

УДК 004.7

## **ВЕБ-РЕСУРСЫ РОССИЙСКОГО ВУЗА: САМООРГАНИЗАЦИЯ ИЛИ АДМИНИСТРАТИВНОЕ ВОЗДЕЙСТВИЕ? ИССЛЕДОВАНИЕ НА ПРИМЕРЕ САНКТ-ПЕТЕРБУРГСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА**

**А.А. Печников**

*Институт прикладных математических исследований  
Карельского научного центра РАН  
pechnikov@krc.karelia.ru*

### ***Аннотация***

Веб-пространство организации – это множество веб-сайтов организации, связанных посредством гиперссылок. Основной вопрос, рассматриваемый в статье, заключается в том, является ли такое веб-пространство самоорганизующимся, то есть происходит ли упорядочение его элементов за счёт их внутреннего взаимодействия без внешних воздействий или же внешние (так называемые «административные») воздействия настолько сильны, что их влияние можно обнаружить. В статье предлагается общий подход, демонстрируемый на примере веб-пространства Санкт-Петербургского государственного университета. На основе проведенного анализа на вопрос, являются ли значимыми административные воздействия на веб-пространство университета, в данном случае дается положительный ответ.

**Ключевые слова:** *гиперссылка, веб-сайт, веб-граф, динамическая модель, административное воздействие*

### **ВВЕДЕНИЕ**

Веб-пространство организации (университета, предприятия, научного института) – это множество веб-сайтов организации, связанных посредством гиперс-

сылок. Как правило, в таком множестве выделяется т. н. «головной сайт» (официальный сайт организации), а также сайты подразделений, проектов, различных мероприятий, форумы, вики-ресурсы и т. д.

Веб-граф является общепризнанной моделью веб-пространства для исследования его основных теоретико-графовых свойств с целью их содержательной интерпретации и выработки управленческих решений, направленных на улучшение характеристик присутствия в Вебе. Определения веб-графа могут варьироваться в зависимости от целей, стоящих в конкретном исследовании. В данной работе мы определим веб-граф как граф, у которого вершинами служат веб-сайты, а ребра соединяют те вершины, между которыми имеются гиперссылки.

Под гиперссылками в данном случае мы понимаем не все ссылки между сайтами. На различных страницах одного сайта могут встречаться гиперссылки на один и тот же внешний адрес, имеющие одинаковый контекст (в частном случае – анкор), и количество таких «одинаковых» гиперссылок может быть равно количеству страниц на сайте (например – ссылка на сайт вышестоящей организации). Из такого множества гиперссылок, имеющих одинаковый адрес-приёмник и контекст, сделанных с данного сайта, в нашем исследовании мы рассматриваем только одну, – ту, которая находится на странице, имеющей максимальный уровень (наивысшим считается уровень начальной страницы сайта). Таким образом, здесь веб-граф – это ориентированный граф с кратными дугами без петель.

Методы построения веб-графа информационного веб-пространства университета на примере Санкт-Петербургского государственного университета (СПбГУ) детально описаны в [1], и здесь мы будем использовать полученные ранее результаты.

Регламентируемыми (администрируемыми) веб-ресурсами будем называть веб-ресурсы, для которых существует (или, точнее сказать, должен существовать) официальный документ, в котором изложены цели и задачи веб-ресурса, основные структурные составляющие, правила обновления, добавления и изменения информации и т. д. [2]. В частности, официальные сайты вузов следует отнести к регламентируемым веб-ресурсам.

Под управляемостью процессами в заданном фрагменте Веба понимается оформление управленческих решений в виде официальных документов, определяющих их реализацию и дальнейшее исполнение. Такие официальные документы будем называть административными воздействиями по управлению фрагментом Веба (или просто административными воздействиями).

В данной работе рассматриваются следующие вопросы:

- методика построения веб-пространства вуза и веб-графа как его модели;
- сравнение веб-графа вуза и веб-графа Барабаши–Альберт с целью установления факта, является ли веб-пространство вуза самоорганизующейся системой;
- анализ изменений веб-графа университета во времени с выявлением значимых процессов и явлений (таких, как появление большого количества ссылок и/или сайтов за ограниченный период времени);
- обнаружение в веб-пространстве университета документов, свидетельствующих об управленческих решениях, направленных на реализацию и исполнение решений относительно университетских веб-ресурсов (постановлений, инструкций, решений, приказов);
- сопоставление значимых изменений веб-пространства и управленческих решений во времени.

Последовательное рассмотрение этих вопросов позволяет ответить на основной вопрос данной работы, а именно, насколько значимыми являются административные воздействия на изменения, происходящие в веб-пространстве университета.

Представляется, что на примере СПбГУ удалось получить положительный ответ на данный вопрос, что, 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 актуальной областью исследований последние двадцать лет, и актуальность исследований бесспорна, поскольку гиперссылки являются единственным способом взаимодействия между сайтами. В книге [3] всесторонне рассматривается комплекс проблем анализа гиперссылок на конкретных примерах, включая исследования университетского веб-пространства Великобритании. В данном кратком обзоре мы не будем останавливаться на большом массиве работ, посвященных теоретическим аспектам данной проблемы, таким, как модели Эрдеша–Реньи, Барабаши–Альберт, Боллобаша–Риордана [4], тем более, что модель Барабаши–Альберт будет несколько подробнее описана далее, а сделаем акцент на работах, посвященных реальным фрагментам Веба.

