

ОРГАНИЗАЦИЯ ИССЛЕДОВАТЕЛЬСКОЙ ДЕЯТЕЛЬНОСТИ

НОВЫЕ ИНДЕКСЫ ПУБЛИКАЦИОННОЙ АКТИВНОСТИ

© 2020 г. В. В. Миронов

Рязанский государственный радиотехнический университет им. В.Ф. Уткина, Рязань, Россия

E-mail: mironov1vv@mail.ru

Поступила в редакцию 05.03.2020 г.

После доработки 27.03.2020 г.

Принята к публикации 12.04.2020 г.

В статье кратко описаны результаты исследования информативности широко известного и просто-го в применении индекса цитирования Хирша, отмечены его достоинства и недостатки. Как альтернатива представлены четыре новых трёхмерных индекса цитирования: M -индекс, взвешенный M -индекс, G -индекс сверхвысокого цитирования и взвешенный G -индекс. Для удобства сравнения публикационной активности учёных введены их одномерные псевдонормы (оценки). Вводимые показатели просты в использовании, включают привычную для научного сообщества методологию индекса Хирша, но существенно информативнее последнего.

Ключевые слова: библиометрия, наукометрический показатель, индекс Хирша, новые многомерные индексы цитирования, одномерные индексы цитирования, информативность индексов.

DOI: 10.31857/S0869587320080083

Индекс Хирша (h -индекс) — наукометрический числовой показатель, предложенный в 2005 г. американским учёным аргентинского происхождения Хорхе Хиршем из Калифорнийского университета (Сан Диего, США) для оценки публикационной активности физиков [1]. Он служит количественной характеристикой продуктивности учёного (группы учёных, научной организации и даже страны в целом), исчисляемой на основе его публикаций в любых изданиях и их цитирования. Ввиду простоты вычисления и необходимости количественно оценивать работу специалистов индекс Хирша быстро распространился на все виды научной деятельности.

Обратим внимание, что h -индекс был предложен для учёных, работающих в одной и той же

предметной области и имеющих близкий научный стаж от года первой публикации. Сейчас это требование зачастую нарушается, ибо идея Хирша оказалась, при поверхностном взгляде, весьма продуктивной и простой по алгоритму расчёта.

ОСОБЕННОСТИ ИНДЕКСА ХИРША

Так уж складывается научная система, что индекс цитирования, с одной стороны, помогает чиновникам управлять финансированием работ и карьерным ростом руководителей, а с другой стороны, учитывая излишнее внимание к нему, порождает определённые проблемы, в частности, стимулирует неоправданное самоцитирование, тиражирование статей (их пересказ без новых результатов — самоплагиат), использование подчинённых и учеников для искусственного повышения этого показателя.

Вычисляется индекс Хирша так: если h статей учёного из их общего числа N_p цитируются h (или более) раз и каждая из оставшихся $N_p - h$ статей цитируется менее (или ровно) h раз, то h -индекс учёного равен натуральному числу h (в противном случае, если нет статей или ссылок на них, $h = 0$).

Принято считать, что индекс Хирша оценивает так называемое ядро публикационной активности учёного и формируется через Интернет на



МИРОНОВ Валентин Васильевич — доктор физико-математических наук, профессор кафедры высшей математики РГРТУ им. В.Ф. Уткина.

основе открытых наукометрических баз данных, например, Google Scholar [2], eLIBRARY.ru [3], ADS NASA [4], а также платных баз Scopus или Web of Science Group.

Заметим, для одного и того же учёного h -индекс может быть различным, если его подсчитывать, используя разные базы данных или разные по широте охвата публикации [5]. При этом самоцитирования (ссылки автора на свои же статьи) в индексе не учитываются [6].

Очевидно, этот показатель не оптимален для своих же целей, его недостатки были отмечены в основополагающей статье [1]. Главный из них состоит в том, что соотношение h -индексов учёных часто не соответствует их вкладу в науку (подразумевается, что чем он выше, тем больше вклад специалиста в развитие определённого направления). К другим недостаткам индекса можно отнести его автоматизированный расчёт: он возможен только по одной базе и только по публикациям, проиндексированным в ней, иначе потребуются ручной разбор дубликатов как публикаций, так и цитирований.

