

УДК: 004.351

Алекперова И.Я.

Институт Информационных Технологий НАНА, Баку, Азербайджан

airada.09@gmail.com**АНАЛИЗ ФАКТОРОВ, ВЛИЯЮЩИХ НА РЕЙТИНГ СТАТЕЙ В ВИКИ-СРЕДЕ**

В работе анализируется рейтинг различных статей в вики-среде и на основе этих анализов выявляются некоторые факторы, которые и влияют на рейтинг. Факторы, которые обуславливают рейтинг той или иной статьи, могут зависеть как от времени и языкового раздела Википедии, так и от темы, которая описывается в статье. Некоторые выводы сделаны на основе сравнительного анализа.

Ключевые слова: Википедия, вики-пользователи, вики-среда, вики-анализ, рейтинг статьи.

Введение

Как известно, Википедия – это свободная общедоступная мультязычная универсальная интернет-энциклопедия [1]. Но сама по себе Википедия не высказывает никаких мнений и заключений, а дает только информацию, основанную на авторитетных источниках. Все, что пишется в Википедии, пишется добровольцами, которые не получают за это зарплату [2]. Википедия ориентирована на создание онлайн-энциклопедии, написанной самими пользователями, которая в свою очередь создает вики-среду. Вики-среда представляет собой автоматизированную социальную среду, позволяющую общаться группе пользователей, объединенных общим интересом. Каждый посетитель сайта энциклопедии может внести свой посильный вклад: подправить статью, добавить или удалить информацию, участвовать в обсуждениях [3].

Стоит особо отметить, что не все может стать предметом новой статьи. Статья в Википедии может быть написана и сохранена, если этот предмет достаточно подробно освещается во внешних источниках. Об этом гласит главный критерий значимости, который был создан сообществом вики-среды [4]. И естественно, что каждую статью в Википедии посещает определенное количество читателей. О том, сколько вики-пользователей посетило ту или иную статью, можно узнать на сайте stats.grok.se (в первом списке выбирается язык Википедии, во втором – месяц посещений, в третьем – название статьи) [5]. Конечно, число посетителей во всех статьях неодинаково. В этой работе будут проанализированы факторы, которые влияют на рейтинг статей в вики-среде.

О рейтинге статей в вики-среде

В последние годы вики-среда стала существенным инструментом информационного влияния. Она не только объединяет людей и дает возможность по обмену информацией, свободному ее обсуждению, но также служит для дезинформации, пропаганды, управления мнениями [6]. С началом любого конфликта в мире резко повышается активность вики-пользователей и рейтинг некоторых статей значительно увеличивается.

В первую очередь рейтинг зависит от значимости темы статьи для читателей, желающих прочесть статью на том или ином языке. К примеру, за октябрь 2012 года статью «Россия» в Русской Википедии посетили в среднем 11 000 раз (рис. 1), статью «Азербайджан» около 2000 раз, а статью про относительно малоизвестную африканскую страну Бурунди в среднем 200 раз (рис. 2). Это означает, что для русскоязычных читателей статьи про Россию и прочие темы, касающиеся российской тематики, намного значимее и интереснее.

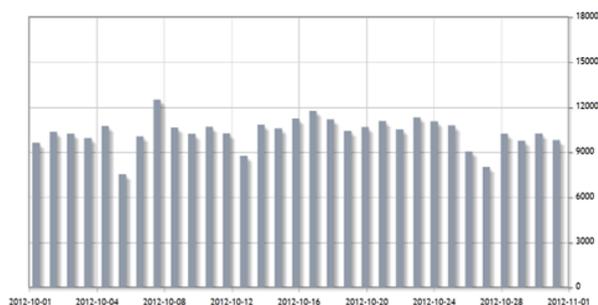


Рис. 1. Рейтинг статьи «Россия» в Русской Википедии за октябрь 2012 года.

