

Изучение мотиваций создания внешних гиперссылок на университетских веб-сайтах: опыт Нигерии

Статья рекомендована А. М. Елизаровым 04.02.2013



Энтони Метумараибе НВОХИРИ

Аспирант факультета прикладной математики — процессов управления Санкт-Петербургского государственного университета



Андрей Анатольевич ПЕЧНИКОВ

Доктор технических наук, ведущий научный сотрудник Института прикладных математических исследований Карельского научного центра РАН

Аннотация.

В работе исследуются официальные сайты нигерийских университетов в целях выяснения мотиваций создания исходящих гиперссылок, сделанных с этих сайтов.

Предложенная классификация мотиваций гиперссылок применена к множеству из более чем 6300 ссылок, сделанных с сайтов 86 нигерийских университетов. Наличие нескольких потенциальных мотиваций приводят авторов к модификации классификационной схемы в сторону объединения как нескольких мотиваций в одну, так и университетов в региональные группы. Применение статистических методов позволяет получить более четкую картинку мотиваций создания внешних гиперссылок на сайтах нигерийских университетов и дать ряд рекомендаций для разработчиков, направленных на улучшение использования внешних гиперссылок.

Ключевые слова:

гиперссылка, университетский веб-сайт, университеты Нигерии, мотивация.

Введение

Несмотря на нынешний интерес к исследованиям гиперссылок в научных исследованиях и многим результатам, указывающим, что гиперссылки связаны с ранжированием результатов и результативностью поиска в Вебе [1–5], о мотивах создания гиперссылок сказано немного [6; 7]. Отсутствие понимания мотиваций создания гиперссылок является одним из главных препятствий формирующегося научного направления «вебометрика», и оно должно быть преодолено, несмотря на его очевидную сложность (см.: [6]).

На практике очень трудно (наверное, и невозможно) оценить и классифицировать веб-коммуникации в целом, что обусловлено как неоднородностью форм публикаций в Вебе, так и техническими проблемами, связанными с получением соответствующей информации, даже без учета размеров Веба.

Некоторые авторы предприняли попытку развить подход, предложенный в [6], в направлении исследования и понимания мотиваций создания гиперссылок применительно к университетам Нигерии. По данным Национальной университетской комиссии Нигерии в стране насчитывается 125 университетов [8]. Список доменных

* Исследования по разработке ИТ-систем для науки и образования и их практическое воплощение проводятся при поддержке Министерства образования и науки Российской Федерации, ГК № 07.514.11.4041.

имен, указанных в [8], являлся неполным и неточным, он был принят лишь в качестве основы и многократно дополнялся и уточнялся в процессе исследования. Так, из рассмотрения были исключены неработающие официальные сайты университетов. Поскольку уточнение списка продолжается, в статье мы остановились на целевом множестве, содержащем 86 доменных имен официальных сайтов университетов Нигерии.

Нигерия состоит из 36 штатов и одной федеральной столичной территории, которые по географическим и этнокультурным признакам составляют пять регионов: Север, Юго-Запад, Юг, Юго-Восток и так называемый Центральный пояс. Из 86 исследуемых сайтов мы проанализировали 13 сайтов с Севера страны, 16 — из Центрального пояса, 28 — с Юго-Запада, 14 — с Юго-Востока и 15 сайтов — с Юга страны.

В статье изложены результаты анализа более чем 6300 исходящих гиперссылок, сделанных с 86 официальных сайтов нигерийских университетов, в целях выяснения мотиваций их создания. Представляется, что авторам удалось обнаружить существенное различие в мотивациях создания внешних гиперссылок на официальных университетских сайтах в зависимости от регионов Нигерии. Это, в свою очередь, позволило сформулировать ряд рекомендаций для разработчиков веб-сайтов, направленных на улучшение использования внешних гиперссылок как одного из основных инструментов коммуникации в Вебе.

Связанные исследования

Анализ гиперссылок в Вебе является новой областью исследований, возможно, лишь внешне напоминающей анализ научного цитирования [9; 10]. «Веб-ссылки являются радикально новым явлением, хотя частичные аналогии могут быть проведены с уже существующими подобными объектами, включая... цитирование в журналах» [6].