К примеру, формальный h -индекс основателя современной высшей алгебры Э. Галуа (1811–1832), будь он подсчитан сегодня, стал бы равным 4. В то же время Cited Reference Search по Web of Science Core Collection выделяет 54 работы, которые необходимо объединить вручную. Кроме того, вклад Э. Галуа фиксировался в трудах его последователей, получавших ссылки, а также в именовании артефактов, разработанных автором, без необходимости устанавливать ссылки на исходные публикации.

Другой пример: h -индекс А. Эйнштейна в 1906 г., разумеется, при условии его существования в те времена, равнялся бы 5, несмотря на очень высокий показатель цитирования учёного в 1905 г. [7]. При этом вклад обоих гениев в науку чрезвычайно высок.

Из этих нереферентных для методологии Хирша примеров вытекает, что за пределами внимания его метода остаётся весьма важная библиографическая информация, свидетельствующая о весомости вклада учёных в развитие науки.

Положительные стороны h -индекса также очевидны: простота вычисления, доступность данных и неплохая, на первый взгляд, оценка публикационной активности специалиста.

Недостатки количественной оценки вклада учёного в науку, предложенной Х. Хиршем, породили многочисленные идеи по её улучшению на основе совершенствования авторской методологии и других принципов [8–18]. Особо выделим оригинальную работу Л. Вальтмана и Н.Й. ван Экка [19]. Ещё около четырёх десятков публикаций с различными предложениями по модернизации идеи Хирша или удалёнными от неё ново-

ведениям в библиометрию можно найти на сайте [20].

Выдвинутые предложения не лишены серьёзных недостатков. Главные из них — отход от привычной и повсеместно принятой системы Хирша, вычислительная сложность и небольшое количество преимуществ над h -индексом, а также психологическое неприятие дискуссионных предложений. Это относится, по моему мнению, и к модернизации Вальтмана и ван Экка.

Однако неугасающий интерес научной общности и организаторов науки, в том числе чиновников, к учёту научной и публикационной активности сотрудников говорит о том, что эта тема, если так можно выразиться, фрактальна: всякие попытки решить её до конца обречены на провал. Значит, нужно остановиться на адекватной и приемлемой мере оценки научной активности учёного и смириться с её недостатками.

В этом контексте можно предложить два новых трёхмерных индекса: M -индекс (стандартный) и G -индекс (нестандартный, или индекс сверхвысокого цитирования) для оценки ядра вклада учёного в науку по данным цитируемости его работ при большом и небольшом числе публикаций. С целью детализации информации и учёта квартилей журналов, где опубликованы цитируемые работы, вводятся обобщения этих индексов в виде взвешенного M -индекса (стандартный) и взвешенного G -индекса цитирования.

Новые индексы, сохраняя плодотворную идею Хирша о ядре цитирования работ учёного, простоту и наглядность оценки, на порядок превосходят классический h -индекс Хирша по *качеству* оценки вклада или влияния учёного на развитие той или иной отрасли науки на основе данных о цитировании его работ (корреляция “чем больше публикаций, тем больше вклад” сомнению здесь не подвергается). Кроме того, путём введения псевдонорм индексов строится одномерный индекс цитирования (оценка публикационной активности учёного), позволяющий *количественно* оценить ядро, по значению которого удобно ранжировать представителей науки.

Новизна идеи достигается за счёт рассмотрения новых понятий — так называемых хвоста и подвала ядра цитирования, а также учёта квартилей журналов, где была напечатана цитируемая статья.

В рамках предлагаемой методологии рекомендуется учитывать базу данных о публикациях учёного в российских журналах Перечня ВАК и базы более высокого уровня — Scopus, Web of Science и другие, хотя это не принципиально. Для определения новых индексов может быть использована любая база данных о публикациях.

М-ИНДЕКС ЦИТИРОВАНИЯ КАК ОБОБЩЕНИЕ ИНДЕКСА ХИРША

Введём необходимые определения, данные для алгоритмов и сделаем некоторые комментарии к ним.

По определению *М*-индекс публикационной активности учёного, рассчитанный на основе анализа цитирований его научных публикаций, есть упорядоченная тройка натуральных чисел

$$M = (h, p, q), \quad (1)$$

где *h* — индекс Хирша, *p*, *q* — дополнительные индексы, смысл и определение которых излагаются ниже.

Расположим все статьи автора *NN* списком сверху вниз по мере убывания числа ссылок на его публикации — цитирования в журналах Перечня ВАК и в базах Scopus, Web of Science и др. Статьи с одинаковым числом цитирований распределяются в произвольном порядке. Получим *P*-список публикаций автора *NN* с общим числом работ *N_p* (по определению *N_p* > 0).