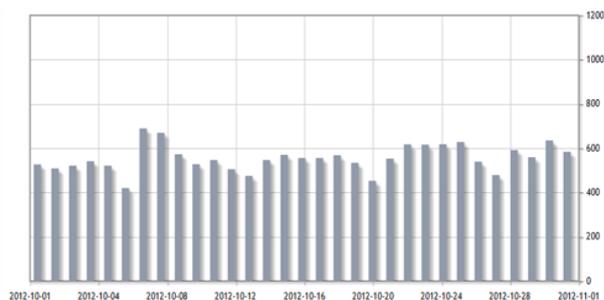


Рис. 2. Рейтинг статьи «Бурунди» в Русской Википедии за октябрь 2012 года.

Прочтение статьи в то или иное время также зависит от того, насколько интересна и актуальна тема статьи в данный момент. К примеру, статью «Баку» с 1 по 21 мая 2012 года просмотрели в среднем 800 человек. А 22 мая в Баку стартовал конкурс песни «Евровидение». О том, насколько эта тема была интересна для общества, наглядно показывает статистика посещений статьи «Азербайджан» в дни проведения конкурса. Так, 22 и 24 мая, то есть в дни полуфинала «Евровидения», статью о Баку в Русской Википедии посмотрели более 3000 раз, а 26 мая, в день финала, – 7655 раз (рис. 3).

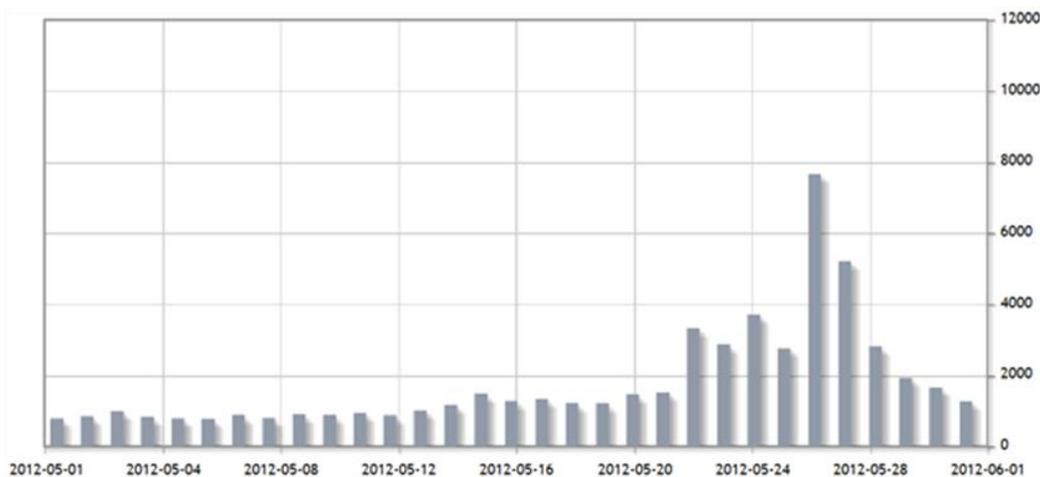


Рис. 3. Количество посещений статьи «Баку» в Русской Википедии за май 2012 г.

Из рис. 3 видно, что рейтинг статьи после конкурса начинает падать. 27 мая статью «Баку» посмотрели 5208 раз, 28 мая – 2809, а 31 мая – уже 1265. Это означает, что после окончания конкурса интерес к статье начинает падать и эта тема может считаться более неактуальной. Вывод этот был сделан на основе вики-анализа – анализа статистики посещения статьи в вики-среде. В связи с этим можно сделать вывод о том, что, проанализировав статью о том или ином объекте в вики-среде, можно с определенной степенью уверенности говорить об актуальности и значимости того или иного объекта для общества в определенный период времени.

Рейтинг статьи также резко возрастает при получении статьей в том или ином разделе статуса «хорошая» или «избранная». Для того чтобы статья получила статус «хорошая», она должна удовлетворять определенным критериям. К примеру, тема статьи должна быть раскрыта, статья должна быть законченной и не содержать откровенно недописанных разделов, объем статьи должен быть не меньше 8 тыс. знаков и т.д. [7]. Для того чтобы статья удовлетворяла этим требованиям, один из вики-пользователей, как правило, основной автор статьи, номинирует статью, а другие участники проекта

высказывают свои замечания [8]. После устранения всех недочетов статья получает статус. И информация о ней на несколько дней становится доступной на заглавной странице Википедии. К примеру, 4 ноября статус «хорошая» был присвоен статье «Дворец Шекинских ханов» (рис. 4).