Одна из наиболее известных сегодня классических моделей так называемого Большого Веба, построенная с использованием теоретико-графового подхода, – это модель «галстук-бабочка» [5]. В дальнейшем модель применялась не только для всего Веба (что вскоре стало трудно осуществимым), но и для анализа его фрагментов (например, [6]).

Веб-граф как модель фрагмента веб-пространства на некотором заданном множестве сайтов является распространенной теоретико-графовой моделью, с успехом используемой в течение последних 10–15 лет. На эту тему написано большое количество работ, исследующих разные аспекты; мы ограничимся тремя ссылками [7–9]. Более 10 лет назад начались исследования по изучению самоорганизации веб-сообществ, представляющие большой научный и практический интерес [10].

Из российских работ, наиболее близких к содержанию данной статьи, можно выделить работу [11], в которой на примере веб-графа Сибирского отделения РАН моделируются сценарии присоединения к этому графу новых вершин и исследуются происходящие в результате этого изменения рангов вершин, вычисляемых алгоритмом PageRank.

Веб-пространства крупных университетов мира на сегодня становятся достаточно большими объектами, обладающими собственной спецификой. В отличие от описанных выше исследований, в данной статье предлагаются подходы к

исследованию веб-пространства одного вуза (на примере СПбГУ), а не группы вузов, как ранее. При этом веб-граф рассматривается в «динамике» с изменением по годам, а не как статический граф на некоторый момент времени. Предпринимается попытка использования информации об управляющих воздействиях, обнаруживаемой на головном сайте исследуемого веб-пространства. Подчеркнем новизну подходов, реализованных в данной работе. Сама постановка вопроса о проверке того, насколько значимыми являются административные воздействия на изменения, происходящие в веб-пространстве университета, представляется новой.

### **ВЕБ-ГРАФ СПБГУ**

В соответствии с [1] кратко опишем подход к построению множества веб-сайтов информационного пространства университета. Вначале сканируется головной сайт вуза, и все полученные внешние гиперссылки сохраняются в специализированной базе данных. Головному сайту присваивается номер уровня, равный 0. Далее проводится анализ внешних гиперссылок, сделанных с головного сайта (в данном случае это официальный сайт СПбГУ [spbu.ru](http://spbu.ru)), на предмет поиска:

а) доменных имен, аффилированных с доменом головного сайта (для [spbu.ru](http://spbu.ru), например, [it.spbu.ru](http://it.spbu.ru)),

б) доменных имен, не аффилированных с доменом головного сайта, но содержащих вхождения подстрок домена головного сайта (например, для [spbu.ru](http://spbu.ru) это подстроки 'spbu' и обнаруженный сайт [www.spbumag.nw.ru](http://www.spbumag.nw.ru)) с дальнейшей содержательной проверкой сайтов,

в) доменных имен, не аффилированных с доменом головного сайта и не содержащих вхождения соответствующих подстрок, на которые имеются ссылки с головного сайта, и обнаруживаемые прямым просмотром анкоров гиперссылок (что является ненадежным поиском сайтов).

В нашем случае для СПбГУ получили множество из 97 доменных имен сайтов, которым присвоен номер уровня, равный 1. Далее проводится анализ внешних гиперссылок, сделанных с сайтов уровня 1, по тем же признакам а)–в). Для СПбГУ получили множество из 162 доменных имен сайтов уровня 2. Аналогично строятся множество из 25 доменных имен сайтов уровня 3 и множество из одного

доменного имени сайта уровня 4. На уровне 5 сайты отсутствуют. Полученное множество доменных имен веб-сайтов принимается в качестве множества вершин, а соединяющие эти вершины гиперссылки принимаются в качестве дуг веб-графа.

Веб-граф университета представляет собой  $G=G(V,E)$ , где  $V$  – множество вершин, соответствующих сайтам, а  $E$  – множество дуг, то есть пар вершин  $(v,w) \in V$ ; дуга  $(v,w)$  существует, если существует гиперссылка с сайта  $v$  на сайт  $w$ . По построению граф  $G=G(V, E)$  является ориентированным графом с кратными дугами без петель. Кратные дуги возникают в том случае, когда с сайта  $v$  на сайт  $w$  сделано более одной гиперссылки. Построенный веб-граф СПбГУ содержит 286 вершин и 19394 дуги (с учетом кратных дуг).

### **СРАВНЕНИЕ С МОДЕЛЬЮ БАРАБАШИ–АЛЬБЕРТ**

Под самоорганизацией веб-пространства как системы понимается упорядочение его элементов за счёт их внутреннего взаимодействия без внешних воздействий. А.-Л. Барабаши и Р. Альберт [12] предложили одну из известных моделей случайного графа, основанную на принципе предпочтительного присоединения, когда с нового сайта преимущественно делаются ссылки на те сайты, на которые уже имеется много ссылок. Предпочтительное присоединение может рассматриваться как внутреннее взаимодействие. Поэтому, рассматривая граф Барабаши–Альберт в качестве одной из возможных моделей самоорганизации веб-пространства, будем полагать, что если граф веб-пространства обладает свойствами графа Барабаши–Альберт, то можно предположить, что и веб-пространство обладает свойством самоорганизации, то есть развивается без внешних воздействий.

