

MOTROSHILOVA N. RUSSIAN SCIENTIFIC CITATION INDEX IN REFERENCE TO PHILOSOPHY

The article deals with the analysis of the existing forms and methods of Russian Scientific Citation Index service (eLibrary) from the point of view of representation the real effectiveness of science and research activities.

Key words: scientometrics, citation databases, bibliometric systems, Russian scientific citation index.

Б.И. БЕДНЫЙ, доктор ф.-м. наук
Ю.М. СОРОКИН, доктор ф.-м. наук
ННГУ им. Н.И. Лобачевского
(национальный исследовательский университет)

О показателях научного цитирования и их применении

Обсуждаются показатели научного цитирования, применяемые для количественной оценки результативности научной деятельности. Понятие «импакт-фактор» рассматривается как группа показателей, различающихся алгоритмом вычисления. Проанализированы свойства индекса Хирша. Показано, что значения показателей, связанных с измерением количества публикаций и ссылок (цитирований), существенно зависят от выбора библиографической рейтинговой базы, на которой они рассчитываются. Оцениваются возможности и ограничения для применения наукометрических показателей в вузовской практике*.

Ключевые слова: научный журнал, цитируемость, рейтинговая база, классический импакт-фактор, импакт-фактор предметной области, относительный импакт-фактор, индекс Хирша.

В последние годы в связи с расширением конкурсного финансирования научных исследований и разработок научные работники и преподаватели высшей школы все чаще обращаются к наукометрическим показателям (индикаторам), применяемым для количественной оценки результативности научной деятельности организаций, научных коллективов и отдельных исследователей. При подготовке заявок на выполнение научных проектов и отчетных документов по выполненным проектам обычно необходимо указывать количество научных статей, отражаемых в определённых международных и российских базах данных, импакт-факторы журналов, в которых эти статьи опубликованы, показатели цитирования и индексы Хирша руководителей и основных исполнителей проектов. Однако, судя по тому, какие показате-

ли запрашиваются организаторами конкурсов, а также по личному опыту авторов этой статьи, многократно обсуждавших вопросы практического применения наукометрических показателей с университетскими коллегами и сотрудниками академических институтов, далеко не все, кто оперирует этими показателями, имеют ясное представление об их смысле, методике измерения и значимости при решении тех или иных задач в области организации научных исследований.

Известно, что, несмотря на колоссальный прогресс в развитии науки, она до сих пор не смогла найти оптимальный способ измерения значимости собственных результатов, качества научной продукции ученых, их вклада в новое научное знание. Одной из шкал для оценки научной продуктивности является число публикаций в

* Авторы выражают признательность М.Б. Сапунову, иницировавшему постановку вопросов, рассмотренных в данной статье.

авторитетных научных изданиях и журналах, осуществляющих содержательную экспертизу поступающих в редакцию статей (эта шкала традиционно используется руководством вузов при проведении конкурсных процедур по замещению вакантных должностей профессорско-преподавательского персонала). Вместе с тем, как писал один из основоположников науковедения Д. Прайс, «с самого начала и с величайшей готовностью согласимся, что это плохая шкала. Кто бы осмелился уравнивать одну статью Эйнштейна по теории относительности хотя бы сотней статей бакалавра Джона Доу о константах эластичности для различных древесных пород в лесах Нижнего Базутоленда, по одной константе на статью. Кроме того, шкала плоха уже потому, что само ее существование толкает людей на публикации просто ради престижа. Но при всем том шкала дает все же исходный пункт для анализа ... поскольку существует довольно строгое соотношение между элитностью ученого и его продуктивностью как автора статей» [1, с. 317].

По мере глобализации науки, стремительного роста объема научной периодики и сокращения времени обмена информацией значимость метода подсчета публикаций в научных журналах падает. Во-первых, требует уточнения само понятие «научный журнал». Попытка такого рода уточнения (фильтрации) научных изданий проведена, например, ВАК в виде условий, необходимых для их включения в Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертаций (так называемый Перечень ВАК) [2]. Во-вторых, научный уровень рецензирования поступающих в редакцию материалов не может быть сопоставимым для разных журналов. Чем выше уровень журнала, тем более качественную экспертизу статей он проводит. В результате для любой области исследований в научном сообществе, в раз-

ных его группах давно и устойчиво существует некий интуитивный рейтинг профильных журналов, не имеющий, однако, количественного выражения. Этот рейтинг влияет на практику оформления научных результатов и их использования, но за пределами самой науки (научного направления, дисциплины, специальности) не информативен.