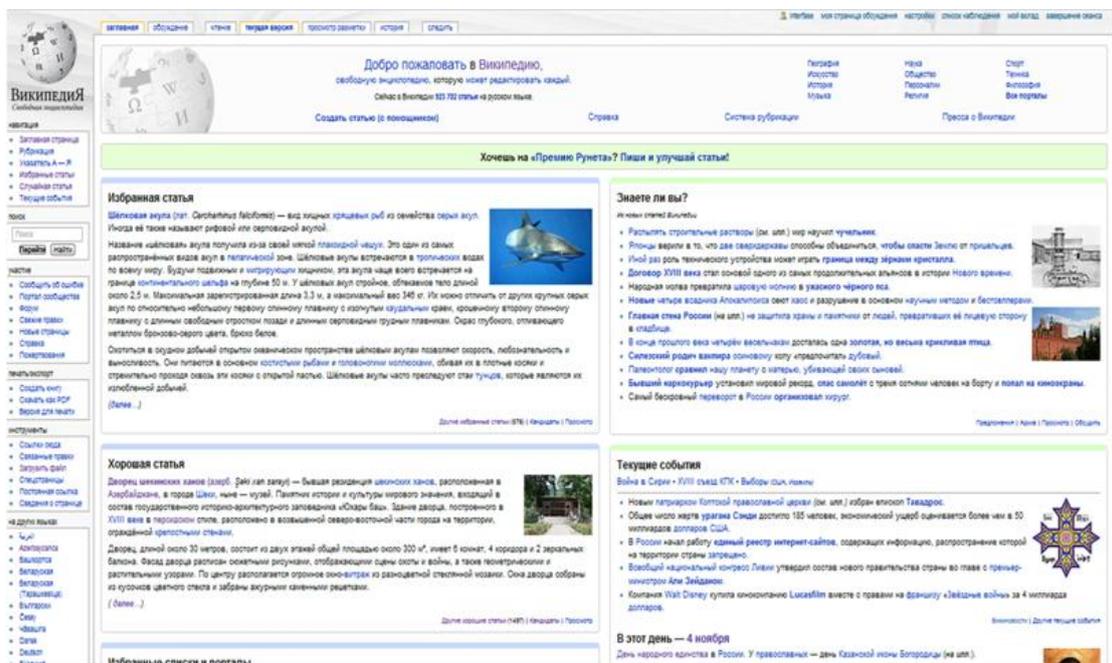


Рис. 4. Заглавная страница Русской Википедии на 4 ноября 2012.
Статья «Дворец Шекинских ханов» – хорошая статья

Теперь, если посмотреть статистику посещения статьи про дворец (рис. 5), то можно увидеть, что еще 3 ноября статью посетили всего 20 раз, а уже 5 ноября, через день после получения статуса «хорошая», когда информация о ней была уже на заглавной странице Русской Википедии, ее посмотрели 1205 раз. Этот пример наглядно показывает, как получение статуса «хорошая» или «избранная» может резко увеличить интерес к ней читателей, заглянувших на заглавную страницу Википедии. А это значит, что читатели уже были знакомы с сайтом Википедии и сначала заглянули на ее заглавную страницу и только потом на страницу статьи про дворец.

