

Оценка качества доступов как фактор выбора при продвижении библиотеки в социальных сетях (на примере сайта Отделения ГПНТБ СО РАН)

Комаров Сергей Юрьевич

Библиотекарь 1 категории ГПНТБ СО РАН

Отдел создания, развития и продвижения информационно-
библиографических ресурсов

Группа информационных систем и электронных ресурсов

skomarov87@gmail.com

Основные тематические исследования

- А.Е.Гуськов, Е.С.Быховцев, Д.В.Косяков Альтернативная вебметрика: исследование веб-трафика сайтов научных организаций // НТИ. Серия 1: Организация и методика информационной работы. – 2015. – №2. – Ст.12-28.
- А.Р.Фахруллина, Э.Ю.Степанова, И.Б.Юлуев Применение социальных сетей по технологии Web 2.0 в учебном процессе технического университета // Молодежный вестник УГАТУ. – 2015. – №1(13). – Ст.126-132.
- А.С.Шуклина, Ф.А.Перепелица, Е.В.Шалобаев Исследование и сравнительный анализ продвижения интернет-проектов в социальных сетях // Альманах научных работ молодых ученых Университета ИТМО в 5 т. – 2016. – Ст.280-281.
- Г.В.Можаева, П.Н.Можаева-Ренья Twitter в информационной среде Digital Humanities // Информационный бюллетень ассоциации «История и компьютер». – 2016. – №45. – Ст.46-48.
- И.В.Юрик Как продвигать библиотеку в Twitter. Секреты кухни корпоративного микроблога // Современная библиотека. – 2013. – №1. – Ст.22-29.
- И.К.Алдарова Социальные сети как инструмент современного маркетинга // Бизнес-образование в экономике знаний. – 2017. – №2. – Ст.6-10.
- С.А.Кошечкин, Ю.А.Тузоз Создание механизма генерации трафика клиентов из социальных сетей // Промышленный и b2b маркетинг. – 2014. – №4. – Ст.280-290; Они же Генерация трафика клиентов из социальных сетей // Маркетинговые коммуникации. – 2014. – №4. – Ст.202-217.
- С.В.Бредихин., Н.Г.Щербакова Две компоненты анализа сетевого трафика // Вестник НГУ. Серия «Информационные технологии». – 2008. – №1. – Ст.10-14.
- С.П.Главацкий Статистический анализ трафика социальных сетей // Наукові праці ОНАЗ ім. О.С.Попова. – 2013. – №2. – Ст.94-99.
- A.Chandra, M.Suaib A survey on Web spam and Spam 2.0 // International journal of advanced computer research. – 2014. – V.4. – №2. – I.15. – Pp.634-644.
- H.Gao et.al. Detecting and characterizing social spam campaigns // Proceedings of the 17th ACM Conference on computer and communication security. – 2010. – Chicago, Illinois, USA. – Pp.681-683.
- H.Kanchan, L. Ahuja, S.K.Muttoo A Fuzzy Logic Approach for detecting Redirection Spam // International Journal of Electronic Security and Digital Forensics. – 2016. – V.8. – I.3. – Pp.191-204.
- H.Kanchan, L. Ahuja, S.K.Muttoo Approaches for Web spam detection // Int. Journ. of Computer Applications. – 2014. – V.101. – №1. – Pp.38-44.
- M.Egel, K.Kolbich, K.Platzer Removing web spam links from search engine results // Journal in Computer Virol. – 2011. – V.7. – I.1. – Pp.51-62.

