

УДК 004.738

ГРНТИ 50.39.17. Режимы и дисциплина взаимодействия элементов ВС.
Алгоритмы и протоколы ВС.

Зиен Роман Маратович

Студент

Санкт-Петербургский горный университет

Россия, г. Санкт-Петербург

АНАЛИЗ АЛГОРИТМА ПОИСКОВОЙ ИНФОРМАЦИИ

Аннотация

SEO с английского переводится как «поисковая оптимизация». Это процесс получения трафика из бесплатных, органических, редакционных или естественных результатов поиска в поисковых системах. Чем раньше (или выше рейтинг на странице результатов поиска) и чаще сайт появляется в списке результатов поиска, тем больше посетителей он получит от пользователей поисковой системы, и эти посетители могут быть преобразованы в клиентов. SEO может быть нацелено на различные виды поиска, включая поиск изображений, локальный поиск, поиск, поиск видео, академический поиск, поиск новостей и отраслевые вертикальные поисковые системы. Как стратегия интернет-маркетинга, SEO учитывает, как работают поисковые системы, что ищут люди, фактические поисковые термины или ключевые слова, введенные в поисковые системы, и какие поисковые системы предпочитают их целевая аудитория. Оптимизация веб-сайта может включать редактирование своего содержимого, HTML и связанного кода, чтобы повысить его актуальность для конкретных ключевых слов и устранить препятствия для индексирования поисковых систем. Продвижение сайта для увеличения количества обратных ссылок или входящих ссылок - еще одна тактика SEO. По состоянию на май 2020 года мобильный поиск наконец-то превзошел поиск на компьютере. Google разрабатывает и продвигает мобильный поиск как будущее во всех

своих продуктах, и многие бренды начинают использовать другой подход к своим интернет-стратегиям.

Ключевые слова: SEO, HTML, факторы, объем, сканирование.

Zien Roman Maratovich
Student
Saint-Petersburg Mining University
Russia, St. Petersburg

ANALYSIS OF SEARCH ENGINE OPTIMIZATION ALGORITHM

Abstract

SEO stands for “search engine optimization.” It is the process of getting traffic from the free, organic, editorial or natural search results on search engines. The earlier (or higher ranked on the search results page), and more frequently a site appears in the search results list, the more visitors it will receive from the search engine's users, and these visitors can be converted into customers. SEO may target different kinds of search, including image search, local search, video search, academic search, news search and industry-specific vertical search engines. As an Internet marketing strategy, SEO considers how search engines work, what people search for, the actual search terms or keywords typed into search engines and which search engines are preferred by their targeted audience. Optimizing a website may involve editing its content, HTML and associated coding to both increase its relevance to specific keywords and to remove barriers to the indexing activities of search engines. Promoting a site to increase the number of backlinks, or inbound links, is another SEO tactic. As of May 2020, mobile search has finally surpassed desktop search, Google is developing and pushing mobile search as the future in all of its products and many brands are beginning to take a different approach on their internet strategies.

Keywords: SEO, HTML, Factors, Volume, Crawl.

I. ВСТУПЛЕНИЕ

SEO - это маркетинговая дисциплина, ориентированная на рост видимости в обычных (неоплачиваемых) результатах поиска. SEO включает в себя как технические, так и творческие элементы необходимые для повышения рейтинга, привлечения трафика и увеличения осведомленности в поисковых системах [1]. Есть много аспектов SEO, от слов на странице до других сайтов и ссылок на сайт в Интернете. Иногда SEO - это просто способ, чтобы убедиться, что ваш сайт структурирован таким образом, что поисковые системы разбираются и находят его. SEO – это не только лишь создание веб-сайтов, удобных для систем поиска. Речь идет о том, чтобы сделать любой сайт лучше для людей. Основной целью создания сайта любой компанией является информирование о компании и ее продуктах / услугах. Для этого интернет-маркетинг играет важную роль в привлечении посетителей на сайт с помощью методов SEO. Теперь становится важным, чтобы сайт был указан на верхних страницах поисковых систем и опережал сайт конкурента [9].

Процесс привлечения внимания поисковых систем веб-сайтов требует особого персонализированного обслуживания. Для этого применяются соответствующие инструменты и методы SEO, чтобы достичь более высокого рейтинга и сделать поиск по сайту эффективным. Хотя основной принцип работы всех поисковых систем одинаков, незначительные отличия между ними приводят к серьезным изменениям релевантности результатов. Для разных поисковых систем учитываются разные факторы [1].