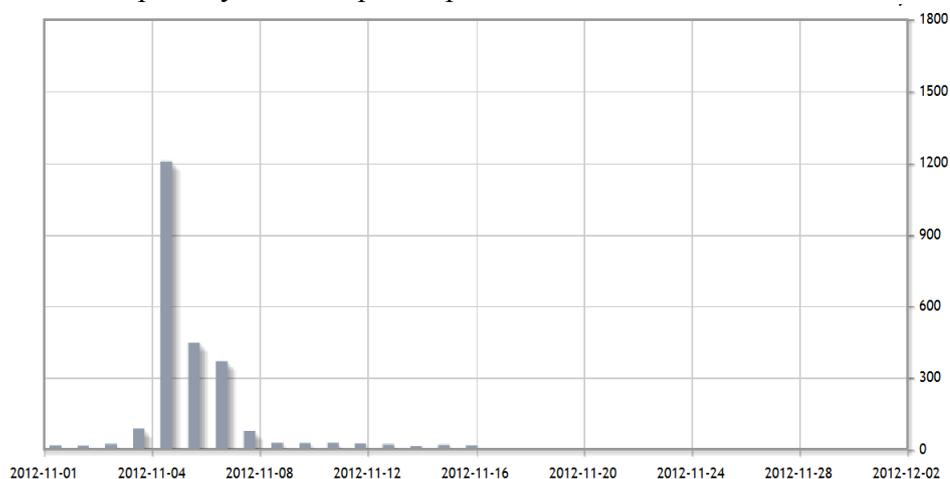


Рис. 5. Количество посещений статьи «Дворец Шекинских ханов» в Русской Википедии за ноябрь 2012 г.

В отличие от предыдущего примера, когда посетители статьи про Баку, возможно, и не заходили предварительно на заглавную страницу Википедии, а приходили на страницу через поисковики (многие, к примеру Google, [при поиске первым выдает статью из Википедии](#)), то посетители статьи про Дворец Шекинских ханов предварительно заходили на заглавную страницу Википедии. Из этого можно сделать вывод о том, что не только вызывающая интерес к статье причина может быть различной, но и интересы самих вики-пользователей могут отличаться.

Таким образом, анализ рейтинга той или иной статьи может оказаться очень информативным. Так, большое количество посещений в том или ином языковом разделе Википедии может говорить о важности и значимости объекта этой статьи. Резкое же увеличение рейтинга относительно незначимой статьи в тот или иной промежуток времени может означать резкий рост интереса к этой статье общества в этот период, а резкое снижение числа посетителей – уменьшение интереса. Это в свою очередь зависит от событий, связанных с объектом статьи в этот период времени.

Факторы, влияющие на рейтинг статей

Следует также отметить, что рейтинг статьи на одну и ту же тему в разных языковых разделах может кардинально различаться. Во-первых, это зависит от того, что количество говорящих на том или ином языке неодинаково. Во-вторых, это зависит от интереса носителей разных языков к теме статьи.

Проведем сравнительный анализ резкого роста рейтинга про Баку в дни проведения «Евровидения» в нескольких языковых разделах. Так, 26 мая 2012 года (в дни проведения «Евровидения») статья про город Баку в Русской Википедии была просмотрена 7655 раз, а в Английской Википедии – 43 878 раз (рис. 6).

Такое различие в числе просмотров зависит от того, что носителей английского языка около 2,5 млрд человек, а носителей русского языка – около 173 млн человек [9], в 14 раз меньше, чем носителей английского. Естественно, что и количество посещений статьи в Английской Википедии в 17 раз больше, чем в Русской.

Посмотрим теперь количество посещений статьи в Немецкой Википедии. Так, 26 мая статья про Баку была просмотрена 33 002 раза (рис. 7). Однако в отличие от носителей английского языка, количество носителей немецкого – около 109 млн человек, то есть меньше, чем носителей русского.

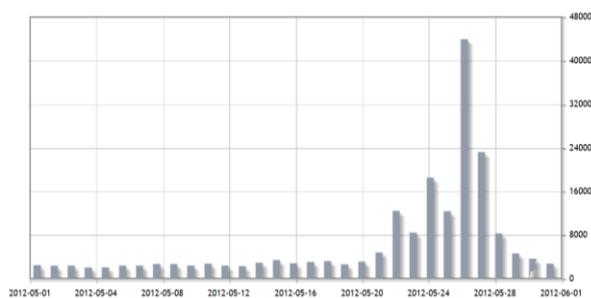


Рис. 6. Количество статей про Баку в Английской Википедии за май 2012 г.

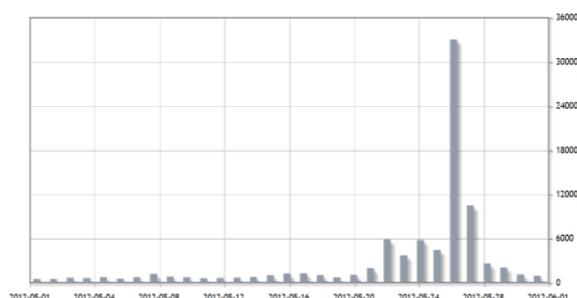


Рис. 7. Количество статей про Баку в Немецкой Википедии за май 2012 г.