Используемые термины

- **Реферер** - запись в лог-файлах сервера о посетителе сайта. Если поле реферера пусто, то это обычно означает, что пользователь набрал адрес непосредственно в своем браузере. Реферер используется при анализе статистики сайта
- **Хит** — это наиболее точный фрагмент данных в веб-аналитике. По своему содержанию хит — это запрос небольшого графического файла с сервера веб-аналитики. Вместе с каждым таким запросом передаются данные о действиях пользователя на веб-сайте или в мобильном приложении.
- **Сеанс** — это последовательность взаимодействий пользователя с веб-сайтом за указанный промежуток времени. В рамках одного сеанса может произойти несколько просмотров страниц, событий, социальных взаимодействий и транзакций электронной торговли.
- **Отказ** — это сеанс с просмотром одной страницы на вашем сайте. В Google Analytics (и других веб-аналитических системах) отказом считается сеанс, в ходе которого был активирован только один запрос к серверу Google Analytics (например, если пользователь открыл одну страницу сайта и покинул его, не активировав других запросов).
- **Конверсия**, буквально — это переход из одного состояния в другое. В контексте аналитики конверсией называется выполнение каких-то целевых действий. Конверсии, обычно, делятся на макро-конверсии - выполнение целей, направленных на прямой доход, и микро-конверсии - указывают на заинтересованность, вовлеченность посетителя в материалы сайта.
- **Ghost referral spam** — искаженные данные о внешних пользователях, основанные не на реальных посещениях целевого сайта, а на прямой отправке ложной информации в систему веб-аналитики посредством использования уязвимостей в Measurement Protocol — протокол передачи статистических данных, позволяющий отправлять необработанные данные напрямую на серверы Google Analytics посредством HTTP-запросов.
- **Crawler referral spam** - искаженные данные о внешних пользователях, основанные на реальных посещениях целевого сайта, но осуществленных не пользователями, а специальными программами-ботами, направленными на продвижение рекламной либо вредоносной информации.
- **Languafe referral spam** — новый тип спама, основанный на вредоносных доступах к целевому сайту из источника, в адресе либо в описании которого намеренно искажаются либо прямо подменяются буквы и символы — например, латиница заменяется кириллицей, или буквы заменяются внешне похожими на них символьными кодами, что затем используется для преодоления защитных механизмов как целевого сайта в целом, так и встроенных в него систем веб-аналитики.

Предлагаемые автором термины

- **«Позитивный» внешний трафик** - внешний трафик, ориентированный на сугубо научные, или, как минимум, научно-популярные, ресурсы академических библиотек, и происходящий из источников, также относимых к научным или научно-популярным.
- **«Негативный» внешний трафик** - внешний трафик, никак не связанный ни с научными, ни с научно-популярными интересами и происходящий из источников, ориентированных исключительно на рекламный (включая технологии вирусного SEO-продвижения) либо откровенно неприемлемый (порнография и схожие с ней сферы) контент

Методика исследования

Автор предлагает методику 3-ступенчатого анализа количественных и качественных показателей реферального трафика наиболее значимых соцсетей, используя в качестве примера сайт Отделения ГПНТБ СО РАН – <http://www.prometeus.nsc.ru>. Первый этап представляет собой первичный анализ с привлечением данных лог-анализатора AWStats 7.3 (build 20140126) и выделением только показателей числа просмотров страниц и сделанных хитов с ограниченной оценкой качества трафика Отделения ГПНТБ СО РАН, идущего из социальных сетей, на 2 и 3 этапах используются более глубокие данные, полученные путем обращения к возможностям Google Analytics и Яндекс-Метрики.