Ряд важных особенностей веб-графа Барабаши–Альберт изложим, следуя работе [4]. Во-первых, веб-граф – это весьма разреженный граф, у которого на  $n$  вершинах примерно  $k \cdot n$  ребер, где  $k \geq 1$  – некоторая константа. Во-вторых, диаметр веб-графа исключительно скромный. «Кликакая» по ссылкам, можно с любого сайта на любой другой перейти за 5–7 нажатий клавиши компьютерной мыши. Несколько правильнее сказать, что в веб-графе есть гигантская компонента (сильной связности), и уже ее диаметр невелик. Таким образом, веб-граф очень специфичен: будучи разреженным, он, тем не менее, в известном смысле тесен. В-тре-

тых, у веб-графа весьма характерное распределение степеней вершин. Эмпирическая вероятность того, что вершина веб-графа имеет степень  $d$ , оценивается как  $c/d^\lambda$ , где  $\lambda \approx 2.1$ , а  $c$  – нормирующий множитель, вычисляемый из условия «сумма вероятностей равна 1».

Для веб-графа СПбГУ при  $n=286$  вершин имеем 19394 дуги, т. е.  $k=67.8$ , тогда как в максимальном случае полного графа (даже без кратных дуг) их должно быть более 80000, т. е.  $k$  потенциально может увеличиться еще в 4 раза. Если взять веб-граф СПбГУ без кратных дуг, то количество дуг оказывается равным всего 1717, плотность графа равна 0.022, то есть свойство разреженности графа становится очевидным.

Компонента сильной связности (КСС) веб-графа СПбГУ достаточно велика, она содержит 218 вершин, и её диаметр равен 6. Отметим, что потенциально КСС имеет все предпосылки к росту, поскольку, если рассмотреть этот же веб-граф, но с неориентированными дугами, его компонента связности содержит уже все вершины (и её диаметр равен 5).

Таким образом, для веб-графа СПбГУ мы имеем практически полное выполнение первых двух свойств, характерных для модели Барабаши–Альберт: разреженный граф с КСС диаметром 6.

Более сложно обстоит дело с распределением степеней вершин. Используя метод идентификации степенного закона распределения, предложенный М. Ньюменом [13], получаем «теоретическое» степенное распределение степеней вершин в графе СПбГУ вида  $p(d)=0.39/d^{1.39}$ , где  $d$  – степень вершины ( $d$  – натуральное число). Очевидно, что показатель степени  $\lambda$  в данном случае гораздо меньше 2.1. Покажем на рисунке 1 различие данного распределения и значений частот, вычисленных на основе реальных данных (показаны значения  $d$  от 1 до 50).

Не погружаясь в математические подробности, можно сделать вывод о том, что степенной закон распределения степеней вершин в случае веб-графа СПбГУ не выполняется. Таким образом, веб-граф СПбГУ нельзя считать случайным веб-графом Барабаши–Альберт, а это значит, что предположение о его самоорганизации можно отвергнуть.

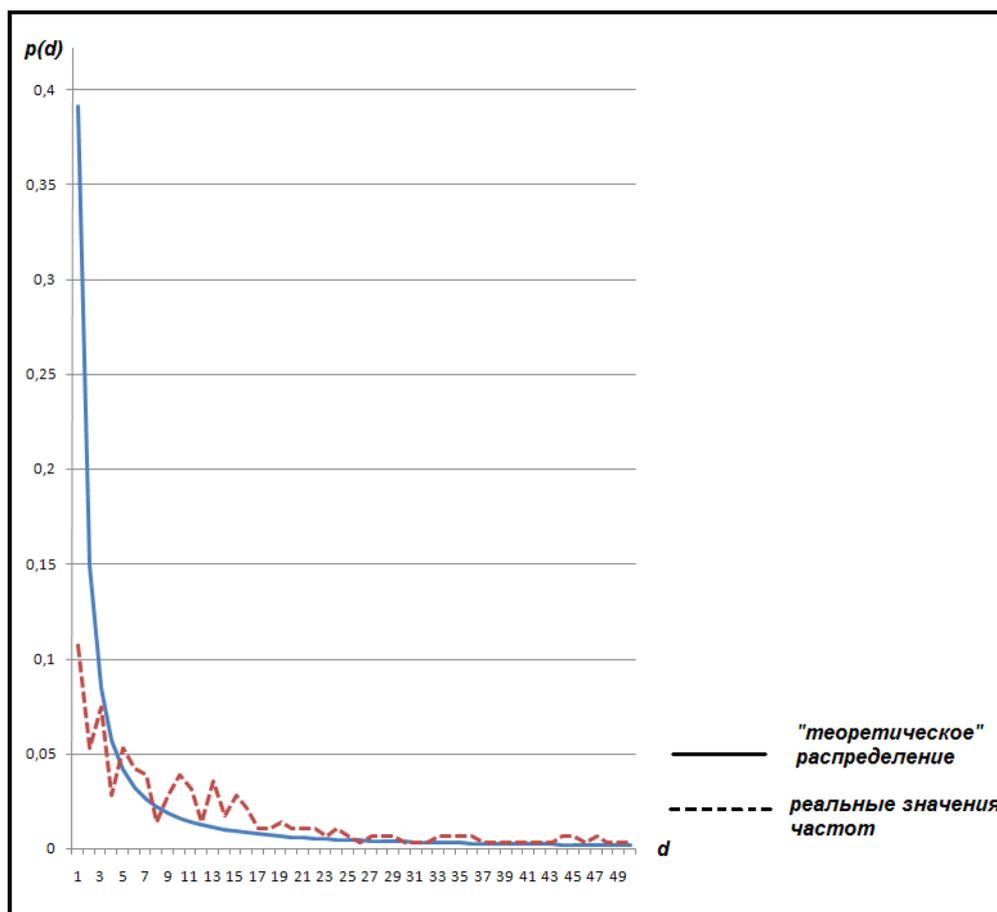


Рис. 1. Сравнение степенного закона и реальных частот распределения степеней вершин

### ДИНАМИЧЕСКАЯ МОДЕЛЬ ВЕБ-ГРАФА УНИВЕРСИТЕТА

До сих пор мы рассматривали  $G=G(V,E)$  как статический веб-граф, сформированный на некоторый момент времени (в нашем случае – на конец 2015 года). На самом деле все несколько сложнее, поскольку сканирование сайтов занимает определенное время, – на сканирование сайтов СПбГУ понадобилось около месяца, – и уже за это время могли появиться или исчезнуть сайты, а на отсканированных сайтах могли исчезнуть некоторые гиперссылки и появиться новые.