II. КАК РАБОТАЕТ СИСТЕМА ПОИСКА

Первая основная истина, которую вам нужно знать, чтобы изучить SEO, заключается в том, что поисковые системы - это не люди. Хотя это может быть очевидно для всех, разница между тем, как люди и поисковые системы не просматривают веб-страницы. В отличие от людей, поисковые системы анализируют текст. Хотя технологии быстро развиваются, поисковые системы

далеки от умных существ, способных почувствовать красоту крутого проектирования или наслаждаться звуками и движением в фильмах. Вместо этого поисковые системы сканируют Интернет, глядя на определенные элементы сайта (в основном текст), чтобы понять, что это за сайт. Это краткое объяснение и не самое точное, потому что, как мы увидим дальше, поисковые системы выполняют несколько действий для предоставления результатов поиска - сканирование, индексирование, обработка, вычисление релевантности и получение данных [2].

Во-первых, поисковые системы просматривают Интернет, чтобы увидеть, что там есть. Эта задача выполняется программным обеспечением, называемым краулер или паук (или бот Google, как в случае с Google). Пауки переходят по ссылкам с одной страницы на другую и проиндексируют все, что попадется на пути. Имея в области анализа количество страниц в Интернете (более 20 миллиардов), для паука невозможно посещать сайт ежедневно, чтобы просто посмотреть, не появилась новая страница или существующая страница была изменена, иногда поисковые роботы могут не посещать свой сайт на месяц или два [10].

После того, как страница просканирована, следующим шагом будет индексирование ее содержания. Проиндексированная страница хранится в гигантской базе данных, откуда ее впоследствии можно будет извлечь. По сути, процесс индексации - это определение слов и выражений, которые лучше всего описывают запрос и присваивают странице определенные ключевые слова. Для человека не сложно обрабатывать такие объемы информации, но в целом поисковые системы прекрасно справляются с этой задачей [8].

Иногда они могут неправильно понять смысл страницы, но, если человек помогает им, оптимизируя его, будет легче правильно классифицировать страницы, и – получать более высокие рейтинги. Когда приходит поисковый запрос, поисковой движок обрабатывает его - то есть сравнивает строку поиска в поисковом запросе с проиндексированными страницами в базе данных. Поскольку вполне вероятно, что более одной страницы (практически это

миллионы страниц) содержит строку поиска, движок поиска начинает вычислять релевантность каждой из страниц в своем индексе со строкой поиска. Есть разные алгоритмы расчета релевантности. Каждый из них имеет разную относительную частоту для таких общих факторов, как плотность ключевых слов, ссылки или встреченные теги. Это причина, почему разные поисковые системы дают разные результаты поиска страницы для одной и той же строки поиска.

Более того, это известный факт, что все основные поисковые системы, такие как Yahoo!, Google, Bing и др. периодически меняют свои алгоритмы.

Последний шаг в деятельности поисковых систем – получение данных (результатов). По сути, это не более чем просто отображение их в браузере – т.е. бесконечные страницы результатов поиска отсортированы от наиболее релевантных до наименее релевантных сайтов [2,3].

III. СОДЕРЖАНИЕ SEO ДЛЯ ДВИЖКОВ ПОИСКА

SEO больше не означает разброс ключевых слов. Новейшие поисковые системы сканируют страницы почти так же, как читатели. Якоб Нильсен, исследователь и эксперт по человеко-машинному взаимодействию Технического университета Копенгагена обнаружил, что почти 80 процентов посетителей веб-сайта сканировали страницу быстрее чем читали его построчно. Они провели свои первые доли секунды на странице, решая, стоило ли это их времени. Программисты поисковых систем до сих пор используют это исследование, чтобы разрабатывать алгоритмы, обеспечивающие более значимые рейтинги. То же самое, что у посетителей поисковой системы – то что привлечет внимание глаз. Вверху слева – угол страницы – самая ценная часть страницы, так как на нее в первую очередь обращаются глаза читателя. Стоило бы поставить важный текст там, чтобы поисковые системы и люди видели его немедленно. Также хорошее место для текста – в рамке и детализированные списки, оба из которых в равной степени обращают на себя внимание.