Общие показатели соцсетей по AWStats

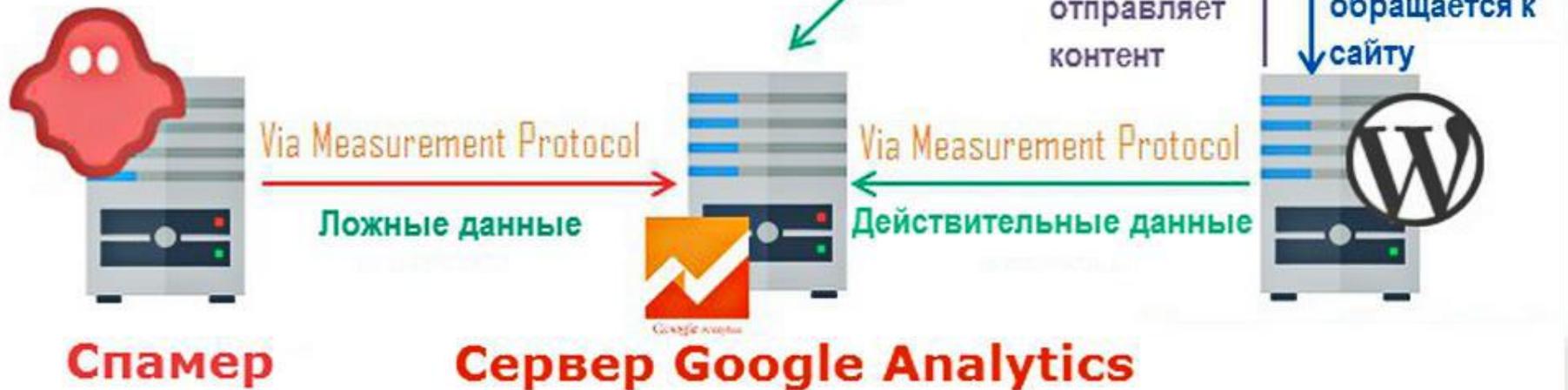
Год	Соцсеть	Число внешних url из соцсети	«Страницы» и «хиты», общее число	«Позитивные» «страницы» и «хиты»	«Негативные» «страницы» и «хиты»	Процентное соотношение «позитива» и «негатива» по страницам и хитам
2015	Вконтакте	98	826/826	657/657	169/169	80%/20%
2016	Вконтакте	55	577/577	45/45	532/532	7.8%/92.2%
2017	Вконтакте	139	2223/2223	3/3	2220/2220	0.13%/99.87%
2015	Твиттер	7	36/36	0/0	36/36	0%/100%
2016	Твиттер	28	43/43	2/2	41/41	5%/95%
2017	Твиттер	32	23/23	6/6	17/17	26%/74%
2015	Фейсбук	43	167/192	1/1	166/191	0.59% и 0.52%/99.4%/99.4%
2016	Фейсбук	34	274/141	4/4	270/137	1.45% и 2.83%/98.54%/97.61%
2017	Фейсбук	94	629/629	5/5	624/624	0.8%/99.2%
2015	Лайвджурнал	255	846/1900	82/484	764/1416	9.7% и 90.3%/25% и 75%
2016	Лайвджурнал	206	267/2039	190/967	77/1072	71.2% и 28.8%/47.4% и 52.6%
2017	Лайвджурнал	140	113/1152	91/772	22/380	80.5 %и 19.5%/67% и 33%
2015	Лайвинтернет	77	193/575	171/254	22/321	88.6 %и 11.4%/44.1% и 55.9%
2016	Лайвинтернет	87	158/473	142/191	16/282	90% и 10%/40.4% и 59.6%
2017	Лайвинтернет	56	146/334	134/198	12/136	92% и 8%/59.3% и 40.7%