Для построения динамической модели веб-графа необходимо знать, когда появился тот или иной сайт, а также когда появилась та или иная гиперссылка. Например, зная датировки с точностью до года, мы могли бы, исключая вершины и/или дуги в соответствии с их датировками (начиная с 2015 года в обратном порядке), получить последовательность веб-графов, представляющих динамиче-

скую модель, развернутую в дискретном времени с шагом в 1 год. Датировка сайтов – решаемая задача, а вот датировка гиперссылок – нет, и поэтому потребует определенных допущений.

Датировка сайтов проводилась с помощью веб-ресурса Internet Archive: Wayback Machine [14], позволяющего по доменному имени сайта увидеть копии некоторых его страниц (включая начальную), собранные специализированным роботом и сохраненные в архиве с указанием даты сохранения с точностью до дня. В случае с СПбГУ для датировки сайтов мы использовали год первого появления копии сайта в архиве.

Таким способом удалось датировать 247 сайтов СПбГУ из 286. Еще 33 сайта, данные о которых отсутствуют в Internet Archive, были датированы по прямым или косвенным признакам, таким, как дата проведения конференции (для сайта конференции), дата первой новости, появившейся на сайте (для сайта кафедры) или годы выполнения проекта (для сайта проекта). Не удалось датировать лишь 6 сайтов, которые далее не рассматриваются (заметим, что это не самые важные сайты веб-пространства университета). Таким образом, мы получили веб-граф СПбГУ на 2015 год с датировками возникновения сайтов-вершин, содержащий 280 вершин и 9697 дуг.

Если из веб-графа 2015 года удалить все вершины, соответствующие веб-сайтам, датированным 2015 годом, то, очевидно, получим множество вершин веб-графа на 2014 год. Для веб-графа СПбГУ их будет 271. При этом количество дуг веб-графа СПбГУ уменьшится до 9400. Среди этих дуг могут быть как те, которые соответствуют гиперссылкам, появившимся в 2014 году (и ранее), так и те, которые появились позже, уже в 2015 году. К сожалению, возможностей датировки появления гиперссылок мы не имеем, значит, не можем удалить ссылки, которые относятся к 2015 году. Если бы нам это удалось, то мы без особого труда получили бы последовательность веб-графов СПбГУ, изменяющуюся во времени с временным шагом 1 год.

Поэтому вместо графа с кратными дугами будем рассматривать граф с одинарными дугами (то есть дуги любой кратности заменим единственной дугой). Что еще дает такая замена графа с кратными дугами графом без кратных дуг? Она дает возможность предложить похожую на реальность (не более того) гипотезу о

том, что дуги в веб-графе появляются вместе с вершинами (в один и тот же год). Поэтому, когда мы удаляем вершины, датированные некоторым годом, то удаляем и все дуги, (как бы) появившиеся вместе с ними.

При такой замене мы теряем характеристики веб-графа с кратными дугами, такие, как Page Rank и модулярность (точнее – их значения изменятся при такой замене), но сохраним не менее важные – связность, диаметр и среднюю длину пути.

Датировка веб-сайтов СПбГУ говорит нам о том, что первым сайтом был официальный сайт СПбГУ (spbu.ru), который датируется 1996 годом. Пусть  $G_{year}(V_{year}, E_{year})$  – веб-граф СПбГУ на год 'year' без кратных дуг. Тогда  $G_{1996} = G_{1996}(V_{1996}, E_{1996})$ , где  $V_{1996} = \{spbu.ru\}$ , а  $E_{1996} = \{\emptyset\}$ . В 1997 году веб-граф не изменился, поскольку не появилось новых сайтов, а в 1998 году появилось еще 7 сайтов, т. е. в  $G_{1998}(V_{1998}, E_{1998})$  имеем  $|V_{1998}| = 8$ , а  $|E_{1998}| = 17$ .

На рис. 2 приведены веб-графы СПбГУ, позволяющие наглядно представить изменения веб-пространства университета с 2008 по 2014 годы с двухгодичным интервалом. Из рисунка очевидно нарастание мощности множеств  $V$  и  $E$ , а также улучшение связности со временем.