Идея придания рейтингу научных журналов универсального количественного выражения принадлежит Ю. Гарфилду [3, 4]. В 1963 г. эта идея была реализована в экспериментальном порядке, а с 1964 г. возглавляемый им Институт научной информации в Филадельфии стал регулярно выпускать так называемый «Индекс научных ссылок» (Science Citation Index, SCI). В этом ресурсе систематизирована информация о цитировании научных статей по библиографической базе данных, которая первоначально охватывала около 600 научных журналов. Выросший отсюда современный библиографический ресурс Web of Knowledge компании Thomson Reuters, включающий, в частности, базу Web of Science (WoS) с библиометрическим инструментом Journal Citation Reports (JCR), учитывает («расписывает») более 12000 наиболее авторитетных в мире научных журналов. Сегодня по широте охвата научной периодики с этим ресурсом конкурирует лишь созданный гораздо позднее европейской компанией Elsevier ресурс SCOPUS (SCP).

Идея Гарфилда заключалась в оценке значимости научного материала (статьи, журнала) по числу его упоминаний (ссылок, цитирований) в научной периодике. Отметим, что с позиций наукометрии, рассматривающей науку как информационный процесс, цитируемость действительно является мерой полезности публикации [5]. Как известно (см., например, [6]), упоминание в ссылках работ предшественников и коллег является нормой, обязательной для научной публикации. Поэтому подсчет количества ссылок (метод «цитат-индекс»

[7]) позволяет оценить влияние публикации или группы публикаций на развитие науки как информационного процесса, выявить и оценить воздействие результата на научное сообщество, его полезность для других ученых. Конечно, высокая цитируемость – это еще не гарантия какого-то особого качества работы, это лишь сигнал, свидетельствующий о том, что данная работа с большой вероятностью может оказаться полезной для развития научного знания¹.

Подсчет цитирования публикаций не призван заменить их оценку. Он, скорее, расширяет возможности таких оценок, делая их более объективными [9]. В определенной степени «цитат-индекс» – это завуалированный метод экспертных оценок, и, естественно, он включает элемент субъективности, которая, однако, «... гасится на больших массивах публикаций в результате усреднения экспертиз разных авторов» [7]. Поэтому для проведения формальных оценок результативности научной деятельности и авторитетности ученых важно измерять не столько цитируемость отдельных работ, сколько цитируемость значительных массивов публикаций (авторов, научных коллективов, научных журналов). Вместе с тем отметим, что в социологии науки надежно доказана высокая степень корреляции между цитируемостью публикаций и общественным признанием вкладов их авторов в развитие науки [10, 11].

Различные методики подсчета цитирований научных публикаций получили в настоящее время широкое признание не потому, что полученные с их помощью оценки являются однозначными или исчерпывающими, а потому, что они являются пока

единственными формализованными оценками научных результатов средствами самой науки.

Система показателей, основанных на подсчете числа ссылок и цитат, библиометрическая по методам, наукометрическая по целям, продолжает развиваться (см., например, [12–16]). Не претендуя на полноту изложения, далее мы остановимся на анализе лишь двух наиболее распространенных сегодня показателей – «импакт-фактор» и «индекс Хирша». Эти показатели алгоритмически просты, имеют ясный и наглядный смысл и наиболее часто используются при проведении оценок результативности научной деятельности.

Импакт-факторы и рейтинговые базы

Среднюю цитируемость опубликованных в журнале статей (количество ссылок в расчете на одну статью) принято называть импакт-фактором этого журнала. Значение импакт-фактора характеризует важность, авторитетность научного журнала. Применяются разные алгоритмы подсчета импакт-фактора, отличающиеся процедурой усреднения. Поэтому правильнее говорить об импакт-факторе как о группе понятий, определяемых параметрами усреднения. Дадим определение импакт-фактора, обобщающее формулировку, предложенную в [17, 18].

Пусть S – библиографическая рейтинговая база, используемая при подсчетах, $PUB_S(t)$ – число статей, опубликованных в данном журнале в году t и включенных («расписанных») в S , $CIT_S(T, t)$ – число ссылок (цитирований), сделанных в году T

¹ Использование для этой цели других сигналов, генерируемых, например, инструментариумом семантических сетей [8], может только приветствоваться, если соответствующая идея обретёт черты объективного алгоритма, исключающего корпоративные интерпретации и административную составляющую. При этом следует иметь в виду, что указанный алгоритм, неизбежно использующий аппарат теории графов, будет явно уступать методу «цитат-индекс» в смысле прозрачности.

² Не любая публикация учитывается и включается в базу данных; это зависит от установленных при ее формировании правил.

на все статьи из группы $PUB(t)$ во всех расписанных в S журналах. Тогда импакт-фактор данного журнала:

$$IF_n(T, S) = \frac{\sum_{\tau=\tau_0}^n CIT_S(T, T-\tau)}{\sum_{\tau=\tau_0}^n PUB_S(T-\tau)}. \quad (1)$$

Здесь τ – время запаздывания ссылок, τ_0 – параметр запаздывания (минимальное время запаздывания ссылок, принятое при подсчете конкретного импакт-фактора), n – порядок импакт-фактора (максимальное время запаздывания, выбранное при подсчете импакт-фактора; этот параметр характеризует ширину интервала усреднения – «окно цитирования»). По смыслу введенных параметров T , τ , n – целые числа, причем $\tau \geq \tau_0$, а $n \geq \tau$.