Выводы из обработанных данных

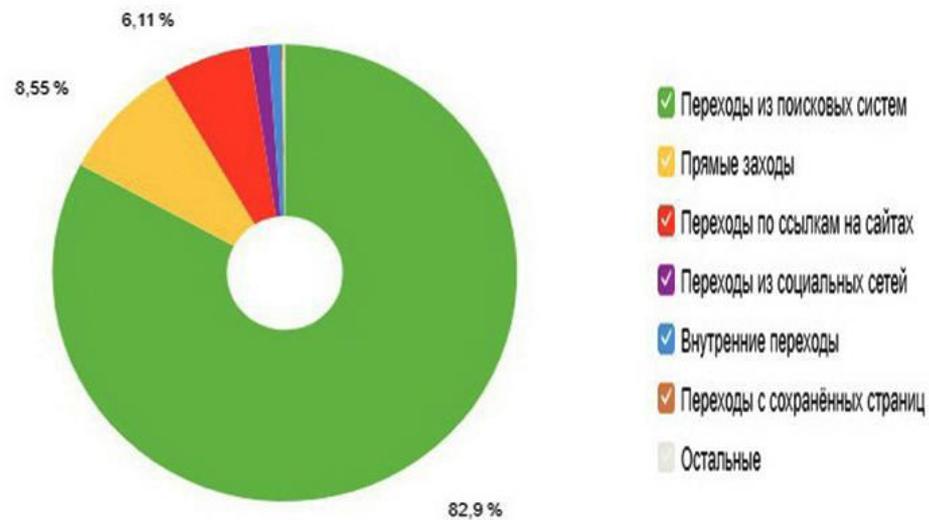
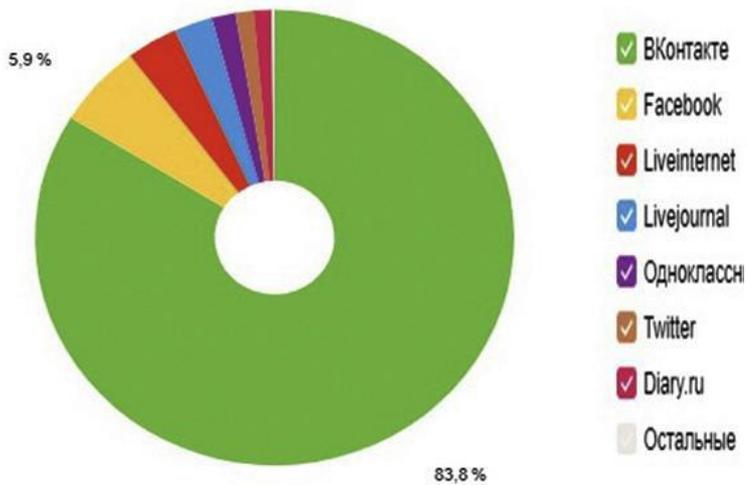
- Соцсети дают крайне малую долю внешнего трафика – буквально, 1.4% за три года (переходы из поисковых систем – 83%);
- В то же время, как канал негативного внешнего трафика, соцсети (большинство) имеют крайне деструктивное значение (постоянно возрастающее), так как предельно затруднены эффективные блокировки их контента;
- Наибольший процент негатива приходит из социальной сети Twitter;
- Социальная сеть Вконтакте (крупнейший источник трафика по группе социальных сетей – 83.8% доступов) демонстрирует «плавающий» внешний трафик;
- Социальная сеть LiveInternet наоборот демонстрирует стабильно позитивные внешние доступы, прямо ассоциированные с конкретными ресурсами сайта, и должна приоритетно использоваться для их продвижения;
- Применительно к научной библиотеке, достоверно прослеживаются лишь «виртуальные» конверсии, связанные с социальными сетями.

Принцип работы Ghost Spam («призрачного» спама)

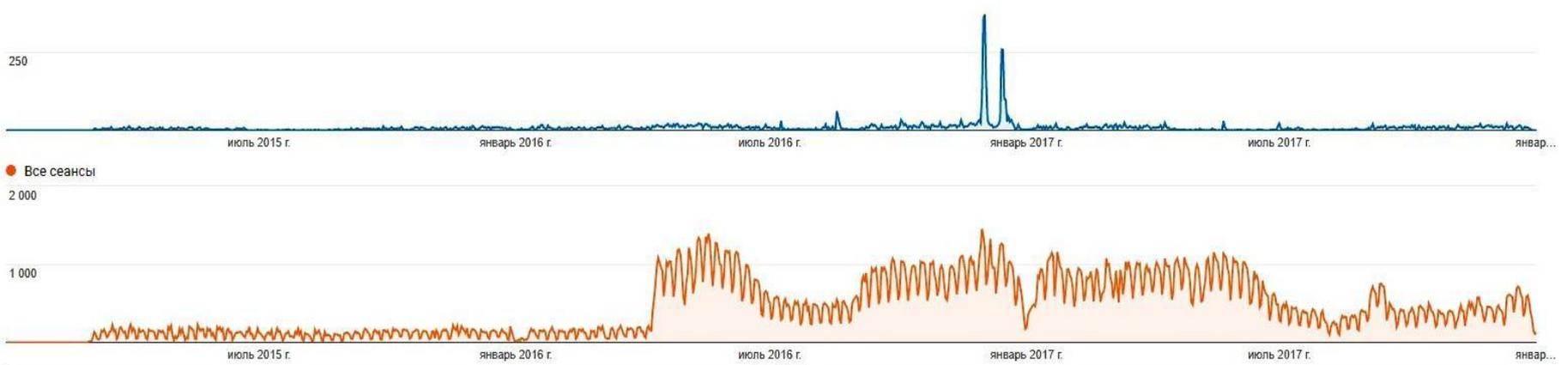
Как работает Ghost Spam?