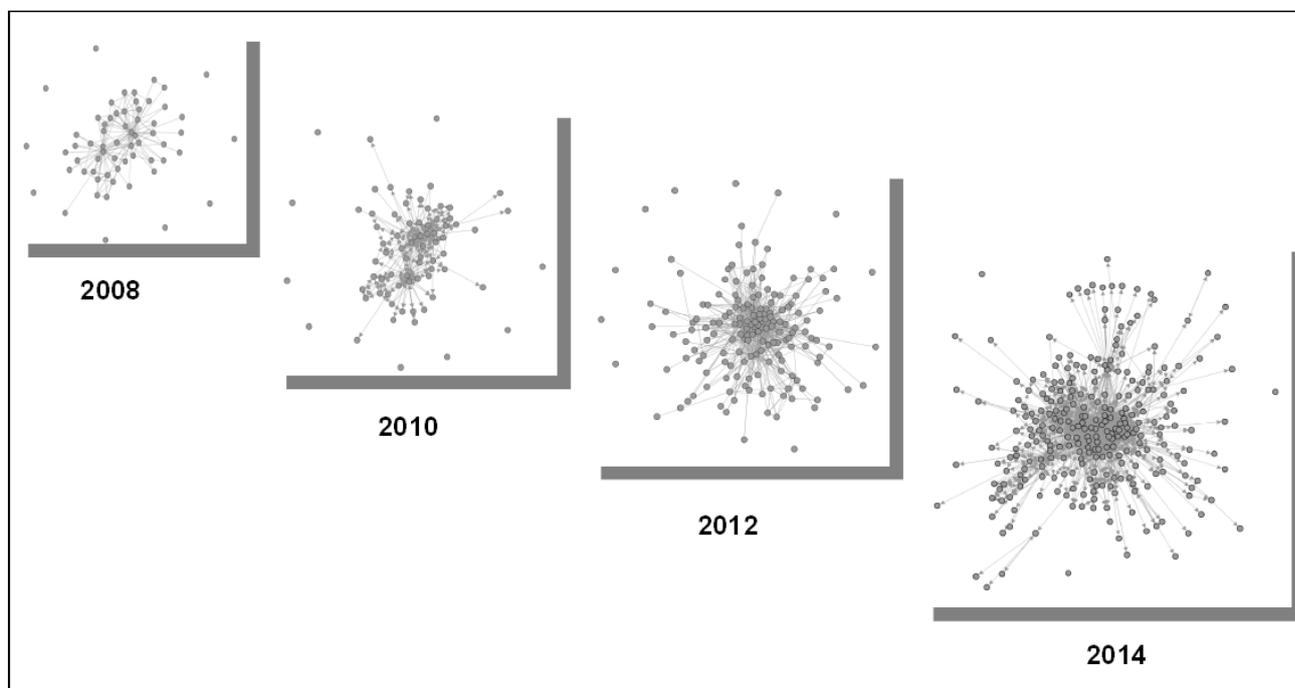


Рис. 2. Динамика веб-графа СПбГУ

Более точно изменение ряда характеристик веб-графов с 1996 по 2015 годы можно увидеть в приведенной таблице 1.

Таблица 1. Характеристики веб-графов с 1996 по 2015 годы

год	количество вершин	количество дуг	диаметр	связные компоненты
1996	1	0	0	0
1997	1	0	0	0
1998	8	17	4	1
1999	11	26	4	1
2000	13	31	4	1
2001	27	98	4	1
2002	30	105	4	2
2003	36	122	5	2
2004	41	140	5	3
2005	42	143	5	3
2006	50	174	7	5
2007	56	199	7	6
2008	62	208	7	10
2009	67	220	7	12
2010	74	291	5	10
2011	112	663	6	10
2012	165	991	6	10
2013	245	1428	6	6
2014	271	1623	6	4
2015	280	1717	6	1

Наглядное представление об изменении количества дуг и вершин веб-графа СПбГУ с 1996 по 2015 годы дают графики на рисунке 3. На верхнем графике показаны абсолютные значения, на нижнем – прирост количества дуг и вершин за год.

Очевиден резкий прирост количества дуг и вершин в период с 2011 по 2013 годы, например, количество дуг в веб-графе в 2010 году было равно 291, а в 2012 – уже 663. Наибольший прирост вершин произошел в 2013 году – добавилось сразу 80 веб-сайтов.