Представим *P*-список в виде таблицы, которую назовём *P*-таблицей. В ней по вертикали расположены статьи (удобнее присваивать им номера) автора *NN*, вошедшие в *P*-список; а по горизонтали в каждой строке, соответствующей выбранной статье, — коэффициенты *k_i* = 1, каждый из которых означает, что статья процитирована кем-то один раз.

За единицу измерения индексов примем величину *e* = 1 × 1, где единица слева — одна статья, которая цитируется один раз (единица справа).

Ядро цитирования *М*-индекса — это квадратная таблица размером *h* × *h*, где сторона *h* соответствует индексу Хирша. При этом “физические” размерности (измерения) сторон квадрата разные: первое *h* — число статей, второе *h* — минимальное число цитирований каждой стати, при этом хотя бы одна из них имеет ровно *h* цитирований.

Рубеж цитирования — это число *μ*, которое является предельным для учёта числа цитирований той или иной статьи. Другими словами, если число цитирований статьи превысило число *μ*, то эти цитирования не входят в зачёт *М*-индекса.

Для чего вводится рубеж цитирования? Дело в том, что, начиная с некоторого предельного числа цитирований, возникает своего рода мода на упоминание статей автора по соответствующей теме. Учёные, особенно молодые, считают престижным сослаться именно на этого известного, “модного” автора, придавая тем самым вес своим работам. Такая мода в своё время была на великих учёных. Сегодня к наиболее цитируемым авторам относится небольшая группа лауреатов Нобелевской премии, а также быстро прогрессирующие

специалисты в тех или иных областях науки; их имена, списки можно найти в Интернете, сохраняя при этом определённый скептицизм по отношению к подобного рода классификациям.

Исходя из собственного опыта и анализа открытых индексов цитирования учёных, автор статьи предлагает рассмотреть два случая рубежа цитирования.

Случай А. Если индекс *h* ≤ 50, то фиксированный рубеж цитирования *μ* = 100. Как следствие, для таких авторов индекс Хирша в данном способе подсчёта всегда будет не больше рубежа цитирования *h* ≤ *μ*.

Случай В. Если индекс *h* > 50, то плавающий рубеж цитирования

$$\mu = 2h. \quad (2)$$

Анализируя ситуацию, в том числе принимая во внимание метазакон Мерфи “чем больше ты на публике, тем больше тебя приглашают”, введение предельного числа *μ* при определении и продвижении *М*-индекса цитирования можно считать обоснованным, равно как и его ограничение (2).

Однако оговоримся, есть ряд работ учёных, которые оказывают колоссальное влияние на развитие той или иной отрасли науки или направления. Ограничение числом *μ* количества ссылок на такие работы, вне сомнения, снижает уровень влияния их авторов. Вот почему для учёного, число цитирований работ которого превосходит 2*μ* (это уже не мода!), рекомендовано вычислять его *G*-индекс цитирования (индекс сверхвысокого цитирования, по моей терминологии).

Хвост ядра цитирования — это таблица, где по вертикали расположены все статьи (или их номера), названия которых входят в ядро цитирования, а по горизонтали — все цитирования упомянутых статей, не вошедшие в ядро.

Подвал ядра цитирования — это таблица, где по вертикали расположены все статьи (или их номера), названия которых не вошли в ядро цитирования, а по горизонтали — все цитирования упомянутых статей, при этом каждая имеет ровно *h* цитирований.

Для подвала, как и для хвоста цитирования, вводится понятие “рубеж подвала”, под которым подразумевается число учитываемых в подвале статей, обозначенное *r*. Для обоих случаев *А* и *В* полагаем, что *r* ≤ *h*. Если таковых статей больше *r*, их просто не учитываем в подвале.

Рубеж подвала можно рассматривать как меру борьбы с тиражированием статей, когда автор, желая повысить свою публикационную активность, издаёт практически одну и ту же статью под разными названиями.

Для удобства изложения термины “ядро цитирования”, “хвост цитирования”, “подвал цитирования” и другие перенесём в P -таблицу.

Из определений вытекает следующее утверждение: если для одной статьи из ядра в хвосте учтены t цитирований, то для случая A имеет место неравенство

$$h + t \leq \mu, \quad (3)$$

а для случая B неравенство

$$t \leq h. \quad (4)$$

АЛГОРИТМ ВЫЧИСЛЕНИЯ M -ИНДЕКСА

Представим все статьи автора NN в виде P -таблицы.