Согласно (1) импакт-фактор порядка n за год T может быть определен на библиографической рейтинговой базе данных S при условии, что эта база содержит данный журнал не только за год T , но и, по крайней мере, за n лет, предшествующих году T .

Как видно из (1):

- для одного и того же журнала можно определить множество импакт-факторов (однако лишь некоторые из них получили широкое распространение);
- любой конкретный импакт-фактор является функцией времени (года цитирования T);
- импакт-фактор может быть вычислен не только для журнала, но и для иного объекта (группы журналов, страны, корпорации), если только количество публикаций для этого объекта в «окне цитирования» не является нулевым.

«Классический» импакт-фактор (по Гарфилду [4]) определяется из (1), если $\tau_0 = 1$, $n = 2$. Смысл его очень прост: число ссылок, полученных в году T из всех расписанных в базе журналов, на статьи, опубликованные в обследуемом журнале в те-

чение двух предыдущих лет, делится на число этих самых статей. Таким образом, окно цитирования принимается равным двум годам, а параметр запаздывания ссылок – одному году. Под термином «импакт-фактор», если не сделано каких-либо уточнений, принято подразумевать именно этот показатель.

В связи с определением термина «импакт-фактор» отметим некоторые важные обстоятельства, не слишком известные широкой аудитории, оперирующей этим понятием.

1. Определение импакт-фактора теряет смысл без указания на библиографическую базу S . Значение импакт-фактора журнала существенно зависит от выбора библиографической рейтинговой базы, на которой он рассчитывается. Поэтому при сравнении импакт-факторов разных журналов строится количественный рейтинг на априорно выбранной базе S .

2. Важнейшими признаками авторитетности базы являются представительность, полнота и качество расписанного в ней контента.

3. Представительная база должна быть международной и мультидисциплинарной, иначе само понятие импакт-фактора теряет отношение к науке. Сужение границ представительности по национальному или тематическому признаку приводит к искажённым (в смысле пригодности для сравнения) результатам.

4. Условие полноты означает наличие в базе достаточно глубокого архива журналов и сохранение данных о каждой публикации по всем библиографическим признакам.

5. Условие высокого качества контента означает, что рейтинговая база должна обязательно содержать «входной фильтр», обеспечивающий отбор расписываемых журналов как по формальным, так и по содержательным признакам. При нарушении регулярности выхода номеров журналов и сроков их поступления в базу теряется со-

поставимость показателей (1) во времени. При нарушении требований к оформлению материалов происходит потеря или искажение данных при алгоритмической обработке. При отсутствии содержательного фильтра база грозит быстрое «зашумление».

Одним из критериев отбора журналов, принятых в базах WoS и SCP, является превышение импакт-фактором журнала-кандидата некоего минимального (порогового) значения [19–22].

В качестве иллюстрации представительности отметим, что базы WoS и SCP в настоящее время расписывают приблизительно одинаковую по значимости совокупность: более 12 тыс. журналов из примерно 100 тыс. журналов, выходящих в мире (фильтр SCP более мягкий, общее количество журналов здесь больше, отмечается существенный крен в сторону европейской периодики). При этом в обеих базах отсутствуют как национальные, так и предметные «провалы». Есть все основания считать эти базы лидерами по представительности³. Однако, если ставить узкую цель поиска по автору, теме или публикации, ни один из лидеров не может считаться доминирующим по представительности [23]. С точки зрения представительности заметным недостатком всех упоминаемых рейтинговых баз является почти полное отсутствие в них монографий (в базах WoS и SCP процесс их ввода начался сравнительно недавно), что особенно существенно в гуманитарной области, где новые научные результаты нередко отражаются именно в этом классе публикаций.

Конечно, условие представительности находится в определённом противоречии с

качеством контента базы. Баланс этих двух требований главным образом и определяет неформальный рейтинг базы. Более мягкий по сравнению с WoS, фильтр SCP (об этом можно судить хотя бы по тому, что число российских журналов, расписанных в SCP, примерно вдвое больше, чем в WoS) отражает конкурентную политику более молодой европейской базы.

По критерию полноты база SCP существенно уступает WoS из-за разной глубины архивов: в SCP глубина архива – 1995 г., WoS – 1979 г. Фактический архив этих баз глубже декларируемого, однако архив предыдущих лет не удовлетворяет всем требованиям полноты. Там либо не представлены все включённые в базу источники (SCP), либо из всех авторов публикации отражён лишь первый (WoS).