Жирный текст заставляет людей и машины замечать его, но стоит использовать эти теги разумно. Тот же жирный текст выглядит как реклама и вызовет поисковые системы, чтобы обесценить ваш сайт [4].

Курсив и полужирный HTML-теги должны окружать значимые концепции, а не акценты слова. Выделение жирным шрифтом или курсивом означает ничего для поисковой системы, поэтому лучше применять эти теги к важным понятиям и подзаголовкам [3].

Поиски теперь используется для связанных терминов и релевантных фраз, а не только ключевых слов. Человек улавливает смысл из контекста и легко различает термин "вырезка" применительно к вопросу от того же слова, которое относится к киноленте или видеоигровой графике. Включите в свой SEO-текст синонимы и релевантные термины, позволяющие поисковым системам распознать цель вашего сайта. К счастью, есть способ работать с этими условиями в контенте без отслеживания ключевых слов и ключей в процентах: просто написать, что интересно, скопировать то, что люди любят читать. Если писать для читателей, то поисковые системы последуют этому [5].

IV. ФАКТОРЫ УСПЕХА SEO [3]

Поисковая система подразделяется на три основные группы.

Периодическая таблица факторов успеха SEO:

- SEO на странице
- SEO вне страницы
- Нарушения

Внутри каждой группы есть подгруппы, поскольку каждая глава этого Руководство по SEO объясняет. Эти подгруппы содержат один или больше индивидуальных факторов SEO с определенным весом или важностью. Нарушения, хотя и представляют собой самостоятельную группу, являются группой и подгруппой, с которой они связаны.

V. ПРОЦЕДУРА SEO

А. Стратегии начальной настройки

Б. Ключевые слова и конкурентный анализ

С. Оптимизация сайта

1. Оптимизация сайта на странице

А. Идентификация ключевого слова на странице

Б. Улучшение содержания по целевому ключевому слову

С. Заголовок / мета-теги и описание

Д. Процесс названия файлов

Е. Анализ веб-страницы

2. Оптимизация сайта вне страницы

А. Создание ссылок

Д. Платный маркетинг / продвижение

1. Кампания с оплатой за клик

2. Партнерский маркетинг

Е. Отчеты о проверке местоположения и позиционировании



Рис 1: Жизненный цикл SEO

VI. SEO ПОИСК - МАРКЕТИНГ

Ни один фактор SEO не гарантирует поисковой системе рейтинги [4,11]. Хороший HTML-заголовок не поможет, если страница имеет некачественный контент. Наличие большого количества ссылок не поможет, если все они низкого качества. Имея несколько положительных факторов можно увеличить шансы на успех, в то время как наличие негативных факторов могут ухудшить эти шансы.

Чтобы избежать нежелательного содержания в поисковых индексах, веб-мастера могут указать паукам не сканировать определенные файлы или каталоги через стандартный файл robots.txt в папке корневого каталога домена. Дополнительно страницу можно исключить из базы данных поисковой системы, используя метатег, специфичный для роботов. Когда движок поиска посещает сайт, файл robots.txt находится в корне directory - это первый просканированный файл. После того как файл robots.txt проанализирован, движок инструктирует робота о том, какие страницы не показывать. Сканер поисковой системы может сохранять кэшированные копии этого файла, он может иногда сканировать страницы, которые веб-мастер не дает сканировать [6]. Страницы, сканирование которых обычно запрещено, включая страницы для входа в систему, например, корзины покупок и пользовательский контент, например, результаты внутреннего поиска. В марте 2007 г. Google предупредил веб-мастеров, что они должны предотвратить индексирование результатов внутреннего поиска, потому что эти страницы считаются поисковым спамом[7].