В P -таблице выделим (по методологии Хирша) h статей из их общего числа N_p , на каждую из которых приходится h или более цитирований, а на каждую из оставшихся $N_p - h$ статей — h или менее цитирований. Следовательно, индекс Хирша автора NN равен h , а сформированное ядро цитирования M -индекса — квадрат $h \times h$. При этом площадь ядра $S_0 = h^2$.

Пересчитаем все цитирования, то есть все единицы статей, вошедших в хвост и не вошедших в ядро. Напомним, в такой пересчёт может войти не более μ цитирований на каждую статью и каждая статья из ядра имеет максимально длинный хвост цитирования по сравнению со статьями, не вошедшими в ядро. Пусть таких цитирований в хвосте будет S_1 штук.

Пересчитаем все цитирования статей, вошедших в подвал и не вошедших в ядро. Напомним, в такой пересчёт может войти не более h статей и каждая имеет ровно h цитирований. Пусть таких цитирований в подвале будет S_2 штук.

По определению M -индекс публикационной активности учёного, рассчитанной на основе анализа цитирований его научных публикаций, есть упорядоченная тройка натуральных чисел

$$M = (h, p, q), \quad (5)$$

где h — индекс Хирша,
индекс

$$p = [\sqrt{S_1}], \quad (6)$$

индекс

$$q = [\sqrt{S_2}], \quad (7)$$

в формулах (6), (7) символ [...] — стандартная функция целой части числа.

Введение дополнительных индексов p и q именно в видах (6), (7) обосновано тем, что ядро Хирша — квадрат площадью $S_0 = h^2$, а площади

дополнительного цитирования в хвосте и подвале с учётом единицы измерения равны S_1 и S_2 соответственно.

Индекс (5) позволяет определить качественную и количественную оценку ядра цитирования и его окружения в виде хвоста и подвала. Но тройка чисел (5) психологически воспринимается, особенно неподготовленным пользователем, с трудом.

Для удобства сравнения публикационной активности учёных — обозначим её $A(NN)$ для учёного NN — в целях поддержания финансирования работ или кадрового роста персонала введём одномерный псевдонормированный индекс δ (публикационная активность учёного NN), положив, что

$$\delta = [\|M\|] = [\sqrt{h^2 + p^2 + q^2}], \quad (8)$$

в формуле (8) символ [...] — стандартная функция целой части числа.

Для двух учёных NN_1 и NN_2 с M -индексами их публикационной активности $M_1 = (h_1, p_1, q_1)$ и $M_2 = (h_2, p_2, q_2)$ и индексами δ_1 и δ_2 соответственно положим, что

$$A(NN_1) > A(NN_2) \stackrel{\text{def}}{\Leftrightarrow} \begin{cases} \delta_1 > \delta_2, \\ \delta_1 = \delta_2 \wedge h_1 > h_2, \\ \delta_1 = \delta_2 \wedge h_1 = h_2 \wedge p_1 > p_2, \\ \delta_1 = \delta_2 \wedge h_1 = h_2 \wedge p_1 = p_2 \wedge q_1 > q_2. \end{cases} \quad (9)$$

В противном случае публикационная активность авторов считается одинаковой: $A(NN_1) = A(NN_2)$.

В формуле (9) знак $\wedge$ — стандартный символ бинарной операции конъюнкции двух высказываний, более слабый по сравнению с соседними символами равенства или неравенства.

С целью детализации информации индексы p , q , δ можно вычислять с точностью до десятых долей.

СВОЙСТВА НОВЫХ ИНДЕКСОВ

Следующие утверждения являются простым следствием определений:

- для любого автора, у которого существуют индексы цитирования, $\delta \geq h$;
- в подвале ядра не может быть статей с числом цитирований больше, чем h ; в противном случае либо эта статья была бы уже учтена в ядре цитирования, либо ядро имело бы размеры $s \times s$ при $s > h$, что противоречит определению ядра; напомним, что здесь для индекса Хирша определена верхняя граница учёта $\max h = 50$, при зна-

чении $h > 50$ рекомендовано перейти к определению G -индекса;

- для рекомендованного индекса Хирша предлагаемый индекс δ не превосходит значения $\max \delta = 85$; из соответствующих определений следует, что для рекомендованных индексов $\max S_0 = \max S_1 = \max S_2 = 50 \times 50$, следовательно, в силу равенств (6), (7), (8) индекс δ ограничен числом 85.