Отметим еще одно важное для корректного расчета импакт-фактора обстоятельство. В базе, реализующей алгоритм (1), должна быть определена дата обновления данных. Поскольку для подсчёта $IF_n(T, S)$ надо иметь библиографические данные на все выпуски включённых в базу журналов за год обследования, требуется высокая оперативность всех элементов цепочки: «выход журналов – внесение данных в базу – подсчёт IF_n ». Вероятно, любому читателю этой статьи известны российские научные журналы, заведомо не удовлетворяющие критерию оперативности выхода всех номеров, принятому в базах JCR и SCP. И это одна из многих проблем национальной, пока более скромной во всех отношениях базы «Российский индекс научного цитирования» (РИНЦ)⁴. Глубину архивов и полноту отражения для РИНЦ в целом оце-

³ Следует упомянуть о сопоставимой представительности Google Scholar, однако этот ресурс не имеет ни ссылочного научного инструментария, ни собственной чётко очерченной базы источников, ни каких бы то ни было входных фильтров; он работает как мощный гипертекстовый поисковик и для наукометрического анализа непригоден.

⁴ «Особенностью национальной научной периодики» (по крайней мере, значительной её части), отражаемой в РИНЦ, является то, что используемые при обработке данных электронные версии многих российских журналов запаздывают даже по отношению к печатным, тогда как в остальном мире имеет место обратное.

нить затруднительно: в настоящее время она очень неровная, хотя по каждому источнику в отдельности соответствующую информацию получить можно. В качестве иллюстрации неполноты отметим, что значительная часть отражённых в РИНЦ публикаций представляет собой ссылочные источники (то есть источники, отражённые на основании ссылочных списков расписанных в базе статей). В связи с отсутствием процедуры содержательной фильтрации [24] база РИНЦ сегодня учитывает около 7 тыс. только российских журналов.

Если говорить о представительности лидирующих международных рейтинговых баз по отношению к российской науке, то, по данным [22], в 2011 г. в базе SCP учитывались 237, а в базе WoS – 117 текущих («активных») российских научных журналов. При этом импакт-факторы этих журналов постепенно возрастают. Если в 1995 г. количество российских журналов с $IF_2(1995, JCR) > 0,25$ и $IF_2(1995, JCR) < 0,1$ было примерно одинаковым, то в 2004 г. число российских журналов с $IF_2(2004, JCR) > 0,25$ приблизительно в 6 раз превосходило число журналов с $IF_2(2004, JCR) < 0,1$ [25]. Нами

отмечены 107 российских журналов в Указателе цитирования по точным и естественным наукам (подбаза JCR Science Edition) с $IF_2(2007, JCR)$ от 0,02 до 2,03 и 6 российских журналов в подбазе JCR Social Science Edition с $IF_2(2007, JCR)$ от 0,02 до 1,58.

С учётом того, что значительная часть работ российских учёных, специализирующихся в области точных и естественных наук, публикуется в зарубежных журналах, имеющих достаточно высокие импакт-факторы, представительность этих областей российской науки в базах-лидерах можно признать вполне адекватной. К сожалению, этого нельзя сказать о социогуманитарных дисциплинах, где в качестве библиографической рейтинговой базы целесообразно выбирать РИНЦ.

Сказанное выше о представительности, глубине архивов и контенте рейтинговых баз вовсе не означает отрицание ценности тех из них, которые имеют ограниченную представительность, недостаточно глубокий архив или слишком мягкий входной фильтр. Поисковый аппарат базы, доступ к контенту – это непосредственное информационное обеспечение научных исследо-

Таблица 1

Некоторые российские журналы, импакт-факторы которых в РИНЦ ниже, чем в JCR

Наименование журнала	$IF_2(2007, JCR)$	$IF_2(2007, РИНЦ)$
Успехи физических наук	2,032	0,514
Успехи химии	1,893	1,049
Журнал экспериментальной и теоретической физики	1,075	0,473
Теоретическая и математическая физика	1,012	0,045
Квантовая электроника	0,985	0,296
Успехи математических наук	0,309	0,103

Таблица 2

Некоторые российские журналы, импакт-факторы которых в РИНЦ выше, чем в JCR

Наименование журнала	$IF_2(2007, JCR)$	$IF_2(2007, РИНЦ)$
Вопросы психологии	0,112	0,831
Журнал органической химии	0,477	0,575
Высшее образование в России	–	0,542
Дефектоскопия	0,156	0,383
Социологические исследования	0,194	0,299
Университетское управление: практика и анализ	–	0,278
Известия РАН. Теория и системы управления	0,145	0,254

ваний. Речь идёт только о несопоставимости наукометрических показателей, вычисленных по единому алгоритму на различных базах. В качестве иллюстрации в *табл. 1* приведены некоторые российские журналы, импакт-факторы которых в базе JCR существенно выше, а в *табл. 2* – существенно ниже, чем в базе РИНЦ.

**«Неклассические»
импакт-факторы**

Классический (двухлетний) импакт-фактор является наиболее известным алгоритмом оценки рейтинга журнала. Однако широко используются также иные показатели этой группы, отличающиеся выбором параметров усреднения. Рассмотрим особенности этих «неклассических» библиометрических показателей.