Доступы из соцсетей в 2015-2018 гг. в сравнении с общими источниками доступов (Яндекс-Метрика)



Страницы с наибольшим числом доступов из соцсетей



Основной параметр: **Общий URL**

Дополнительный параметр

Ещё... [Grid] [Pie] [List] [Refresh]

Общий URL	Сеансы	Просмотры страниц	Сред. длительность сеанса	Страниц/сеанс
1. www.prometeus.nsc.ru/sciguide/index.ssi	3 706 (41,18 %)	10 309 (53,12 %)	00:01:52	2,78
2. www.prometeus.nsc.ru/	1 999 (22,21 %)	4 046 (20,85 %)	00:01:13	2,02
3. www.prometeus.nsc.ru/sciguide/page03.ssi	290 (3,22 %)	413 (2,13 %)	00:01:59	1,42
4. www.prometeus.nsc.ru/sciguide/page01.ssi	253 (2,81 %)	317 (1,63 %)	00:01:15	1,25
5. www.prometeus.nsc.ru/sciguide/page04.ssi	234 (2,60 %)	342 (1,76 %)	00:00:21	1,46
6. www.prometeus.nsc.ru/elibrary/2000vek/137-152	141 (1,57 %)	160 (0,82 %)	00:00:44	1,13
7. www.prometeus.nsc.ru/sciguide/	86 (0,96 %)	226 (1,16 %)	00:02:49	2,63
8. www.prometeus.nsc.ru/biblio/newrus/egrowth.ssi	54 (0,60 %)	61 (0,31 %)	00:00:45	1,13
9. www.prometeus.nsc.ru/biblio/newrus/medins.ssi	52 (0,58 %)	53 (0,27 %)	00:00:08	1,02
10. www.prometeus.nsc.ru/biblio/wex1900/ovchin90.ssi	49 (0,54 %)	52 (0,27 %)	00:00:08	1,06