Индекс Хирша, равный 50, — это очень высокий показатель. Достаточно посмотреть в Интернете h -индекс российских учёных — членов РАН или членов Американской академии наук. Максимальный индекс δ может превосходить этот показатель в 1.7 раза.

Будем говорить, что

- индекс X , где синтаксическая переменная X может принимать значения $X = h, p, q, M, \delta$, устойчив в области D , где синтаксическая переменная D может принимать значения $D = \text{ядро, хвост, подвал}$, если никакие одно-два “случайных” цитирования (новых или старых) статей автора из области D не могут с большой долей вероятности изменить значения индекса X ; в противном случае индекс X неустойчив в области D ;

- индекс X глобально устойчив, если он устойчив при всех значениях переменной D .

Итак, кратко сформулируем два основных тезиса.

Тезис 1. Индексы h, p, q — устойчивые характеристики в своих областях; индекс δ — глобально устойчивая характеристика, более устойчивая, нежели индекс Хирша.

В самом деле, нетрудно посчитать классические (статистические) вероятности изменения характеристик h, p, q, δ при одном-двух “случайных” цитированиях (новых или старых) статей автора из области D и убедиться в относительно небольших значениях этих вероятностей.

Из определения (8) и проведённых подсчётов вытекает, что индекс δ на порядок устойчивее индекса h .

Автор, без сомнения, отметил свойство устойчивости индекса Хирша на собственном опыте.

Тезис 2. Индекс M на порядок информативнее индекса h .

Данный тезис, как и заключительную часть первого, нельзя доказать математически. Для этого нужно дать строгие определения “информативности” и “объективности”, ввести и проранжировать их порядки. В любом случае это спорная тема, и потому здесь она не рассматривается. Тем не менее тезис можно обосновать на содержательном уровне, основываясь на принципах интуиционистской математики [21–23].

По моему мнению, учёт информации, вошедшей в хвост и подвал ядра цитирования, существенно расширяет общее представление о публикационной активности учёного как в количественном, так и в качественном аспектах. И наоборот, неучёт такой информации существенно сужает общее представление о его публикационной активности. Ниже будут приведены примеры, иллюстрирующие эти утверждения.

ВЗВЕШЕННЫЙ М-ИНДЕКС ЦИТИРОВАНИЯ

В ряде случаев возникает желание и необходимость учитывать статус журналов, в которых опубликованы цитируемые статьи. Многие издания из Перечня ВАК относятся к третьему или четвёртому квартилям и лишь отдельные — к первому или второму. Считается, и не без оснований, что авторы, ссылающиеся на те или иные публикации, с большей ответственностью относятся к цитированию статей, опубликованных в серьёзных журналах из первого или второго квартилей. При этом заметим, что число постоянных авторов, публикующих статьи в журналах этой категории, достаточно ограниченное. Такое сужение в какой-то мере компенсирует необычайно возросший круг специалистов по многим направлениям науки, что служит одним из критериев продолжающейся научной революции, а также научных журналов. Их квартиль, по мнению автора, должен учитываться, если необходим более тщательный анализ библиометрической информации.

Как осуществить учёт квартиля журнала в рамках предлагаемых изменений индекса Хирша?

Представим P -таблицу в модернизируемом виде, назвав её PM -таблицей. В ней по вертикали расположены всё те же статьи (или их номера) автора NN , вошедшие в P -список; а по горизонтали в каждой строке, соответствующей выбранной статье, — коэффициенты $k_i = 1/i$, означающие, что статья вышла в журнале с квартилем i и процитирована кем-то один раз.

Положим, что

$$S_0 = k_1 n_{1,0} + k_2 n_{2,0} + k_3 n_{3,0} + k_4 n_{4,0}, \quad (10)$$

где $k_i = 1/i$ — коэффициент журнала i -й квартиля, в котором напечатана процитированная статья из ядра цитирования, а $n_{i,0}$ — число таких цитирований с квартилем k_i ;

$$S_1 = k_1 n_{1,1} + k_2 n_{2,1} + k_3 n_{3,1} + k_4 n_{4,1}, \quad (11)$$

где k_i — имеет тот же смысл, что и в (10), а $n_{i,1}$ — число таких цитирований с квартилем k_i ;

$$S_2 = k_1 n_{1,2} + k_2 n_{2,2} + k_3 n_{3,2} + k_4 n_{4,2}, \quad (12)$$

где k_i — имеет тот же смысл, что и в (10), (11), а $n_{i,2}$ — число таких цитирований с квартилем k_i .