1. *Пятилетний импакт-фактор.* Согласно определению (1), пятилетний импакт-фактор – это показатель $IF_5(T, S)$, в котором окно цитирования расширено до пяти лет, а параметр запаздывания ссылок τ_0 остаётся равным единице. Ранжирование по пятилетнему импакт-фактору снижает рейтинг журналов, получающих основное число ссылок на статьи менее чем трёхлетней давности, и повышает рейтинг журналов, на статьи в которых следует более медленная реакция научного сообщества. Так, например, соотношение $IF_5(T, JCR) > IF_2(T, JCR)$ характерно для подавляющего числа мировых журналов по менеджменту [12]. Следует ожидать такого же результата и для других социальных и гуманитарных наук.

Для журналов с достаточно высоким рейтингом различия пятилетнего и класси-

ческого импакт-факторов, как правило, невелики, причём отклонения могут быть в обе стороны. Заметим, что одной из причин, повышающей пятилетний импакт-фактор некоторых журналов по отношению к классическому, является большое время прохождения рукописей (причина не столько научного, сколько организационного свойства, характерная для многих российских журналов). Как видно из *табл. 3*, для обзорных журналов, получающих большое число ссылок в отечественной периодике, характерно превышение пятилетнего импакт-фактора по отношению к классическому, тогда как один из наиболее оперативных российских журналов («Письма в Журнал экспериментальной и теоретической физики») демонстрирует заметное превышение классического импакт-фактора над пятилетним. Таким образом, соотношение этих двух формальных показателей зависит и от качества публикаций в том или ином журнале, и от организации процесса прохождения рукописей в ссылочной периодике.

Отметим, что преимуществом пятилетнего импакт-фактора является сглаживание выбросов от отдельных статей с аномальной цитируемостью (наибольшее известное значение классического импакт-фактора $IF_2 = 90$ явилось следствием публикации статьи, получившей тысячи ссылок в течение года). Недостатком же пятилетнего импакт-фактора можно считать слишком большое запаздывание (не менее 6 лет) при отражении в рейтинге новых журналов, какое бы признание за это время они ни завоевали.

2. *Оперативный импакт-фактор.* В

Таблица 3

Классические и пятилетние импакт-факторы некоторых российских журналов

Наименование	$IF_2(2010, JCR)$	$IF_5(2010, JCR)$
Успехи физических наук	2,245	2,531
Успехи химии	2,346	2,827
Письма в Журнал экспериментальной и теоретической физики	1,557	1,329
Журнал аналитической химии	0,650	0,647
Социологические исследования	0,147	0,135

последнее время быстро увеличивается число цитирований статей, вышедших в том же году. Это связано с ускорением производственного цикла многих научных издательств, а также с публикацией некоторыми издательствами предварительных версий принятых в печать статей. Показатель, отражающий столь оперативную реакцию научного сообщества с точки зрения общего определения (1), есть $IF_0(T, S)$, где параметр запаздывания ссылок τ_0 и порядок импакт-фактора n приняты равными нулю ($\tau_0 = n = 0$). По общему смыслу представляется логичным называть этот показатель оперативным импакт-фактором (альтернативные названия: immediacy index – в базе JCR, индекс оперативности [12]). Естественно, что значения оперативного импакт-фактора обычно намного меньше, чем классического (табл. 4).

3. *Медианный импакт-фактор*. Ретроспектива цитирования в году T статей из определённого журнала, опубликованных

рения информации, опубликованной в данном журнале.

Ясно, что по общему алгоритму (1) на базе S можно вычислить импакт-фактор, для которого при параметре запаздывания $\tau_0 = 0$ ширина окна цитирования будет равна медиане цитирования ($n = \tau_m$):

$$IF_m(T, S) = \frac{\sum_{\tau=0}^{\tau_m} CIT_S(T, T - \tau)}{\sum_{\tau=0}^{\tau_m} PUB_S(T - \tau)} \quad (2)$$

Этот импакт-фактор по предложению Р. Руссо [16] и принято называть медианным. Медианный импакт-фактор имеет то преимущество, что он «подстраивает» окно цитирования под время старения информации в конкретной предметной области, к которой относится журнал (а эти времена для разных предметных областей сильно различаются), – в отличие от рассмотренных выше импакт-факторов с априори заданными окнами цитирования. Это позво-

Таблица 4

Оперативные и классические импакт-факторы некоторых российских журналов

Наименование	$IF_0(2010, JCR)$	$IF_2(2010, JCR)$
Успехи физических наук	0,651	2,245
Успехи химии	0,057	2,346
Письма в Журнал экспериментальной и теоретической физики	0,232	1,557
Письма в Астрономический журнал: астрономия и космическая астрофизика	0,286	1,091
Океанология	0,324	0,581

в предыдущие годы, отражается специфическим для данного журнала распределением $CIT_S(T, T - \tau)$, форма которого может быть самой разной. Если статьи, опубликованные в данном журнале за период от года T до некоторого года $T - \tau_m$, получают половину общего числа ссылок на данный журнал в году T , то значение τ_m называется медианой цитирования (*median cited age* [16]). Значение τ_m соответствует году, разделяющему все ссылки на данный журнал в году T на две равные части. Как видно из определения, τ_m характеризует время ста-

ляет считать медианный импакт-фактор предпочтительным для наукометрических оценок в области гуманитарных и общественных наук.