То, что рейтинг статьи в немецком разделе выше, чем в русском, означает, что интерес носителей немецкого к статье об актуальном событии в Википедии, да и в Интернете в целом, больше, чем у носителей русского.

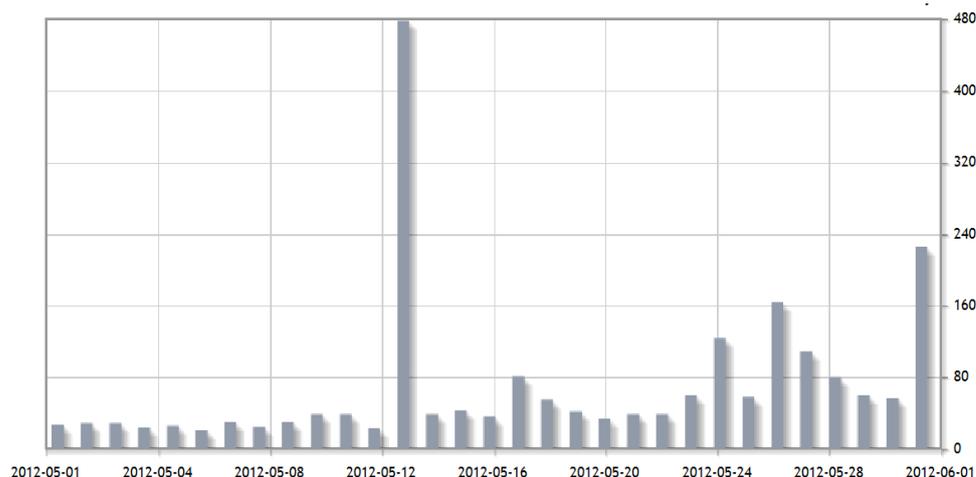


Рис. 8. Количество статей про Баку в Китайской Википедии за май 2012 г.

Однако есть и другие случаи. К примеру, носителей китайского языка около 1,4 млрд человек, намного больше, чем носителей русского языка. Однако статью про Баку в Китайской Википедии просмотрели в мае в целом 2136 раз (рис. 8), что почти в 3,5 раза меньше, чем в Русской Википедии. Такое различие может быть вызвано тем, что, во-первых, большинство носителей китайского языка не так заинтересованы в событии, чем носители русского языка, а во-вторых, тем, что информации о Баку на китайском меньше, чем в английском языке, и многие носители китайского языка ищут информацию о том или ином объекте или событии в других наиболее популярных языковых разделах.

Различие между просмотрами статьи на английском и китайском языках можно легко распознать, если обратить внимание на графики числа статей в том или ином языке и числа говорящих на этом языке. На рис. 9 видно, что английский язык по числу говорящих на нем людей занимает первое место в мире, а китайский – второе.

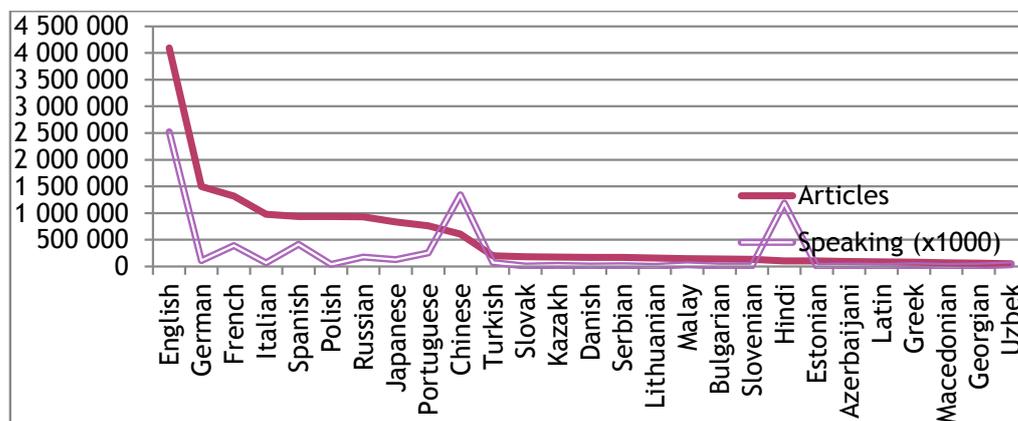


Рис. 9. Зависимость количества статей от языкового фактора

Однако количество статей в китайском разделе Википедии намного меньше количества статей в английском разделе, следовательно, и активность пользователей китайского раздела низкая.