Положим, что индексы

$$h^* = [S_0], \quad (13)$$

$$p^* = [\sqrt{S_1}], \quad (14)$$

$$q^* = [\sqrt{S_2}], \quad (15)$$

$$M^* = (h^*, p^*, q^*), \quad (16)$$

в формулах (13), (14), (15) символ [...] — стандартная функция целой части числа.

И далее снова, но уже для взвешенных индексов цитирования:

$$\delta^* = |||M^*||| = [\sqrt{(h^*)^2 + (p^*)^2 + (q^*)^2}], \quad (17)$$

$$\begin{aligned} & A(NN_1) > A(NN_2) \stackrel{\text{def}}{\Leftrightarrow} \\ & \Leftrightarrow \begin{cases} \delta_1^* > \delta_2^*, \\ \delta_1^* = \delta_2^* \wedge h_1^* > h_2^*, \\ \delta_1^* = \delta_2^* \wedge h_1^* = h_2^* \wedge p_1^* > p_2^*, \\ \delta_1^* = \delta_2^* \wedge h_1^* = h_2^* \wedge p_1^* = p_2^* \wedge q_1^* > q_2^*. \end{cases} \end{aligned} \quad (18)$$

В противном случае публикационная активность авторов считается одинаковой: $A(NN_1) = A(NN_2)$.

С целью детализации информации индексы p^* , q^* , δ^* , как и ранее, можно вычислять с точностью до десятых долей.

Алгоритм вычисления взвешенного индекса цитирования остаётся тем же, меняется только содержание P -таблицы (единицы заменяются соответствующими дробями).

G-ИНДЕКС ДЛЯ УЧЁТА СВЕРХВЫСОКОГО ЦИТИРОВАНИЯ

По определению G -индекс публикационной активности учёного, рассчитанной на основе анализа цитирований его научных статей, есть упорядоченная тройка натуральных чисел

$$G = (H, P, Q), \quad (19)$$

где H — индекс, формируемый на основе индекса Хирша, P , Q — дополнительные индексы, формируемые на основе индексов p , q соответственно.

Во всех новых индексах выбрана другая единица измерения по сравнению с M -индексом, а именно $E = 10 \times 10$, где 10 слева — это 10 статей, каждая из которых цитируется 10 раз (справа). Единицу $E = 10 \times 10$ будем называть также единичным квадратом.

Хвост и подвал ядра цитирования для M -индекса и G -индекса совпадают. Для G -индекса

ограничения хвоста и подвала цитирования не предусмотрены.

ОПРЕДЕЛЕНИЕ КОМПОНЕНТ ВЕКТОРА G И АЛГОРИТМ ВЫЧИСЛЕНИЯ G -ИНДЕКСА

Расположим все статьи автора NN списком сверху вниз по мере убывания числа ссылок на его статьи в журналах Перечня ВАК и в базах Scopus, Web of Science и др. Статьи с одинаковым числом цитирований размещаются в произвольном порядке. Получим список A публикаций автора NN с общим числом работ N_p .

Из списка A выделим (по методологии Хирша) $10H$ статей, если таковые, конечно, имеются, из общего числа N_p , на каждую из которых приходится $10H$ или более цитирований, а на каждую из оставшихся $N_p - 10H$ статей — $10H$ или менее цитирований.

Следовательно, индекс Хирша автора NN равен $10H$, а первая компонента вектора G есть число H . Значит, сформировано ядро цитирования G -индекса — квадрат $10H \times 10H$. При этом площадь ядра $S_0 = (10H)^2$.

Пересчитаем все единичные квадраты, вошедшие в хвост целиком или частично и не вошедшие целиком в ядро. Пусть таких квадратов цитирований в хвосте будет S_1 штук.

Пересчитаем все единичные квадраты, вошедшие в подвал целиком или частично и не вошедшие в ядро. Пусть таких квадратов цитирований в подвале будет S_2 штук.