Остановимся несколько подробнее на цитируемой статье (см.: [6]), поскольку именно предложенные в ней подходы использованы авторами в качестве основы для исследований. Как инструмент сбора гиперссылок авторы названной статьи использовали краулер, с помощью которого были отсканированы сайты университетов Великобритании. Затем из этой коллекции был произведен случайный выбор гиперссылок при определенных ограничениях и условиях (например, были исключены неработающие гиперссылки и гиперссылки, отсылающие на «собственный» сайт). Кроме того, рассматривались только внешние гиперссылки на сайты, размещенные в домене второго уровня as.uk (так называемый «академический» домен, от англ. «academic»), служащего доменом для сайтов учреждений высшего образования, колледжей, научно-исследовательских учреждений и научных сообществ Великобритании. В целях сосредоточения только на академических сайтах были также исключены гиперссылки на электронные журналы, сайты которых расположены в домене as.uk.

После того как был окончательно составлен рассматриваемый набор гиперссылок (отобрано 414 ссылок, исходящих со 170 академических сайтов),

исследователи независимо друг от друга проанализировали эти гиперссылки, оценили и классифицировали мотивации их создания. Оценка проводилась путем изучения исходных и целевых страниц исследуемых гиперссылок. В итоге исследователи совместно разработали схему классификации, где изначально были выделены 10 возможных мотиваций, по которым и были распределены 414 отобранных гиперссылок. Затем эти 10 мотиваций были обобщены до трех в целях уточнения классификации.

В результате авторы рассматриваемой статьи (см.: [6]) пришли к следующим выводам:

- мотивации для создания ссылок на основе изучения только исходных и целевых страниц трудно классифицировать потому, что (а) существуют различные возможные мотивации, (б) сложно правильно угадать истинные намерения авторов в некоторых случаях и (в) потенциально существуют многократные мотивации; отсюда авторы делают вывод, что лучшая схема классификации могла бы привести к «более надежным результатам»;
- мотивация создания большинства гиперссылок между сайтами британских университетов не является подобной мотивации цитирования журнальных статей в научных публикациях;
- напрямую учебные материалы одного университета очень мало используются в других университетах, несмотря на «обширное финансирование объединенного межучрежденческого обучения», и это, возможно, является показателем того, что финансирование подобных проектов объединенного межуниверситетского обучения неэффективно.

Задачи и методы исследования

Исследование мотиваций создания гиперссылок, сделанных с сайтов нигерийских университетов, подразумевает ответы на следующие два вопроса:

- какие мотивации стоят за размещением гиперссылок на нигерийских университетских сайтах?
 - какие мотивации создания гиперссылок наиболее характерны для этногеографических регионов Нигерии?
Как уже было сказано, нами использовались подходы, предложенные в [6], однако есть существенные отличия:
 - исследуемое множество гиперссылок содержит все внешние гиперссылки, сделанные с сайтов нигерийских университетов, а не только гиперссылки, связывающие эти сайты;
 - предложена существенно иная схема классификации гиперссылок.
- В качестве инструмента для сбора гиперссылок использовался краулер BeeBot [11], с помощью которого было отсканировано 86 сайтов

№	Название мотива	Описание мотивации	Ahmadu Bello University, Zaria (www.abu.edu.ng)	Babcock University, Ilishan-Remo (www.babcockuni.edu.ng)	Ladoke Akintola University of Technology, Ogbomoso (www.lautech.edu.ng)	Kwara State University, Ilorin (www.kwasu.edu.ng)	Nnamdi Azikiwe University, Awka (www.unizik.edu.ng)
1	Учебные материалы для обучающихся	Ссылка на учебные материалы (электронный учебник, лекция, учебный план...)			4		
2	Информация для обучающихся и поступающих	Ссылка на информацию для обучающихся и поступающих (работа, каталоги студенческих ресурсов, новости)	11	2	4		5
3	Ресурсы поддержки научных исследований	Ссылка на ресурсы информационной и финансовой поддержки исследований, фонды, гранты	1				
4	Организация-партнер	Ссылка на веб-ресурс организации, с которой осуществляется совместная работа			2		
5	Библиотеки и электронные журналы	Ссылка на веб-ресурс библиотек и электронных журналов		15	2	2	
6	Публикации сотрудников	Ссылка на опубликованную в Вебе публикацию автора, работающего в организации-владельце сайта	2		1		
7	Подобная по виду деятельности организация	Ссылка на университет, колледж, институт, библиотеку				2	
8	Цитируемая статья	Ссылка на цитируемую статью, размещенную в Вебе				1	
9	Научное мероприятие	Ссылка на веб-ресурс с информацией о проведении научной конференции, семинара					
10	Личные страницы	Ссылка на персональную страницу сотрудника, расположенную на другом веб-ресурсе		2	2	6	
11	Доступ к программному обеспечению	Ссылка на веб-ресурс, предоставляющий возможности загрузки ПО	3	1	1		1
12	Служебный сайт	Закрытые и почтовые порталы	1	1	8	3	4
13	Альтернативный сайт	Ссылка на веб-ресурс организации, не рассматриваемый как официальный					
14	Вышестоящая организация	Ссылка на веб-ресурс вышестоящей организации					
15	Нижестоящая организация	Ссылка на веб-ресурс структурного подразделения данной организации	1	3	3	2	
16	Органы государственной власти	Ссылка на веб-ресурсы органов государственной власти	1				