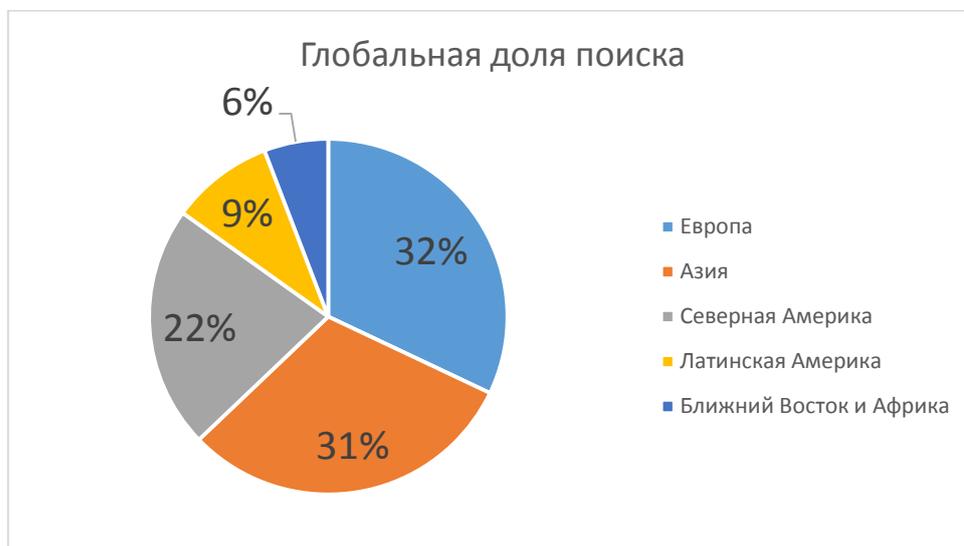


Рис.2 . Поисковые истории

VII. ЗАКЛЮЧЕНИЕ

Доступность информации в Интернете растет резко. Но будет полезно, если пользователь найдет правильную информация за меньшее время. Итак,

выделиться среди конкурентов - означает увеличить видимость сайтов, поисковая оптимизация (SEO) играет важную роль.

SEO - это рентабельный процесс продвижения в Интернете. Это помогает привлечь клиентов на сайт.

SEO может осуществляться двумя способами - на странице и вне страницы. Оптимизация на странице может выполняться на странице, например, проверять битые теги, ссылки, проверять метатеги, альтернативные теги и т. д.

Оптимизация вне страницы означает предоставление качественных обратных ссылок на конкретную веб-страницу для улучшения рейтинга веб-сайта.

SEO - это процесс, состоящий из нескольких этапов: исследование, постановка цели, оптимизация страницы, контент-разработка, создание ссылок, а затем отслеживание всего процесса.

В этой статье показано SEO-исследование, которое помогло повысить рейтинг веб-сайта и, таким образом, получить веб-сайт продвижение. Так что этот документ можно использовать как основу для SEO.

Эффективное SEO требует управления тремя основными пунктами: содержание, ключевое слово и построение ссылок. С помощью этих трех главных пунктов можно создать базовую структуру SEO, что будет поддерживать высокий рейтинг веб-страниц в будущем.

Использованные источники

1. S.Mukherjee, "A probabilistic model for optimal searching of the deep Web", 2003.
2. Fuxue Wang; Yi Li; Yiwen Zhang; Coll. of Econ. & manage, "An empirical study on the search engine optimization technique and its outcomes," Artificial Intelligence, Management Science and Electronic Commerce AIMSEC), Aug. 2011.
3. Ma Xiao ling, Wu Yong he." For research of search engine optimization (SEO)". Information Retrieval. 2005, 12(1):6~ 10

4. Graf, Cousin. Search Engine Optimization. Beijing: Qinghua University Press, 2007.
5. Liu Xiaolin." Chinese search engine key technologies". Hubei University of Technology Reported.2006, 6(1):9~12
6. Yi Dan. 'Search engine optimization (SEO) Research and Implementation". Petroleum Industry Computer Applications. 2006, (1):132~ 134
7. Grappone J, Couzin G. Search Engine Optimization: An Hour a Day. Yang Mingjun translation. Beijing: Tsinghua University Press, 2007.
8. Fei Wei, Huang Ruhua. "Based on user behavior analysis of search engine optimization strategy". Library and Information Service, 2005 (10):314~320
9. Прохорова А.М. Регистрация сервера в поисковых системах как часть инструмента поисковой оптимизации сайта / А. М. Прохорова // Наука, техника и образование. 2016. №8 (26).
10. Чамина О. Г., Бексаева Е. А. SEO-аналитика веб-ресурсов электронной коммерции / О. Г. Чамина, Е. А. Бексаева // Модели, системы, сети в экономике, технике, природе и обществе. 2015. №3 (15).
11. Яковлев А. А. Раскрутка и продвижение сайтов: основы, секреты, трюки / А. А. Яковлев — СПб.: БХВ-Петербург, 2007. — 336 с.