Строк на странице: 10 К строке: 1 1-10 из 578

Пример выборки AWStats для анализа качества трафика

Отчетный период: - Год - 2017 OK



Вернуться на главную страницу

Фильтр : liveinternet.ru

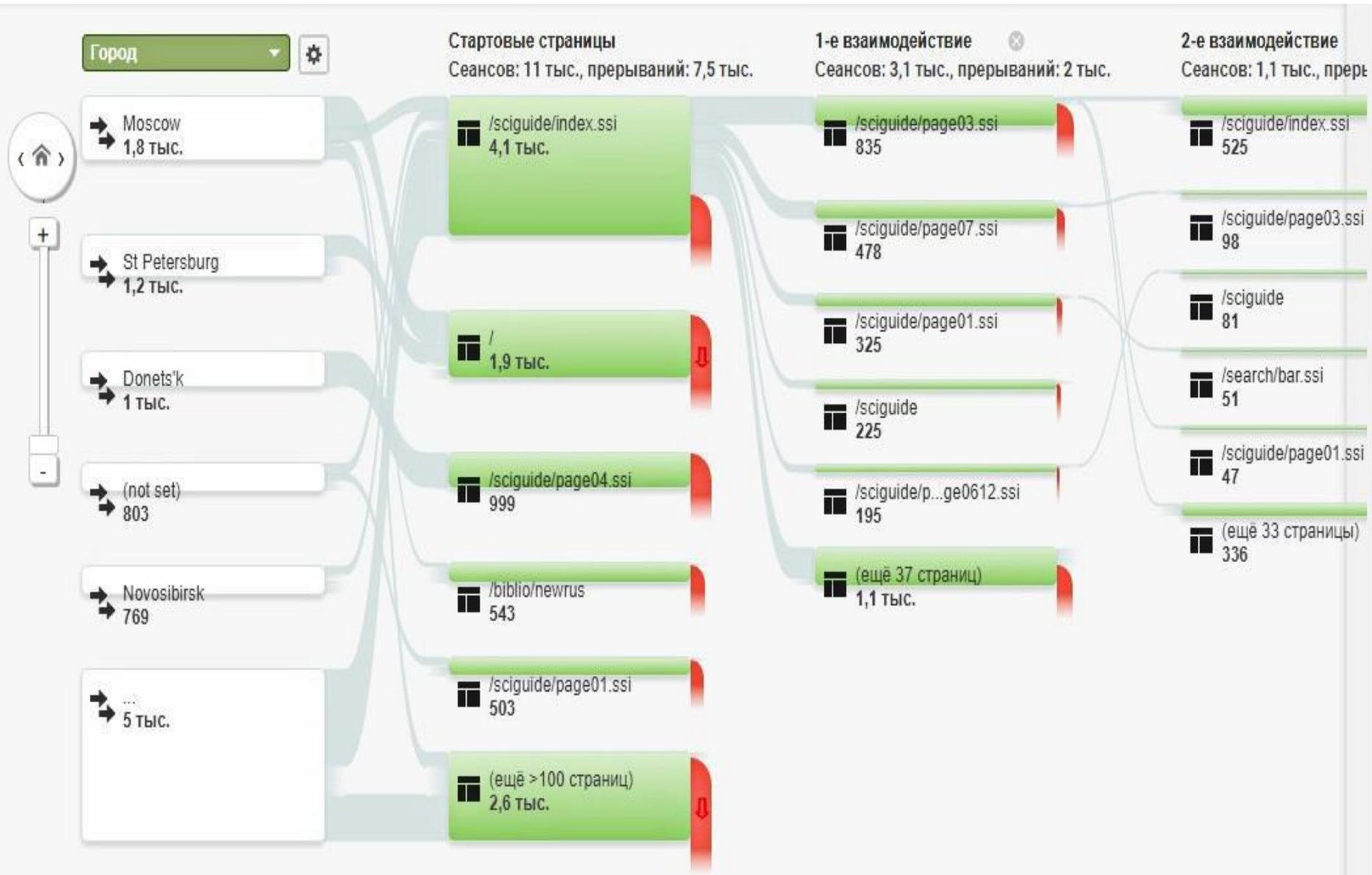
Фильтр исключений :

OK

Ссылки из внешней страницы (остальные web сайты исключая поисковые системы)