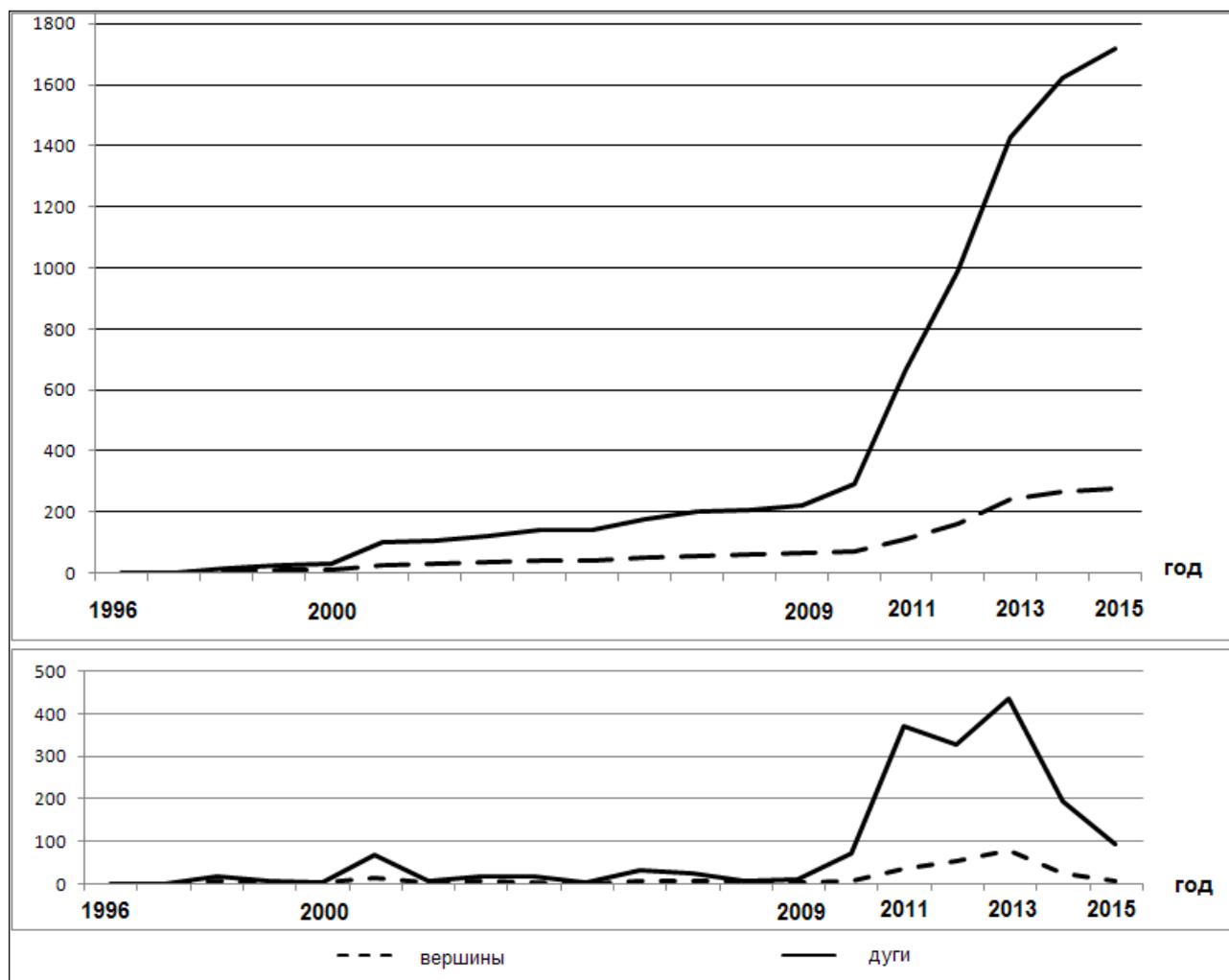


Рис. 3. Изменение количества дуг и вершин в веб-графе СПбГУ

Такие резкие скачки требуют рациональных объяснений, которые мы попробуем дать с точки зрения административных воздействий.

### АДМИНИСТРАТИВНЫЕ ВОЗДЕЙСТВИЯ

Для обнаружения административных воздействий на веб-пространство СПбГУ с использованием Яндекса и Google был осуществлен поиск документов на сайтах домена spbu.ru, содержащих словосочетания «единое информационное пространство», «веб-ресурсы университета», «официальный сайт университета», «официальный портал», «приказ», «положение» и их различные комбинации. Наиболее значимым результатом поиска в контексте данного исследования представляется приказ первого проректора И.А. Горлинского №665 от 17.03.2011 г. «Об

унификации доменных имен веб-ресурсов и повышения рейтинга Санкт-Петербургского государственного университета в информационном пространстве» ([http://www.bio.spbu.ru/news/pdf/prikaz\\_665\\_1\\_17\\_03\\_2011.pdf](http://www.bio.spbu.ru/news/pdf/prikaz_665_1_17_03_2011.pdf)). В приказе говорится: «... регистрацию всех веб-ресурсов Университета производить только в официальном международном домене “spbu.ru”». Обратим внимание на то, что год опубликования приказов совпадают с первым пиком нижнего графика на рис. 3.

Из 38 сайтов, появление которых датировано нами 2011 годом, половина – это 18 сайтов институтов и факультетов и 1 сайт научной библиотеки. Естественно, напрашивается вопрос, неужели раньше эти факультеты и библиотека не имели собственных сайтов? Ответ достаточно прост: до 2011 года их сайты имели другие доменные имена, например, сайт юридического факультета имел имя [jurfak.spb.ru](http://jurfak.spb.ru), а с 2011 года [law.spbu.ru](http://law.spbu.ru), сайт научной библиотеки – [lib.spbu.ru](http://lib.spbu.ru), с 2011 года [library.spbu.ru](http://library.spbu.ru). Очевидно, что в 2011 году был выполнен приказ об унификации доменных имен для этих сайтов. В 2012 и 2013 годах продолжился процесс унификации доменных имен, причем многие из них ранее были поддоменами домена [ru.ru](http://ru.ru). В частности, с домена [ru.ru](http://ru.ru) на домен [spbu.ru](http://spbu.ru) было переведено 30 сайтов.

Отметим также, что на официальном сайте СПбГУ найдено несколько разделов, посвященных ректорским совещаниям, на которых затрагивались вопросы университетского веб-пространства. В частности, в материалах ректорского совещания от 02.04.2012 (<http://spbu.ru/structure/dekanskie/6287-6287#p6>) говорится: «... ректор напомнил деканам, что с главной страницы портала СПбГУ можно легко попасть на страницу/сайт любого факультета — достаточно зайти в раздел «Факультеты». А вот на страницах/сайтах далеко не всех факультетов есть обратные ссылки на главную страницу сайта СПбГУ (нет, например, у биологов, математиков ...). На странице факультета любого зарубежного университета есть такие ссылки на главную страницу вуза». К сожалению, решений ректорских совещаний, которые можно было бы рассматривать как прямые административные воздействия на веб-пространство СПбГУ, найти не удалось. Однако очевидно, что такие обсуждения на самом высоком уровне не прошли бесследно. Обратим внимание, что

время проведения указанного ректорского совещания попадает в интервал с 2011 по 2013 годы, когда происходил наибольший прирост количества дуг в графе.

