

НОВОЕ ПОКОЛЕНИЕ МЕТРИК ЦИТИРОВАНИЯ. ПОСТРОЕНИЕ IF-SCORING RULES

THE NEW GENERATION OF CITATION METRICS. CONSTRUCTION OF IF-SCORING RULES

В.М. Московкин

Белгородский государственный национальный исследовательский университет, Белгород, Россия, moskovkin@bsu.edu.ru

Н.А. Голиков

независимый исследователь, Харьков, Украина

V.M. Moskovkin, National Research Belgorod State University, Belgorod, Russia

N.A. Golikov, Independent researcher, Kharkov, Ukraine

Keywords: citation metrics, h-index, scoring rules, summation-based rankings, Data Mining, Machine Learning, Scopus, SCIMAGO, Google Scholar, impact factor

Исследование выполнено в рамках тематического плана НИУ "БелГУ" на 2013 год, проект № 6.8599.2013

Начиная с классической работы J.E. Hirsch (2005), возник бум по модификации h-index и созданию ему подобных показателей. Как показано в работе (L.Waltman, N.J.van Eck, 2012) в 2011 и 2012 гг. почти каждая четвертая статья, опубликованная в журналах *Scientometrics* и *Journal of Informetrics*, цитировала вышеуказанную работу. В научной литературе, помимо h-index, можно встретить множество ему подобных индексов (m , g , e , w , q^2 , hg и др.) для оценки продуктивности исследований в терминах количества публикаций и цитируемости. Суть всех этих исследований хорошо описал T. Marchant (2009): "Многие исследователи, анализируя ранее предложенные индексы, находят их слабые стороны и предлагают новые их версии или новые индексы. Не удивительно, что аргументация сторонников этих индексов часто надуманна: они предлагают новые индексы, не страдающие

слабостями старых, но это не гарантирует того, что новые индексы не будут иметь другие слабости". Поэтому, как указывает Т. Marchant, необходимо построение теории библиометрического ранжирования.

Ущербность h-index на фундаментальном уровне наглядно показана в работе (L.Waltman, N.J.van Eck, 2012). Оказывается, что этот индекс не удовлетворяет трем простым постулатам:

1. Если два ученых достигают одного и того же относительного улучшения их научной результативности, то их ранжирование (с помощью рангов выраженных h-index) друг относительно друга должно оставаться неизменным.

2. То же самое относительно абсолютного улучшения научной результативности.

3. Если ученый X_1 имеет ранг выше, чем у ученого Y_1 , а ученый X_2 имеет ранг выше, чем у ученого Y_2 , тогда исследовательская группа, состоящая из ученых X_1 и X_2 , должна иметь совокупный ранг выше, чем у исследовательской группы, состоящей из ученых Y_1 и Y_2 .

Все эти недостатки устраняются с помощью построения нового поколения метрик цитирования, основанных на так называемых scoring rules (summation – based rankings). Обобщенный подход для построения таких метрик предложен в работе (Marchant, 2009). В нем учитывается весь спектр публикаций автора или коллектива авторов вместе с количеством цитирований каждой статьи.

В развитие этого подхода мы построили семейство метрик цитирования, основанных исключительно на импакт-факторах журналов, в которых опубликованы статьи автора или коллектива авторов, и на импакт-факторах журналов, статьи из которых цитируют статьи рассматриваемого автора или коллектива авторов.

Нами с помощью алгоритмов распознавания образов и слабого искусственного интеллекта, работающих в системах Data Mining и Machine Learning, разработан специальный алгоритм для расчета предложенного семейства метрик цитирования, идентифицирующий названия журналов в которых опубликованы статьи и определяющий их импакт-факторы, размещенные на платформах SCIMAGO и Scopus. Эксперименты по расчету предложенного семейства метрик цитирования, которые мы назвали IF-scoring rules, проведены для пяти высоко цитируемых ученых Белгородского го-

сударственного национального исследовательского университета с помощью поисковых систем Google Scholar и Scopus как для стандартного импакт-фактора IF, так и для новых его версий – SJR и SNIP. Предлагаемое нами название IF-scoring rules связано с тем, что наше семейство scoring rules строится исключительно на основе импакт-факторов журналов.

Литература

1. Hirsch, J. E. (2005). An index to quality an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569–16572.
2. Watman, L., Van Eck, N. J.(2012). The inconsistency of the h-index. *Journal of American Society for Information Science and Technology*, 63(2), 406–415.
3. Marchant, T. (2009). Score-based bibliometric rankings of authors. *Journal of American Society for Information Science and Technology*, 60(6), 1132–1137.