

ОБ ОСОБЕННОСТЯХ ПРОЕКТА «ВЕБОМЕТРИЧЕСКИЙ РЕЙТИНГ НАУЧНЫХ УЧРЕЖДЕНИЙ РОССИИ»

А.А. Печников, А.А. Огийко

*Институт прикладных математических исследований Карельского научного центра РАН
Петрозаводск*

Важным направлением вебметрических исследований являются задачи ранжирования веб-сайтов университетов и научных учреждений. Основоположниками этого направления справедливо считаются исследователи из испанской группы Cybermetrics Lab [1]. Проект «Ranking Web of Universities» был официально начат Cybermetrics Lab в 2004 году [2]. Популярность этих исследований настолько велика, что в проекте «Национальный рейтинг университетов» значения вебметрических рейтингов веб-сайтов учитываются в качестве одного из индикаторов [3].

В России известны исследования по ранжированию веб-сайтов Сибирского отделения РАН [4], Дальневосточного отделения РАН [5] и формированию российского национального вебметрического индекса научных и образовательных учреждений [6]. Группой по вебметрике Института прикладных математических исследований КарНЦ РАН (ИПМИ КарНЦ РАН) [7] также проведен ряд исследований в этом направлении [8,9]. Исследования [5-9] в основном используют подходы, предложенные Cybermetrics Lab (отметим, что в 2013 г. эти подходы существенно изменились, но об этом стоило бы сделать отдельный доклад). Для построения функций ранжирования используются значения таких индикаторов, как S - количество страниц сайта, R — количество полнотекстовых файлов на сайте, V — количество внешних гиперссылок, сделанных на сайт и Sc — индикатор «научности» сайта (количество ссылок на сайт, обнаруживаемых *Google Scholar*). Первые три индикатора измеряются с помощью наиболее распространенных поисковых машин. Единицами анализа являются официальные сайты организаций, идентифицируемые по их доменным именам.

С конца 2012 года группа по вебметрике ИПМИ КарНЦ РАН начала работу по проекту «Вебметрический рейтинг научных учреждений России», поддержанному РГНФ [10]. Существенным отличием данного проекта от перечисленных ранее является выбор в качестве единицы анализа научного учреждения не только официального сайта, а всех (по возможности) веб-сайтов, составляющих веб-пространство данного учреждения. Поясним сказанное примером. Официальный сайт Института проблем управления РАН имеет доменное имя *www.ipu.ru*, а сайт «Управление большими системами», входящий в веб-пространство института — *ubs.mtas.ru*, и, естественно, не ассоциируется поисковыми системами как поддомен головного сайта, из чего следует, что достаточно большие значения его индикаторов не учитываются при ранжировании в проектах [3,6]. А ведь значение только S для сайта *ubs.mtas.ru* составляет примерно десятую часть сайта *www.ipu.ru*, и таких «неассоциируемых» сайтов журналов Института проблем управления РАН, известных авторам, можно насчитать еще 4: «Проблемы управления» (*pu.mtas.ru*), «Автоматика и телемеханика» (*ait.mtas.ru*), «Датчики и системы» (*www.datsys.ru*) и «Автоматизация в промышленности» (*www.avtprom.ru*).

Формирование списка сайтов научного учреждения, составляющих его веб-пространство, является непростой задачей, однако у группы по вебметрике ИПМИ КарНЦ РАН имеются определенные методы и программное обеспечение, позволяющие (хотя бы частично) её решать. Среди них можно отметить разработанные методы исследования фрагментов регламентируемого Веба [11], программные средства для сканирования сайтов и формируемые базы данных гиперссылок (гостевой вход можно выполнить вот с этой страницы: <http://webometrics-net.ru/section.php?id=27>).

Предполагается также, что указанные методы и программное обеспечение позволят решить и такую задачу, как определение нового индикатора, который предварительно можно назвать индикатором внутренне ссылочной популярности V_{IN} — количество гиперссылок, сделанных на единицу анализа научного учреждения с других единиц анализа научных учреждений, участвующих в ранжировании.

За пределами данных тезисов остались такие важные вопросы, как применение поисковых машин для измерения значений вебметрических индикаторов, автоматизация сбора значений индикаторов, построение функции ранжирования и ряд других. Надеемся, что будет возможность их обсуждения как во время работы секций, так и в кулуарах конференции.

Работа выполняется при поддержке Российского гуманитарного научного фонда (грант № 12-03-12001).

ЛИТЕРАТУРА

1. InternetLab: Observatorio de Ciencia y Tecnología en Internet [сайт]. URL: <http://internetlab.cchs.csic.es> (дата обращения: 20.03.2013).

2. Ranking Web of Universities [сайт]. URL: <http://www.webometrics.info> (дата обращения: 20.03.2013).
3. Методика построения рейтинга классических университетов (2009 год) [Электронный ресурс] / Национальный рейтинг университетов [сайт]. URL: <http://www.univer-rating.ru> (дата обращения: 20.03.2013).
4. Шокин Ю.И., Клименко О.А., Рычкова Е.В., Шабальников И.В. Рейтинг сайтов научных организаций СО РАН // Вычислительные технологии. 2008. Т. 13, № 3. С. 128 – 135.
5. Ханчук А.И., Наумова В.В. Информационное пространство Дальневосточного отделения РАН // Вестник Дальневосточного отделения Российской академии наук. 2009. № 4. С. 122—129.
6. Антопольский А. Б., Поляк Ю. Е. О новом рейтинге веб-сайтов российских университетов // Дистанционное и виртуальное обучение. 2013. № 1. С. 40—47.
7. Вебметрика. Институт прикладных математических исследований КарНЦ РАН [сайт]. URL: <http://webometrics.krc.karelia.ru> (дата обращения: 21.03.2013).
8. Печников А.А., Илюкевич О.Г. Рейтинг официальных web-сайтов университетов России и Финляндии: сравнительный анализ // Информационные ресурсы России. 2008. № 3 (103). С.25—28.
9. Мазалов В.В., Печников А.А. О рейтинге официальных сайтов научных учреждений северо-запада России // Управление большими системами. М.: ИПУ РАН, 2009. Выпуск 24. С.130—146.
10. Вебметрический рейтинг научных учреждений России [сайт]. URL: <http://webometrics-net.ru> (дата обращения 10.04.2013).
11. Печников А.А. Методы исследования регламентируемых тематических фрагментов Web // Труды Института системного анализа Российской академии наук. Серия: Прикладные проблемы управления макросистемами. Том 59. 2010. С. 134—145.
12. Печников А.А., Чернобровкин Д.И. Адаптивный краулер для поиска и сбора внешних гиперссылок // Управление большими системами. Выпуск 36. М.: ИПУ РАН, 2012. С.301—315.