Таким образом, можно сделать вывод о наличии административных воздействий на веб-пространство СПбГУ, зафиксированных в 2011 и 2012 годах, совпадающих с основными моментами времени резкого увеличения количества дуг и вершин в веб-графе.

### **ЗАКЛЮЧЕНИЕ**

В статье на примере веб-пространства СПбГУ описан достаточно общий подход, позволяющий ответить на вопрос о том, насколько управляемым является его процесс формирования и развития.

Описанный подход к построению веб-графа организации не представляется слишком сложным, тем более, что данные об исходящих с сайта гиперссылках можно сегодня получить из многих источников.

Сравнение веб-графа реального веб-пространства с моделью Барабаши–Альберт наверняка не является единственно возможным подходом подтвердить тот факт, что веб-пространство не обладает свойством самоорганизации. Эта тема может быть продолжена, например, с использованием модели Боллобаша–Риордана, которая учитывает направленность дуг. По крайней мере, мы с уверенностью можем сказать, что веб-граф сайта СПбГУ не является графом Барабаши–Альберт.

Предложенная в статье динамическая модель веб-графа позволяет увидеть «выбросы» в естественном течении процесса, датировать их и сопоставлять с событиями, отраженными в том же самом веб-пространстве организации. Пример СПбГУ показывает интересные возможности этого подхода.

Собственно говоря, управление веб-пространством организации сводится к ряду достаточно очевидных административных воздействий, направленных на увеличение (до некоторого «разумного предела») количества сайтов и гиперссылок между ними (ответ на вопрос о «разумных пределах» далеко не очевиден и выходит далеко за рамки данной статьи, именно поэтому это словосочетание заключено в кавычки).

На примере веб-пространства СПбГУ можно заметить, что возможности административного воздействия в основном были направлены на увеличение количества сайтов посредством «унификации имен» и на сегодня практически исчерпаны. Об этом свидетельствуют и графики на рис. 3; смотрите значения нижнего графика в 2014 и 2015 годах.

Далеко не исчерпана тема обратных ссылок, затронутая на упомянутом ректорском совещании, её можно воплотить в распоряжение, касающееся обязательности гиперссылок на сайты вышестоящих подразделений.

По большому счёту необходимо разработать официальный документ, называющийся, например, «Концепция развития единого информационного веб-пространства СПбГУ», с изложением основных целей, задач и перспектив развития этого важного направления деятельности вуза. Такого документа на официальном сайте СПбГУ найти, к сожалению, не удалось.

### ***Благодарности***

Работа выполнена при поддержке гранта РФФИ № 15-01-06105 «Разработка вебметрических и эргономических моделей и методов анализа эффективности присутствия в Вебе информационных веб-пространств крупных организаций».

### **СПИСОК ЛИТЕРАТУРЫ**

1. *Печников А.А.* Построение и исследование веб-графа информационного веб-пространства Санкт-Петербургского государственного университета // *Фундаментальные исследования*. 2015. № 10 (часть 3). С. 512-517.
2. *Печников А.А.* Методы исследования регламентируемых тематических фрагментов Web // *Труды Института системного анализа Российской академии наук. Серия: Прикладные проблемы управления макросистемами*. 2010. Т. 59. С. 134-145.
3. *Thelwall M.* Link analysis: an information science approach / Amsterdam: Elsevier Academic Press, 2004. 269 p.
4. *Райгородский А.М.* Модели случайных графов и их применения // *Труды МФТИ*. 2010. Т. 2, № 4. С. 130-140.

5. Broder A., Kumar R., Maghoul F., Raghavan P., Rajagopalan S., Stata R., Tomkins A., Wiener J. Graph structure in the Web // *Journal of Computer Networks*. 2000. No 33 (1-6). P. 309-320.

6. Blekanov I.S., Sergeev S.L., Maksimov A.I. Analysis of the topology of large Web segments using Broder's bow-tie model // *Life Science Journal*. 2014. V. 11. No 6 Spec. Iss. P. 258-261.

7. Thelwall M., Harries G. The connection between the research of a University and counts of links to its Web pages: an investigation based upon a classification of the relationships of pages to the research of the host university // *Journal of the American Society for Information Science and Technology*. 2003. V.54, Iss.7. P. 593-699.

8. Ortega J.L., Aguillo I.F. Interdisciplinary relationships in the Spanish academic web space: A Webometric study through networks visualization // *Cybermetrics. International Journal of Scientometrics, Informetrics and Bibliometrics*. 2007. V. 11, Iss. 1. [Электронный ресурс] – Режим доступа: <http://cybermetrics.cindoc.csic.es/articles/v11i1p4.pdf>.

9. Kenekayoro P., Buckley K., Thelwall M. Hyperlinks as inter-university collaboration indicators // *Journal of Information Science*. 2014. V. 40, No 4. P. 514-522.

10. Flake G.W., Lawrence S.R., Giles C.L., Coetzee F.M. Self-organization and identification of Web communities // *IEEE Computer*. 2002. V. 35, No 3. P. 66-71.

11. Веснин Ю.А., Константинова Е.В., Савин М.Ю. О сценариях присоединения новых сайтов к веб-пространству СО РАН // *Вестник НГУ. Серия: Информационные технологии*. 2013. Т. 11, вып. 4. С. 28-37.

12. Barabasi L.-A., Albert R., Jeong H. Scale-free characteristics of random networks: the topology of the world-wide web // *Physica*. 2000. V. A281. P. 69-77.

13. Newman M.E.J. Power laws, Pareto distributions and Zipf's law // *Contemporary Physics*. 2005. V. 46, No 5, September–October. P. 323-351.