Цитируемость. Относительный импакт-фактор

Модный термин «импакт-фактор» зачастую применяют и к авторам, и к научным учреждениям (корпорациям), и к университетам (при этом ясное понимание смысла применяемого термина зачастую отсутствует). Как уже отмечалось, термин «им-

пакт-фактор» без уточнений означает классический $IF_2(T, S)$. Разумеется, его численное значение можно подсчитывать и для автора, и для корпорации, но только по общему алгоритму (1) и с указанием библиографической базы данных. Однако для автора (или небольшой научной организации) этот показатель вполне может быть не только нулевым, но и неопределённым (если в течение двух лет подряд нет опубликованных работ). Кроме того, следует иметь в виду, что при таком применении понятия «импакт-фактор» равные значения этого показателя будут у автора, опубликовавшего за два года одну статью, получившую одну ссылку, и у автора, опубликовавшего за тот же период 10 статей, получивших 10 ссылок. Поэтому применение понятия «импакт-фактор» к авторам и корпорациям не представляется конструктивным.

Нередко под вывеской «импакт-фактор» выдаётся *среднее число ссылок в расчёте на одну статью* (автора, научно-исследовательского института, университета) независимо от времени цитирования и времени публикации этих статей. Однако это совсем иной показатель, не сводящийся к (1) ни при каких значениях параметров. Он вычисляется в WoS и называется *Average Citations per Item*. На наш взгляд, оптимальным русским эквивалентом этого названия может быть термин «*цитируемость*».

В отличие от авторов и корпораций, для журналов величина знаменателя в определении (1) относительно стабильна. Поэтому вполне корректно и плодотворно применение понятия «импакт-фактор» к такому объекту, как *группа журналов, объединённых общим признаком*. Таким признаком может быть предметная область. В качестве примера можно привести систематизацию в JCR примерно 10000 журналов по 230 предметным областям. Для сформированных таким образом групп (предметных областей) можно вычислить любой из

перечисленных выше показателей. В настоящее время в JCR выводятся следующие наукометрические показатели предметных областей:

- медианный импакт-фактор (*Median Impact Factor*),
- классический импакт-фактор (*Aggregate Impact Factor*),
- оперативный импакт-фактор (*Aggregate Immediacy Index*),
- медиана цитирования (*Aggregate Cited Half-Life*).

Поскольку публикационная активность ученых и традиции в цитировании публикаций предшественников и коллег имеют существенную дисциплинарную дифференциацию, сравнение наукометрического показателя конкретного журнала с соответствующим показателем предметной области позволяет более корректно оценить рейтинг журнала на мировом научном поле. Отношение импакт-фактора журнала к импакт-фактору предметной области естественно назвать *относительным импакт-фактором* $IF_{rn}(T, S)$ (с сохранением указанных выше атрибутов порядка n , запаздывания τ_0 и рейтинговой базы S).

Например, относительный классический импакт-фактор журнала «Успехи физических наук» (см. табл. 3) в предметной области Physics, Multidisciplinary, имеющей импакт-фактор $IF_2(2010, JCR) = 2,798$, составляет $IF_{r2}(2010, JCR) = 0,8$.

Индекс Хирша

Результаты научной деятельности авторов или организаций формально можно отражать числом публикаций (за все годы или за последние пять лет, как это принято в российской высшей школе) безотносительно к востребованности этих результатов. В той мере, в какой мы доверяем цитируемости как критерию востребованности и значимости публикаций, результаты научной деятельности можно характеризовать средним числом цитирований в расчёте на одну публикацию, общим числом ци-

тированных или цитирующих статей (с учетом или без учета самоцитирования и цитирования соавторами), а также их распределениями по годам.

В 2005 г. к показателям цитирования научных публикаций добавился быстро завоевавший признание показатель, получивший в англоязычных источниках название *h-index*, а в России более известный как «индекс Хирша» (по имени предложившего его учёного [26]). Этот показатель вычисляется на основе анализа распределения цитирований работ автора или организации.

Согласно данному Хиршем определению, автор (организация) имеет индекс h , если h из его статей цитируются не менее h раз каждая, в то время как оставшиеся статьи цитируются не более чем h раз каждая.

При подсчете индекса Хирша можно воспользоваться эквивалентным данному в [26], но более наглядным определением этого показателя: «индекс Хирша – это число статей, цитируемость которых не меньше их порядкового номера в ранжированном по цитированию списке публикаций». Таким образом, алгоритм вычисления индекса Хирша весьма прост: ранжируем все статьи данного автора (организации) в порядке убывания цитируемости и отбираем статьи с начала списка до тех пор, пока не подходим к статье с цитированием, меньшим порядкового номера этой статьи. Число предшествующих статей и есть индекс Хирша.