Из вики-анализа следует, что хотя и есть зависимость между носителями того или иного языка и числом статей в вики-среде на этом языке, в некоторых случаях эта зависимость нарушается (китайский, хинди и др.). Этот факт, в свою очередь, оказывает непосредственное влияние на рейтинг посещения статьи в этих языковых разделах Википедии.

Вывод

Исследования показывают, что с началом любого события в мире (международные конфликты, президентские выборы, международные конкурсы или соревнования и т.д.) резко повышается активность пользователей и рейтинг некоторых статей в вики-среде. Таким образом, проведя анализ факторов, влияющих на рейтинг статьи об одном и том же объекте в разных языковых разделах Википедии, был сделан вывод о том, что отличие в количестве посещений может быть напрямую связано с количеством носителей того или иного языка. Рейтинг статьи может также зависеть от интереса носителей языка к какому-либо событию или объекту, а также интереса поиска информации об этом объекте или событии в Вики-среде или в Интернете.

Литература

1. Алгулиев Р.М., Алекперова И.Я., Касумова Р.Т. Коллективное знание в Интернете / Шестая международная конференция «Интернет – Образование – Наука-2008». Том 1, Винница, 2008, с. 5–7.
2. Robert Sumi, Taha Yasseri, András Rung, András Kornai, János Kertész, Edit wars in Wikipedia / IEEE Third International Conference on Social Computing (SocialCom), Privacy, Security, Risk and Trust (PASSAT), 9-11 Oct. 2011, pp. 724-727.
3. Buriol L. S., Castillo C., Donato D., Leonardi S., Millozzi S., Temporal Analysis of the Wikigraph // Web Intelligence (2006), Publisher: IEEE CS Press., pp. 45-51.
4. <http://ru.wikipedia.org/wiki/Википедия:Значимость>.
5. <http://stats.grok.se/>
6. [Ələkbərova İ.Y. Viki-mühitdə reallaşdırılan bəzi informasiya müharibəsi texnologiyalarının analizi / İnformasiya cəmiyyəti problemləri, 2011, №2\(4\), səh. 18-28.](#)
7. <http://ru.wikipedia.org/wiki/Хорошие статьи>.
8. <http://ru.wikipedia.org/wiki/Википедия:Кандидаты в хорошие статьи>.
9. http://ostranah.ru/_lists/languages.php.

UOT 004.351

Ələkbərova İ.Y.

AMEA İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan
airada.09@gmail.com

Viki-mühitdə məqalələrin reytinginə təsir edən faktorların analizi

Məqalədə viki-mühitdəki müxtəlif məqalələrin reytinginin analizi aparılmış və analiz əsasında məqalələrin reytinginə təsir edən bəzi faktorlar müəyyən olunmuşdur. Məqalələrin reytinginə təsir edən faktorlar Vikipediyanın dil bölmələrindən və zamandan asılı olduğu kimi, məqalələrdə təsvir olunan mövzulardan da asılıdır. Tədqiqat işində müqayisəli analiz əsasında bəzi nəticələr əldə olunmuşdur.

Açar sözlər: Vikipediya, viki-istifadəçilər, viki-mühit, viki-analiz, məqalənin reytingi.

İrada Y. Alakbarova

Institute of Information Technology of ANAS, Baku, Azerbaijan
airada.09@gmail.com

Analysis of the factors affecting the ranking of articles in the wiki space

This work analyzes the rating of various articles in the Wiki-space and on the basis of this analysis some of the factors, which influence the rating were revealed. The reason that causes the rating to a particular article may depend on the time and the language section of Wikipedia, and on a topic that is described in the article. Some conclusions are made on the basis of comparative analysis.

Keywords: Wikipedia, wiki-users, wiki-community, wiki analysis, rating of article.