14. Internet Archive: Wayback Machine. URL: <http://archive.org/web>.

---

**WEB RESOURCES OF RUSSIAN UNIVERSITIES:  
THE SELF-ORGANIZATION OR ADMINISTRATIVE IMPACT?  
CASE STUDY OF ST. PETERSBURG STATE UNIVERSITY**

**A.A. Pechnikov**

*Institute of Applied Mathematical Research*

*Karelian Research Centre of RAS*

pechnikov@krc.karelia.ru

**Abstract**

The web space of the organization is the many web sites connected by hyperlinks. The main question considered in the article, is whether this web space self-organizing, i.e. the ordering of its elements occurs at the expense of their internal interaction without external influences, or external (so-called "administrative") the impact is so strong that their influence can be detected. The article proposes a common approach demonstrated by the example of the web space of St. Petersburg State University. On the question of whether significant administrative influence on the web space of the University, in this case on the basis of the analysis given a positive response.

**Keywords:** *hyperlink, web site, web graph, dynamic model, administrative impact*

**REFERENCES**

1. *Pechnikov A.A.* Postroenie i issledovanie veb-grafa informacionnogo veb-prostranstva Sankt-Peterburgskogo gosudarstvennogo universiteta // Fundamental'nye issledovaniya. 2015. No 10 (part 3). S. 512-517.
2. *Pechnikov A.A.* Metody issledovaniya reglamentiruemyh tematicheskikh fragmentov Veba // Trudy Instituta sistemnogo analiza Rossiiskoi Akademii Nauk. Serija: Prikladnye problemy upravleniya makrosistemami. 2010. T. 59. S. 134-145.
3. *Thelwall M.* Link analysis: an information science approach / Amsterdam: Elsevier Academic Press, 2004. 269 s.
4. *Raigorodskii A.M.* Modeli sluchainyh grafov i ih primenenija // Trudy MFTI. 2010. T. 2, № 4. S. 130-140.

5. Broder A., Kumar R., Maghoul F., Raghavan P., Rajagopalan S., Stata R., Tomkins A., Wiener J. Graph structure in the web // *Journal of Computer Networks*. 2000. No 33 (1-6). S. 309-320.

6. Blekanov I.S., Sergeev S.L., Maksimov A.I. Analysis of the topology of large Web segments using Broder's bow-tie model // *Life Science Journal*. 2014. T. 11, No 6 Spec. vip. S. 258-261.

7. Thelwall M., Harries G. The Connection between the research of a university and counts of links to its web pages: an investigation based upon a classification of the relationships of pages to the research of the host university // *Journal of the American Society for Information Science and Technology*. 2003. T. 54, Iss. 7. P. 593-699.

8. Ortega J.L., Aguillo I.F. Interdisciplinary relationships in the Spanish academic web space: A Webometric study through networks visualization // *Cybermetrics. International Journal of Scientometrics, Informetrics and Bibliometrics*. 2007. V. 11, Iss. 1. [Electronic resource] – URL: <http://cybermetrics.cindoc.csic.es/articles/v11i1p4.pdf>.

9. Kenekayoro P., Buckley K., Thelwall M. Hyperlinks as inter-university collaboration indicators // *Journal of Information Science*. 2014. V. 40, No 4. P. 514-522.

10. Flake G.W., Lawrence S.R., Giles C.L., Coetzee F.M. Self-organization and Identification of Web Communities // *IEEE Computer*. 2002. V. 35, No 3. P. 66-71.

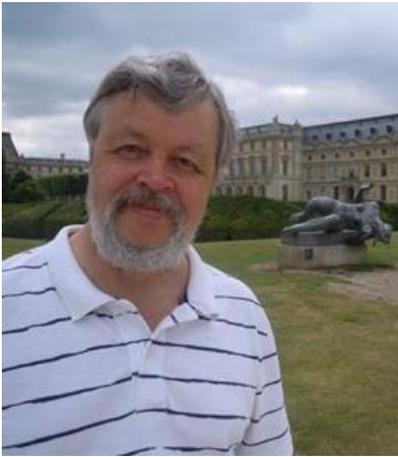
11. Vesnin Yu.A., Konstantinova E.V., Savin M.Yu. O scenarijah prisoedinenija novyh saitov k veb-prostranstvu SO RAN // *Vestnik NGU. Serija: Informacionnye tehnologii*. 2013. T. 11, Vyp. 4. S. 28-37.

12. Barabasi L.-A., Albert R., Jeong H. Scale-free characteristics of random networks: the topology of the world-wide web // *Physica*. 2000. V. A281. P. 69-77.

13. Newman M.E.J. Power laws, Pareto distributions and Zipf's law // *Contemporary Physics*. 2005. V. 46, No 5, September–October. P. 323-351.

14. Internet Archive: Wayback Machine. URL: <http://archive.org/web>.

## СВЕДЕНИЯ ОБ АВТОРЕ



**ПЕЧНИКОВ Андрей Анатольевич** – доктор технических наук, гл. научный сотрудник, руководитель лаборатории телекоммуникационных систем Института прикладных математических исследований Карельского научного центра РАН.

**Andrey Anatolievich PECHNIKOV** – Doctor of Technics (Comp. Sci.), Chief Research Associate, Head in the Laboratory for Telecommunications Systems, Institute of Applied Mathematical Research Karelian Research Centre of RAS.

email: [pechnikov@krc.karelia.ru](mailto:pechnikov@krc.karelia.ru)

*Материал поступил в редакцию 30 ноября 2015 года*