№	Название мотива	Описание мотивации	Ahmadu Bello University, Zaria (www.abu.edu.ng)	Babcock University, Ilishan-Remo (www.babcockuni.edu.ng)	Ladoke Akintola University of Technology, Ogbomoso (www.lautech.edu.ng)	Kwara State University, Ilorin (www.kwasu.edu.ng)	Nnamdi Azikiwe University, Awka (www.unizik.edu.ng)
17	Сайт разработчиков данного сайта	Ссылка на сайт разработчиков сайта (страницы) данной организации		1		2	
18	Веб-ресурсы спонсоров мероприятий	Ссылка на веб-ресурсы спонсоров мероприятий					
19	Коммерческая организация	Ссылка на веб-ресурс коммерческой организации	12				
20	Новостная лента	Ссылка на новостной веб-ресурс (электронная газета, блог)	1	3	3		7
21	Рекламные ссылки и баннеры	Ссылка рекламного характера на информацию о товарах и услугах			1		
22	Информация об отдыхе и туризме	Ссылка на информацию об отдыхе и туризме			2		
23	Неклассифицированный сайт	Все ссылки, не упомянутые выше	2		3	5	1
24	Неработающая ссылка	псевдо-мотивация	1	6	9	3	1
		Общее количество гиперссылок с сайта	36	34	45	26	19

Табл. 1. Основные мотивы для создания гиперссылок

университетов Нигерии, на которых было обнаружено более 6300 исходящих гиперссылок.

Из 86 исследованных сайтов наибольшее количество исходящих гиперссылок имеет сайт University of Agriculture, Abeokuta (www.unaab.edu.ng) — 1027. В то же время восемь сайтов исследуемых университетов имеют всего по одной исходящей гиперссылке (далее они были исключены из рассмотрения). Среднее количество исходящих гиперссылок, размещенных на сайтах 86 исследуемых университетов, равно 74.

Классификационная схема

Изначально были выделены 24 основные мотивации, приведенные в табл. 1.

Последнюю мотивацию правильнее назвать «псевдомотивацией», поскольку в этой группе фиксировалось количество неработающих гиперссылок. Из-за невозможности указать значения по всем 86 университетам

приведены значения только по 5 из них (по 1 университету от каждого этнокультурного региона Нигерии).

Из 6358 исследованных ссылок авторы расходились во мнениях по их классификации в 699 случаях. Это несогласие зачастую было обусловлено существованием частичного совпадения между определенными мотивациями. Приведем несколько примеров:

- ссылку (зачастую оформленную как баннер), отсылающую к товарам или услугам, предоставляемым компанией, можно интерпретировать как ссылку на информацию о товарах, услугах (мотивация «Рекламные ссылки»), либо как ссылку на информацию об отдыхе и туризме (мотивация «Информация об отдыхе и туризме»), либо даже как ссылку на веб-ресурс организации, для которой коммерческая деятельность является основной (мотивация «Коммерческая организация»);
- ссылка со страницы сайта библиотеки одного университета на книгу, находящуюся на странице сайта библиотеки другого университета, может быть классифицирована как ссылка на подобную организацию (мотивация «Подобная по виду деятельности организация»), либо как ссылка на учебные материалы для обучающихся (мотивация «Учебные материалы для обучающихся») или как ссылка на ресурсы поддержки исследований (мотивация «Ресурсы поддержки научных исследований»);
- ссылка с сайта одного университета на сайт другого университета, выступающего как один из спонсоров конференции, проходящей в первом университете, может быть рассмотрена как ссылка на веб-ресурсы спонсоров мероприятий (мотивация «Веб-ресурсы спонсоров мероприятий») или как ссылка на подобную организацию (мотивация «Подобная по деятельности организация»).

