкости расчетов. Другая цель — содержательная интерпретация факторов машинного обучения. Эти цели во многом антагонистичны. Так, для вычислительных целей имеет смысл использовать только один фактор из набора скоррелированных факторов и, следовательно, одному из этих факторов стоит присвоить существенно меньший ранг. А для их содержательной интерпретации напротив, нужно, чтобы все скоррелированные факторы имели одинаковые или близкие ранги.

Мы выдвигаем следующую гипотезу: существуют достаточно много методов решения вычислительной задачи и совсем немного способов, позволяющих решать задачи содержательной интерпретации. Поэтому методы содержательной интерпретации факторов машинного обучения еще нужно дорабатывать.

Поясним, с учетом работы описанного в § 2 алгоритма, каким именно образом предполагается

наполнять перечень факторов, на которых обучается формула.

Согласно общенаучной традиции [5], науки делятся на науки сильной и слабой версий, что, согласно нашей нотации, будет соответствовать ранжированию ученых по двум основным стратам:

— специалистов в области естественных и математических наук отнесем к первой страте;

— исследователей в области гуманитарных и номенклатурных наук — ко второй.

На практике мы предлагаем определять принадлежность ученых к одной из страт по формальным показателям, в частности, по номеру его специальности согласно номенклатуре специальностей научных работников ВАК.

Постулируем условие: ученый первой страты не может быть оценен ниже ученого второй страты, т. е. оценка выбраковывается, если представитель гуманитарных наук, например, историк, поставлен выше представителя естественных наук, к примеру, ученого-химика. Конечно, отдельно взятый гуманитарий может быть «сильнее» некоторого естествознателя. Более того, такое даже регулярно встречается (взять хотя бы Канта или Гегеля), но в среднем все же представитель наук сильной версии оценивается выше, чем представитель наук слабой версии (это не говорит, что кто-то хуже или лучше, это показатель кумулятивности дискурса). Для целей нашего исследования мы отбрасываем противоположные оценки, чтобы и без того «шумную» обучающую выборку сделать менее «шумной». Это один из показателей отбраковки суждений экспертов-ученых.

Для экспертного ранжирования ученых необходимо составить анкету, отвечая на вопросы которой ученые могут оценивать своих коллег. Данные анкет, значения которых не соответствуют указанному условию, отбраковываются.

Таким образом мы сформируем перечень целевых значений параметров для каждого ученого. После этого остальные параметры ученого загружаются из соответствующих наукометрических баз, например, из РИНЦ. Затем на сформированной выборке проводится машинное обучение.

ЗАКЛЮЧЕНИЕ

Итак, благодаря разработке медиационного экспертно-наукометрического подхода в науковедении можно будет, на наш взгляд, проводить более качественное оценивание ученых, тем самым составлять адекватный рейтинг самих ученых, что сейчас делается очень неточно. Более того, с помощью описанного алгоритма и дальнейших разработок в русле заявленного нами подхода можно



определять и приоритетность финансирования исследовательских групп, которая будет базироваться на этом адекватно составленном рейтинге ученых.

В предлагаемом нами медиационном подходе для науковедения в целом снятие имеющихся противоречий или, говоря другими словами, исправление недостатков экспертного и наукометрического подходов, взятых в отдельности, происходит благодаря учету в одном машинном обучении как всех формальных показателей, так и обучающей выборки, полученной путем экспертного ранжирования [6].

Понятно, что в рамках одного исследования мы не сможем сразу ответить на массу возникающих вопросов и проработать все детали. Поэтому в дальнейшем планируем продолжить проработку всех возможных способов сбора и классификации материала для функционирования алгоритмов идентификации. Также продолжим анализ и ранжирование научных построений ученых с учетом разных стадий жизненного цикла исследований. Только весь этот комплекс разработок, проведенных в рамках медиационного подхода, может служить в дальнейшем оптимальным инструментом определения приоритетности научных работ и рейтинга ученых.

ЛИТЕРАТУРА

1. Бурков В.Н. Основы математической теории активных систем. — М.: Наука, 1977. — 255 с.
2. Салтыков С.А., Русяева Е.Ю. Author's Typology of SEO-Strategies as a Tool of Conceptual Creativity // Proc. of First Conf. «Creativity in Intelligent Technologies and Data Science» CIT&DS 2015, Volgograd. — Volgograd: Springer International Publishing AG, 2015. — P. 58—80.
3. Borisov A., Serdyukov P., de Rijke M. Using Metafeatures to Increase the Effectiveness of Latent Semantic Models in Web Search // Proc. of the 25th Intern. Conf. on World Wide Web. — Montréal, Québec, Canada, April 11—15, 2016. — P. 1081—1091. — URL: <https://staff.fnwi.uva.nl/m.derijke/wp-content/papercite-data/pdf/borisov-using-2016.pdf> (дата обращения: 25.10.2017).
4. Пфанцагль И. Теория измерений / при участии В. Баумана и Г. Хубераи; пер. с англ. В.Б. Кузьмина. — М.: Мир, 1976. — 244 с.
5. Новиков А.М., Новиков Д.А. Методология. — М.: URSS, 2014. — 627 с.
6. Салтыков С.А., Русяева Е.Ю. Рафинирование научных построений в теориях принятия решений. — М.: ИПУ РАН, 2016. — 208 с.

Статья представлена к публикации членом редколлегии В.Е. Лепским.

Салтыков Сергей Анатольевич — канд. техн. наук, ст. науч. сотрудник, ✉ sergey.saltykov@gmail.com,

Русяева Елена Юрьевна — канд. филос. наук, ст. науч. сотрудник, ✉ gusyaeva@ipu.ru,

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва.

Новая книга

Бреер В.В., Новиков Д.А., Рогаткин А.Д. Управление толпой: математические модели порогового коллективного поведения. — М.: Ленанд, 2016. — 168 с.

Книга посвящена математическим моделям управления толпой, поведение которой описывается как пороговое (конформное) коллективное принятие решений ее элементами.

На основании результатов анализа взаимосвязи между микро- и макромоделями активных сетевых структур рассматриваются статические (детерминированные, стохастические и теоретико-игровые) и динамические (в дискретном и непрерывном времени) модели управления толпой. Значительное внимание уделяется моделям информационного противоборства. Многие результаты применимы не только для задач управления толпой, но и для управления в социальных группах, онлайн-новых социальных сетях и др.

Для научных и практических работников, а также студентов, аспирантов и докторантов, специализирующихся в области моделирования коллективного поведения.

Скачать PDF: <http://www.mtas.ru/biblio/MMUT.pdf>

Интервью Д.А. Новикова, посвященное выходу в свет новой книги, читайте на <http://www.ipu.ru/press-center/42978>

Уважаемые читатели!

Если Вы не успели подписаться на журнал «Проблемы управления», то через нашего издателя можно оформить льготную подписку в любое время и с любого номера (дешевле, чем через каталоги агентств) или приобрести номера журнала за прошлые годы.

Можно также заказать электронные версии как необходимого Вам номера журнала, так и отдельных статей.

Позвоните издателю по тел. (495) 330-42-66 или пришлите заказ по электронной почте datsys@mail.ru — и подписка будет оформлена за один день. Расходы по пересылке журнала издатель берет на себя. Не забудьте указать свой полный почтовый адрес!