По определению G -индекс публикационной активности учёного, рассчитанной на основе анализа цитирований его научных публикаций, есть упорядоченная тройка натуральных чисел

$$G = (H, P, Q), \quad (20)$$

где H — определён ранее как “укрупнённый” аналог ядра Хирша,

индекс

$$P = [\sqrt{S_1}], \quad (21)$$

индекс

$$Q = [\sqrt{S_2}], \quad (22)$$

где [...] — стандартная функция целой части числа.

Как и ранее, для удобства сравнения публикационной активности учёных — обозначим её $A(NN)$ для учёного NN — в целях поддержания финансирования работ или кадрового роста персонала введём одномерный псевдонормирован-

ный индекс Δ (публикационная активность учёного NN), положив, что

$$\Delta = \|G\| = [\sqrt{H^2 + P^2 + Q^2}], \quad (23)$$

где [...] — стандартная функция целой части числа.

Для двух учёных NN_1 и NN_2 с G -индексами их публикационной активности $G_1 = (H_1, P_1, Q_1)$ и $G_2 = (H_2, P_2, Q_2)$ и индексами Δ_1 и Δ_2 соответственно положим, что

$$A(NN_1) > A(NN_2) \stackrel{\text{def}}{\Leftrightarrow} \begin{cases} \Delta_1 > \Delta_2, \\ \Delta_1 = \Delta_2 \wedge H_1 > H_2, \\ \Delta_1 = \Delta_2 \wedge H_1 = H_2 \wedge P_1 > P_2, \\ \Delta_1 = \Delta_2 \wedge H_1 = H_2 \wedge P_1 = P_2 \wedge Q_1 > Q_2. \end{cases} \quad (24)$$

В противном случае публикационная активность авторов считается одинаковой: $A(NN_1) = A(NN_2)$.

Из определений с очевидностью вытекает равенство

$$H = [h/10], \quad (25)$$

где [...] — по-прежнему стандартная функция целой части числа.

Как и ранее, с целью детализации информации компоненты вектора G можно вычислять с точностью до десятых долей.

По аналогии со взвешенным M -индексом может быть введён взвешенный G -индекс цитирования.

Проиллюстрируем несколькими примерами обоснованность претензий M -индекса на большую объективность в оценке публикационной активности автора по сравнению с индексом Хирша при сопоставимой простоте вычисления (для удобства восприятия единицу E в индексе сверхвысокого цитирования будем обозначать числом 1, как и в случае индекса Хирша или M -индекса):

- автор NN_1 опубликовал 1 работу, которая имеет 1 цитирование — из определений вытекает, что $h = 1$, $M = (1, 0, 0)$, $\delta = 1$;

- автор NN_2 опубликовал 1 работу, которая имеет 100 цитирований — из определений вытекает, что $h = 1$, $M = (1, 9, 0)$, $\delta = 9$;

- автор NN_3 опубликовал 10 работ, каждая из которых имеет по 10 цитирований — из определений вытекает, что $h = 10$, $M = (10, 0, 0)$, $\delta = 10$;

- автор NN_4 опубликовал 10 работ, из которых первая в списке имеет 100 цитирований, вторая — 90, третья — 80 и т.д., последняя, десятая в спис-

ке, — 10 цитирований — из определений вытекает, что $h = 10$, $M = (10, 21, 0)$, $\delta = 23$;

- автор NN_5 опубликовал 10 работ, каждая из которых имеет по 10 цитирований — из определений вытекает, что $h = 10$, $M = (10, 0, 0)$, $\delta = 10$;

- автор NN_6 опубликовал 100 работ, каждая из которых имеет по 10 цитирований — из определений вытекает, что $h = 10$, $M = (10, 0, 10)$, $\delta = 14$;

- автор NN_7 опубликовал 30 работ, из которых 10 имеют по 100 цитирований, а оставшиеся 20 — по 10 цитирований — из определений вытекает, что $h = 10$, $M = (10, 30, 10)$, $\delta = 33$ (заметим, для методики Хирша эта ситуация равносильна случаю, когда каждая из 33 работ автора имела бы минимум по 33 цитирования; если в данном примере все цитируемые публикации были бы в журналах третьего квартиля, то из определений вытекает, что $h^* = 5$, $M^* = (5, 17, 5)$, $\delta^* = 18^*$; таким образом, учёт квартилей журналов может привести к существенному изменению статуса автора публикаций);

- автор NN_8 опубликовал 500 работ, из которых 100 имеют по 500 цитирований, оставшиеся 400 — по 200 цитирований — из определений вытекает, что для индекса Хирша $h = 100$, для M -индекса $M = (100, 100, 100)$, $\delta = 170$, для G -индекса (в новых единицах измерения) $G = (20, 17, 20)$, $\Delta = 33$.