Вследствие этого классификационная схема была упрощена путем объединения строк табл. 1 с сокращением количества мотиваций с 24 до 9. То же самое было проделано и со столбцами классификационной схемы. При объединении столбцов все исследуемые университеты были собраны по регионам в пять указанных групп с соответствующим суммированием количества гиперссылок. Результаты указанных действий приведены в табл. 2.

В табл. 2 мотивы «Учебные материалы для обучающихся», «Информации для обучающихся и поступающих», «Неклассифицированный сайт» и «Неработающая ссылка» остаются такими же, как и в табл. 1.

Мотив «Поддержка научных исследований» в табл. 2 является объединением мотивов «Ресурсы поддержки научных исследований», «Публикации сотрудников», «Библиотеки и электронные журналы» и «Цитируемая статья» из табл. 1. Мотив «Вышестоящая организация» является объединением мотивов «Вышестоящая организация» и «Органы государственной власти». Далее, мотив «Подобная по деятельности организация и нижестоящая организация» является объединением следующих мотивов: «Подобная по виду деятельности организация», «Нижестоящая организация», «Организация-партнер»

и «Альтернативный сайт». Мотив «Служебный сайт» является объединением мотивов «Служебный сайт», «Личные страницы», «Новостная лента», «Научное мероприятие». Наконец, мотив «Организации сферы бизнеса» объединяет мотивы «Доступ к программному обеспечению», «Веб-ресурсы спонсоров мероприятий», «Коммерческая организация», «Рекламные ссылки и баннеры», «Информацию об отдыхе и туризме», «Сайт разработчиков данного сайта».

И без статистической обработки данных табл. 2 можно заметить, что на Юго-Востоке почти не обращают внимания на ссылки на учебные материалы для обучающихся, на Севере большое внимание уделяется поддержке научных исследований, а на Юго-Западе много неклассифицированных гиперссылок. Отметим, однако, что мы анализируем неравные по количеству университетов регионы. В Юго-Западном регионе имеется наибольшее количество университетов, и, следовательно, потенциально значения величин в табл. 2 в столбце «Юго-Запад» должны быть большими. Отметим также, что ссылки, имеющие мотивацию «Поддержка научных исследований», занимают 37% от общего количество исследуемых ссылок, а ссылок с мотивацией «Вышестоящая организация» всего лишь 2%.

Для применения статистических методов вместо табл. 2 была использована табл. 3, в которой количество университетов по каждому региону равно 13 (минимальное значение у Северного региона). Университеты в каждой группе отбирались методом случайного отбора.

Проверка статистической гипотезы о независимости полученных данных с использованием критерия Пирсона [12] показала, что есть существенная разница между наборами данных (ожидаемыми и наблюдаемыми частотами), которая не объясняется только случайными причинами.

№	Название мотива	Центральный пояс					ВСЕГО
		Север	Центральный пояс	Юго-Запад	Юго-Восток	Юг	
1	Учебные материалы для обучающихся	7	8	76	2	28	121
2	Информации для обучающихся и поступающих	14	20	254	129	26	443
3	Поддержка научных исследований	630	719	942	50	21	2362
4	Вышестоящая организация	6	47	28	2	29	112
5	Подобная по деятельности организация и нижестоящая организация	22	49	240	10	52	373
6	Служебный сайт	62	107	755	175	177	1276
7	Организации сферы бизнеса	27	66	525	18	42	678
8	Неклассифицированный сайт	20	74	524	57	24	699
9	Неработающая ссылка	33	47	176	6	32	294
	Общее количество исследуемых гиперссылок	821	1137	3520	449	431	6358

Табл. 2. Обобщенная таблица мотивов и частот

Обсуждение и выводы

Не вдаваясь в математические детали, можно сказать, что для каждого региона можно выделить доминирующую мотивацию создания внешних гиперссылок. В Северном и Юго-Западном регионах и Центральном поясе такой доминантой является поддержка научных исследований, а в Южном и Юго-Восточном регионах — отсылка к служебным сайтам. Отсюда складывается впечатление, что сотрудники этих университетов пока еще не привыкли пользоваться почтовыми сервисами и информационными системами, поэтому их стимулируют к такой деятельности.

При этом обращает на себя внимание большое количество гиперссылок, размещаемых на сайтах университетов Юго-Востока и отсылающих к информации для студентов и абитуриентов, носящей бытовой, а не учебный характер.