Ссылка	Страницы	Процент	Хиты	Процент
http://www.liveinternet.ru/users/barucaba/post124810172	81	46.2 %	81	21.6 %
http://www.liveinternet.ru/users/barucaba/post124810172/	25	14.2 %	25	6.6 %
http://www.liveinternet.ru	17	9.7 %	17	4.5 %
http://www.liveinternet.ru/users/la_belle_epoque/post76992483	11	6.2 %	11	2.9 %
http://www.liveinternet.ru/rating/ru/	4	2.2 %	4	1 %
http://www.liveinternet.ru/users/arin_levindor/post76988245	4	2.2 %	4	1 %
http://www.liveinternet.ru/users/3706516/rubric/1584397/	4	2.2 %	4	1 %
http://www.liveinternet.ru/users/arin_levindor/post76988245/	3	1.7 %	3	0.8 %
http://www.liveinternet.ru/q/	3	1.7 %	3	0.8 %
http://www.liveinternet.ru/journal_proc.php	3	1.7 %	3	0.8 %
https://www.liveinternet.ru/index.html	2	1.1 %	2	0.5 %
http://www.liveinternet.ru/users/3251944/post124818903/	2	1.1 %	2	0.5 %
http://www.liveinternet.ru/stat/helion-ltd.ru/referrers.html	2	1.1 %	2	0.5 %
http://www.liveinternet.ru/stat/portalus.ru/referrers.html	1	0.5 %	1	0.2 %
http://www.liveinternet.ru/rating/ru/humor/	1	0.5 %	1	0.2 %
http://wiki.liveinternet.ru	1	0.5 %	1	0.2 %
http://www.liveinternet.ru/stat/fedordostoevsky.ru/24hours.html	1	0.5 %	1	0.2 %
http://www.liveinternet.ru/users/5153342/post305173802	1	0.5 %	1	0.2 %
http://www.liveinternet.ru/users/lkormacheva/post253519077/	1	0.5 %	1	0.2 %
http://www.liveinternet.ru/stat/prometeus.nsc.ru/	1	0.5 %	1	0.2 %
http://www.liveinternet.ru/users/4455035/post190200814	1	0.5 %	1	0.2 %
http://www.liveinternet.ru/users/3411234/post421173302/&soc=1&ui...	1	0.5 %	1	0.2 %
http://www.liveinternet.ru/journal_post.php	1	0.5 %	1	0.2 %
http://www.liveinternet.ru/users/n-arheolog/post100054025/	1	0.5 %	1	0.2 %
http://www.liveinternet.ru/index.html	1	0.5 %	1	0.2 %
http://wiki.liveinternet.ru/ServisDnevnikovLiveInternet/Fotoal	1	0.5 %	1	0.2 %
http://www.liveinternet.ru/users/astil/post46789643/	1	0.5 %	1	0.2 %
http://www.liveinternet.ru/users/msetynimtio/page1.shtml			1	0.2 %
http://www.liveinternet.ru/users/irena_69/post194026126/			2	0.5 %
http://www.liveinternet.ru/users/msetynimtio/page1.html			1	0.2 %
http://www.liveinternet.ru/users/olga_romeikiene/tags/%FF%ED%E2%...			1	0.2 %
http://www.liveinternet.ru/users/olga_romeikiene/tags/%FF%ED%E2%...			2	0.5 %
http://www.liveinternet.ru/users/3970050/post150350134/			2	0.5 %
http://www.liveinternet.ru/tags/%22F7%E0%F1%F2%ED%E0%FF+%EA%EE%...			1	0.2 %
http://www.liveinternet.ru/users/sarganc/post134228214/			55	14.7 %
http://www.liveinternet.ru/users/nordsahalin/rubric/5942107/page...			1	0.2 %
http://www.liveinternet.ru/users/olga_romeikiene/post92426202/			1	0.2 %
http://www.liveinternet.ru/users/nordsahalin/page1099.shtml			4	1 %
http://www.liveinternet.ru/users/3970050/rubric/1874924/			31	8.2 %

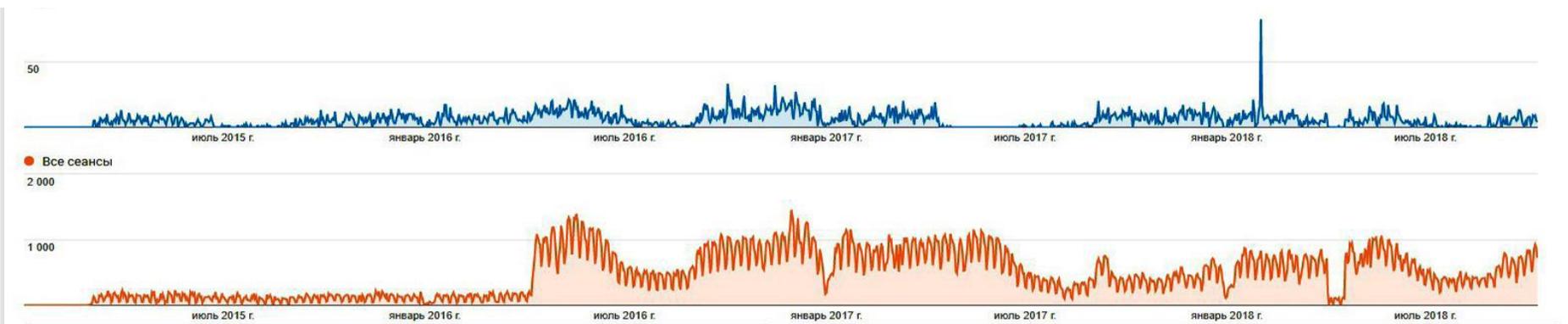
Основные пути пользователей при доступах из соцсетей