Таким образом, представив четыре новых трёхмерных индекса цитирования как альтернативу индексу цитирования Хирша, автор попытался отойти от принятых шаблонов и схем в оценке деятельности научных и педагогических работников, и особенно — перспектив учёного. В этом контексте данная публикация примыкает к работе [24].

БЛАГОДАРНОСТИ

Благодарю кандидата технических наук К.В. Мионову за идею написать статью, полезные советы в её реализации и помощь при анализе большого объёма информации по этой проблеме в Интернете, докторов технических наук А.Н. Власова и А.И. Новикова из РГРТУ им. В.Ф. Уткина за прочтение рукописи и ценные замечания.

Выражаю благодарность рецензентам за глубокий анализ рукописи и конструктивные замечания, устранение которых, несомненно, способствовало улучшению качества статьи.

ЛИТЕРАТУРА

1. Hirsch J.E. An index to quantify an individual's scientific research output // Proceedings of the National Academy of Sciences of the United States of America. 2005. V. 102. № 46. P. 16569–16572.

2. Google Scholar. <https://scholar.google.com/>
3. Российская научная электронная библиотека eLIBRARY.ru. <https://elibrary.ru/>
4. Astrophysics Data System (ADS NASA). <http://adsabs.harvard.edu/>
5. Bar-Ilan J. Which h-index? A comparison of WoS, Scopus and Google Scholar // *Scientometrics*. 2007. V. 74. № 2. P. 257–271.
6. Рейтинг науковців України за показниками наукометричної бази даних Scopus. 05.12.2013. <https://web.archive.org.ru/>
7. Михайлов О.В., Михайлова Т.И. Индекс Хирша в оценке деятельности учёного в национальном исследовательском университете // *Вестник Казанского технологического университета*. 2010. № 11. С. 485–487.
8. Tagiew R., Ignatov D.I. Behavior mining in *h*-index ranking game // *CEUR Workshop Proceedings*. 2017. V. 1968. P. 52–61.
9. Имаев В. Технологии увеличения индекса Хирша и развитие имитационной науки // *В защиту науки*. 2016. № 17. С. 38–51.
10. Дёмина Н. Хиршемания и хиршефобия // *Троицкий вариант – Наука*. 2016. № 218. С. 6.
11. Михайлов О.В. Новая версия индекса Хирша – *j*-индекс // *Вестник РАН*. 2014. № 6. С. 532–535.
12. Egghe L. Theory and practise of the *g*-index // *Scientometrics*. 2006. V. 69. № 1. P. 131–152.
13. Kosmulski M.I. *I*-a bibliometric index // *Forum Akademickie*. 2006. V. 11. P. 31–32.
14. Prathap G. Hirsch-type indices for ranking institutions' scientific research output // *Current Science journal*. 2006. V. 91(11). P. 1439–1440.
15. Jones T., Huggett S., Kamalski J. Finding a Way Through the Scientific Literature: Indexes and Measures // *World Neurosurgery*. 2011. V. 76. № 1, 2. P. 36–38.
16. Холодов А.С. Об индексах цитирования научных работ // *Вестник РАН*. 2015. № 4. С. 310–320.
17. Мазов Н.А., Гуреев В.Н. Альтернативные подходы к оценке научных результатов // *Вестник РАН*. 2015. № 2. С. 115–122.
18. Кузнецов А.В. Для начала надо навести порядок в существующей системе РИНЦ // *Вестник РАН*. 2014. № 3. С. 268–269.
19. Waltman L., van Eck N.J. Robust Evolutionary Algorithm Design for Socio-Economic Simulation: Some Comments // *Comput. Econ*. 2009. V. 33. P. 103–105.
20. Hirsch J.E. An index to quantify an individual's scientific leadership // *PNAS*. 2005. V. 102. № 46. P. 16569–16572.
21. Вейль Г. О философии математики. М.: ГИТТЛ, 1934.
22. Гейтинг А. Интуиционизм. М.: Мир, 1965.
23. Френкель А.А., Бар-Хиллел И. Основания теории множеств. М.: Мир, 1966.
24. Миронов В.В. Информатизация образования: достижения и проблемы // *Информатизация образования и науки*. 2017. № 4(36). С. 3–18.