В работе [26] Хиршем дано статистическое обоснование адекватности предложенного им показателя для $h \gg 1$, основанное на анализе цитирований работ нобелевских лауреатов и членов National Academy of Sciences (США) за 20 лет, из которого вытекает высокая корреляция h -индексов с неформальными научными заслугами учёных.

Отметим ряд интересных свойств индекса Хирша.

1. Это не убывающий со временем по-

казатель, поскольку появление новых статей и новых ссылок на эти статьи не может его уменьшить.

2. В отличие от импакт-факторов, индекс Хирша практически не реагирует на аномально высокое цитирование отдельных работ. Упомянутый выше случай появления статьи, собравшей тысячи ссылок, либо не изменит индекс Хирша вовсе, либо увеличит его не более чем на единицу. Кроме того, он нечувствителен и к ссылкам на малоцитируемые статьи.

3. Это очень «вязкий» показатель, значение которого изменить тем труднее, чем оно выше.

4. Он «отсеивает случайных соавторов», поскольку его величина будет значительной лишь у тех авторов, которые имеют достаточно много публикаций, многие из которых часто цитируются.

Общепризнано, что индекс Хирша обеспечивает более адекватную рейтинговую оценку исследователей, имеющих значительный стаж научной деятельности, чем могут дать такие показатели, как число публикаций или число цитирований. Вместе с тем следует отметить, что вследствие инерционности процесса накопления ссылок индекс Хирша, как и другие связанные с цитированием показатели, вряд ли целесообразно использовать для оценки научной деятельности молодых ученых [27].

Применение индекса Хирша для оценки результативности научной деятельности и построения различного рода рейтингов требует учета следующих важных обстоятельств. Во-первых, этот индекс, как и многие другие наукометрические индикаторы, в отрыве от библиографической базы не имеет смысла. Кроме того, он привязан ко времени подсчёта. Поэтому индекс Хирша, как и группа импакт-факторов, является функцией двух параметров – T и S . Для выполненных Хиршем оценок библиографической базой служил ресурс Института научной информации ISI (ныне WoS), отражавший публикации с 1955 г.

Во-вторых, индекс Хирша не может быть сопоставимым для разных областей науки. Его целесообразно применять для сравнения достижений исследователей, научных коллективов, организаций, работающих в общей предметной области. В связи с этим вряд ли следует серьезно воспринимать предпринимавшиеся попытки использования индекса Хирша для построения рейтинга ректоров российских вузов, а также для ранжирования крупных многопрофильных научно-образовательных организаций.

С точки зрения практического применения важно отметить, что в области точных и естественных наук использование для расчета индекса Хирша баз данных WoS и SCP (в отличие от РИНЦ) в целом дает сравнимые результаты: $h(T, WoS) \approx h(T, SCP)$. В области социогуманитарных наук, вследствие слабой представительности российских журналов в международных библиографических базах, для наукометрических оценок целесообразно использовать данные РИНЦ.

Применение многих наукометрических показателей для сравнения результативности научной деятельности отдельных авторов, научных групп и организаций зачастую осложняется проблемой адекватности поисковых запросов, формируемых в библиографической базе разными операторами. Для российских авторов в РИНЦ это можно считать малосущественным. В англоязычных базах соответствующие ограничения для изучения результативности авторов связаны с необходимостью корректного учёта вариантов написания некоторых фамилий и очистки списка от однофамильцев по дополнительным признакам. Однако при изучении научной продуктивности организаций это может быть серьезной проблемой. Так, один из сотрудников Нижегородского университета (ННГУ) сообщил нам, что в его статьях, отображаемых в WoS, насчитывается 15 англоязычных вариантов наименования ННГУ.

Формирование адекватной маски корпорации, не приводящей к потере данных и надёжно отсекающей «шум», является сложной задачей. Например, при анализе результатов недавней попытки построить рейтинг вузов по данным WoS выяснилось, что применение предложенной разработчиками рейтинга (Сибирский федеральный университет) маски занижает число опубликованных от имени ННГУ и учтённых в WoS статей приблизительно в полтора раза.

В заключение отметим, что, несмотря на отмеченные в этой статье сложности, неоднозначности и оговорки, связанные с применением показателей научного цитирования, следует признать, что в настоящее время, по крайней мере в ведущих университетах страны, профессор, имеющий в каждой из упомянутых в этой статье рейтинговых баз индекс Хирша, не слишком сильно отличающийся от единицы, так же малопривлекателен для работодателя, как и журнал с импакт-фактором $IF_2 = 0,0N$ для исследователя, готовящего к публикации значимый научный результат.