«Языковой» спам (Language Referral Spam) Twitter

Полный URL перехода	Просмотры страниц	Сеансы
1. twitter.com/3mapsVitaly	1 531	95,13 %
2. t.co/Mqle3iPR59	45	2,05 %
3. t.co/	20	1,41 %
4. t.co/PNxwXAcLX6	14	0,51 %
5. t.co/Q1Otd4U4sD	12	0,51 %
6. t.co/aK2tlcXydf	4	0,26 %
7. t.co/13QZaRRnNO	1	0,13 %

Доступы из Вконтакте (2015-2018, Google Analytics)



Основной параметр: **Общий URL**

Дополнительный параметр

Ещё...

Общий URL	Сеансы	Просмотры страниц	Сред. длительность сеанса	Страниц/сеанс
1. www.prometeus.nsc.ru/sciguide/index.ssi	4 025 (52,15 %)	11 115 (63,81 %)	00:01:48	2,76
2. www.prometeus.nsc.ru/sciguide/page01.ssi	311 (4,03 %)	420 (2,41 %)	00:01:23	1,35
3. www.prometeus.nsc.ru/sciguide/page03.ssi	303 (3,93 %)	432 (2,48 %)	00:01:54	1,43
4. www.prometeus.nsc.ru/sciguide/page04.ssi	234 (3,03 %)	342 (1,96 %)	00:00:21	1,46
5. www.prometeus.nsc.ru/archives/news/druzhin.ssi	123 (1,59 %)	165 (0,95 %)	00:00:35	1,34
6. www.prometeus.nsc.ru/sciguide/	111 (1,44 %)	313 (1,80 %)	00:02:45	2,82
7. www.prometeus.nsc.ru/biblio/newrus/egrowth.ssi	73 (0,95 %)	83 (0,48 %)	00:00:35	1,14
8. www.prometeus.nsc.ru/	72 (0,93 %)	311 (1,79 %)	00:03:39	4,32
9. www.prometeus.nsc.ru/acquisitions/11-05-03/cont08.ssi	66 (0,86 %)	78 (0,45 %)	00:00:07	1,18
10. www.prometeus.nsc.ru/biblio/newrus/medins.ssi	58 (0,75 %)	59 (0,34 %)	00:00:07	1,02

Полный URL перехода	Сеансы	Просмотры страниц	Сред. длительность сеанса	Страниц/сеанс
1. vk.com/away.php	2 711 (69,16 %)	7 786 (71,88 %)	00:01:57	2,87
2. away.vk.com/away.php	485 (12,37 %)	1 194 (11,02 %)	00:01:48	2,46
3. vk.com/topic-5632110_25655215	439 (11,20 %)	1 089 (10,05 %)	00:01:55	2,48
4. m.vk.com/away.php	274 (6,99 %)	745 (6,88 %)	00:01:37	2,72
5. vk.com/op_21_2013	7 (0,18 %)	11 (0,10 %)	00:00:21	1,57
6. m.vk.com/topic-5632110_25655215	2 (0,05 %)	4 (0,04 %)	00:00:13	2,00
7. vk.com/im	2 (0,05 %)	3 (0,03 %)	00:00:06	1,50

Уважаемые коллеги, благодарю Вас за внимание и готов ответить на Ваши вопросы.

Комаров Сергей Юрьевич

Библиотекарь 1 категории ГПНТБ СО РАН

Отдел создания, развития и продвижения информационно-библиографических ресурсов

Группа информационных систем и электронных ресурсов

• skomarov87@gmail.com