Подводя итоги, можно сказать, что между университетами имеются немалые различия по регионам в создании гиперссылок. На сегодняшний день невозможно дать содержательное объяснение обнаруженным различиям. Может быть, на Юге вопрос с финансовой поддержкой научных исследований полностью решен, а на Севере наблюдается плохое финансирование, а может быть, наоборот, — на Севере научные исследования ведутся очень интенсивно, а на Юге они полностью отсутствуют. Но при этом результаты исследования позволяют дать ряд рекомендаций разработчикам нигерийских университетских сайтов:

№	Название мотива	Центральный пояс				
		Север	Центральный пояс	Юго-Запад	Юго-Восток	Юг
1	Учебные материалы для обучающихся	7	3	6	2	16
2	Информации для обучающихся и поступающих	14	20	35	129	14
3	Поддержка научных исследований	630	248	466	50	20
4	Вышестоящая организация	6	46	4	2	28
5	Подобная по деятельности организация и нижестоящая организация	22	43	42	10	22
6	Служебный сайт	62	23	122	175	76
7	Организации сферы бизнеса	27	48	91	17	24
8	Неклассифицированный сайт	20	66	10	55	17
9	Неработающая ссылка	33	14	42	6	19
	Общее количество исследуемых гиперссылок	821	511	818	446	236
						2832

Табл. 3. Рандомизированная обобщенная таблица мотивов и частот

- во всех регионах гораздо большее внимание должно быть уделено гиперссылкам на учебные материалы (может быть, проблема еще более глубокая, заключающаяся в малом количестве таких материалов, и, как следствие, — в отсутствии гиперссылок на них);
- следует обратить более пристальное внимание на гиперссылки, сделанные на сайты других учебных организаций и их подразделений (слабая связность веб-пространства университетов Нигерии уже отмечалась в [13]);
- неработающие ссылки (объединенные в специально введенную группу как псевдо-мотивация) не оказывают практически никакого влияния на общий характер поведения разработчиков сайтов, но все-таки с ними надо бороться.

ЛИТЕРАТУРА

1. THELWALL M. **Results from a Web Impact Factor Crawler** // Journal of Documentation. 2001. № 57 (2). Pp. 177–191.
2. THOMAS O., WILLETT P. **Webometric analysis of departments of librarianship and information science** // Journal of Information Science. 2000. № 26 (6). Pp. 421–428.
3. CHU H., HE S., THELWALL M. **Library and information science schools in Canada and USA: a Webometric perspective** // Journal of Education for Library and Information Science. 2002. № 3 (2). Pp. 110–125.
4. SMITH A., THELWALL M. **Web Impact Factors for Australasian Universities** // Scientometrics. 2002. № 54 (1–2). Pp. 363–380.
5. TANG R., THELWALL M. **Exploring the pattern of links between Chinese university web sites** // Proceedings of the ASIST Annual Meeting. 2002. Pp. 417–424.
6. WILKINSON D., HARRIES G., THELWALL M., PRICE L. **Motivations for academic web site interlinking: evidence for the Web as a novel source of information on informal scholarly communication** // Journal of Information Science. 2003. № 29 (1). Pp. 49–56.
7. ПЕЧНИКОВ А. А., ЛУГОВАЯ Н. Б., ЧУЙКО Ю. В., КОСИНЕЦ И. Э. **Разработка инструментов для вебметрических исследований гиперссылок научных сайтов** // Вычислительные технологии. 2009. Том 14. № 5. С. 66–78.
8. **National Universities Commission** // URL: <http://www.nuc.edu.ng/pages/universities.asp> (дата обращения 10.01.2013).
9. WHITE M. D., WANG P. **A qualitative study of citing behavior: contributions, criteria, and metalevel documentation concerns** // Library Quarterly. 1997. № 67 (2). Pp. 122–154.
10. OPPENHEIM C., RENN S. P. **Highly cited old papers and reasons why they continue to be cited** // Journal of the American Society for Information Science. 1978. № 29 (5). Pp. 225–231.
11. ПЕЧНИКОВ А. А., ЧЕРНОБРОВКИН Д. И. **Адаптивный краулер для поиска и сбора внешних гиперссылок** // Управление большими системами. 2012. Вып. 36. С. 301–315.
12. БУРЕ В. М. **Методология статистического анализа опытных данных**. СПб: Изд. СПбГУ, 2007. 141 с.
13. PECHNIKOV A., NWOHIRI A. **Webometric analysis of Nigerian university websites** // Webology. 2012. № 9 (1) // URL: <http://www.webology.org/2012/v9n1/a95.html>.