Литература

1. Прайс Д. Малая наука, большая наука // Наука о науке. М.: Прогресс, 1966.
2. http://vak.ed.gov.ru/ru/help_desk/list/
3. Garfield E. Citation indexes to science: a new dimension in documentation through association of ideas // Science. 1955. V. 122. P. 108–111.
4. Garfield E., Sher I.H. New factors in the evaluation of scientific literature through citation indexing // American Documentation. 1963. Vol. 14. No. 3. P. 195–201.
5. Налимов В.В., Мультченко З.М. Наукометрия. Изучение развития науки как информационного процесса. М.: Наука, 1969.
6. Бернал Дж. Наука и общество: Сб. статей и выступлений. М.: Изд-во иностр. лит., 1953. 299 с.
7. Хайтун С.Д. Наукометрия. Состояние и перспективы. М.: Наука, 1983.
8. Баранов А.Н. Семантическая сеть как

- инструмент библиометрии в гуманитарных науках. URL: <http://iph.ras.ru/upfile/ideol/bibliometr/Baranov.html>
9. *Гарфилд Ю.* Можно ли выявлять и оценивать научные достижения и научную продуктивность? // Вестник АН СССР. 1982. № 7. С. 42–50.
 10. *Shockley W.* On the Statistics of Individual Variations of Productivity in Research Laboratories // Proceedings of the Institute of Radio Engineers. 1957. V. 45. № 279. P. 1409.
 11. *Пельц Д., Энрюс Ф.* Ученые в организациях. Оптимальные условия для исследований и разработок. М.: Прогресс, 1973. 469 с.
 12. *Алескеров Ф.Т., Писляков В.В., Субочев А.Н., Чистяков А.Г.* Построение рейтингов журналов по менеджменту с помощью методов теории коллективного выбора: препринт WP7/2011/04 / НИУ – «Высшая школа экономики». М.: Изд. дом ВШЭ, 2011. 44 с.
 13. *Писляков В.В.* Методы оценки научного знания по показателям цитирования // Социологический журнал. 2007. № 1. С. 128–140.
 14. *Rousseau R.* Journal evaluation: Technical and practical issues // Library Trends. 2002. V. 50. Iss. 3. P. 418–439.
 15. *Glänzel W., Moed H.F.* Journal impact measures in bibliometric research // Scientometrics. 2002. Vol. 53. N. 2. P. 171–193.
 16. *Rousseau R.* Median and percentile impact factors: A set of new indicators // Scientometrics. 2005. Vol. 63. N. 3. P. 431–441.
 17. *Egghe L.* Mathematical relations between impact factors and average number of citations // Information Processing and Management. 1988. V. 24. P. 567–576.
 18. *Egghe L., Rousseau R.* Citation distribution of pure mathematics journals // Informetrics 87/88 /ed. Amsterdam: Elsevier, 1988. P. 249–262.
 19. *Garfield E.* How ISI Selects Journals for Coverage: Quantitative and Qualitative Considerations // Current Contents. 1990. May 28.
 20. *Rotb D.L.* The emergence of competitors to the Science Citation Index and the Web of Science // Current Science. 2005. V. 89. Nr 9, 10. P. 1531–1536.
 21. *Теста Дж.* Процесс отбора журналов в Thomson Reuters. URL: thomsonreuters.com/content/science/pdf/ssr/training/rus-web.pdf
 22. *Кириллова О.В.* Подготовка российских журналов для зарубежной аналитической базы данных Scopus. Рекомендации и комментарии. – <http://elsevierscience.ru/info/add-journal-to-scopus/>
 23. *Jacso P.* As we may search – Comparison of major features of the Web of Science, Scopus and Google Scholar citation-based and citation-enhanced databases // Current Science. 2005. V.89. N 9, 10. P. 1537–1547.
 24. http://elibrary.ru/projects/science_index/science_index_questions.asp
 25. *Солошенко Н.С., Кириллова О.В.* Отражение российских журналов в БД Science Citation Index и SCOPUS // Educational Technology & Society. 2006. Vol. 9. Nr 3. P. 313–320.
 26. *Hirsch J. E.* An index to quantify an individual's scientific research output // Proceedings of the National Academy of Sciences. 2005. Vol. 102. No. 46. P. 16569–16572.
 27. *Бедный Б.И., Миронос А.А., Серова Т.В.* Продуктивность исследовательской работы аспирантов (наукометрические оценки) // Высшее образование в России. 2006. № 7. С. 20–36.

BEDNYI B., SOROKIN YU. ON INDICATORS OF SCIENCE CITATION AND ITS APPLICATION

Indicators of science citation applied for quantitative evaluation of science activities are discussed. The concept “impact-factor” is considered as a group of indicators, which differ by the algorithm of calculation. The properties of h-index are analyzed. It is shown, that the values of indicators, which are connected with the process of publications and citations measurement, depend significantly on the rating base (citation index), which is used for calculation. Capabilities and restrictions for usage the scientometric indicators in higher school are assessed.

Key words: scientific journal, average citations per item, citation index, classic impact-factor, aggregate impact factor, relational impact-factor, h-index.