

Математическая модель расчета интегральной оценки качества внутренней поисковой оптимизации Web-ресурсов

Николенко Дмитрий Михайлович

Волжский политехнический институт (филиал) ФГБОУ ВО «Волгоградский государственный технический университет»

Студент

Лясин Дмитрий Николаевич

Волжский политехнический институт (филиал) ФГБОУ ВО «Волгоградский государственный технический университет»

к.т.н., доцент, доцент кафедры информатики и технологии программирования

Аннотация

В данной статье была представлена и обоснована математическая модель расчета интегральной оценки качества внутренней поисковой оптимизации веб-ресурсов. Для расчета оценки были отобраны и рассмотрены 10 наиболее значимых факторов внутренней поисковой оптимизации, а также, с помощью иерархической процедуры Т. Саати, был определен их вес. Далее приводится результат тестирования модели на специально подготовленном веб-ресурсе и подробности расчетов.

Ключевые слова: поисковая оптимизация, оценка качества внутренней поисковой оптимизации, факторы внутренней поисковой оптимизации

Mathematical model for calculating the integral quality assessment of internal search engine optimization for Web resources

Nikolenko Dmitry Mikhailovich

Volzhsy Polytechnical Institute (branch) of Volgograd State Technical University Student

Lyasin Dmitriy Nikolayevich

Volzhsy Polytechnical Institute (branch) of Volgograd State Technical University Candidate of Technical Sciences, docent of «Informatics and programming technology»

Abstract

This article presents and substantiates a mathematical model for calculating the integral quality assessment of the internal search engine optimization of web resources. To calculate the estimates, 10 of the most significant factors of internal search engine optimization were selected and considered, and their weight was determined using the hierarchical procedure of T. Saati. The following is the result

of testing the model on a specially prepared web resource and details of the calculations.

Keywords: search engine optimization, internal search engine optimization quality assessment, internal search engine optimization factors

ВВЕДЕНИЕ

В современном мире, Интернет является, наверное, наиболее важной площадкой для продвижения товаров и услуг. Многие владельцы веб-ресурсов стремятся заполнить верхние строчки в результатах поисковой выдачи, так как поисковые системы могут помочь с генерацией трафика и привлечением потенциальных клиентов. Чтобы облегчить эту задачу применяются программы для работы с поисковой оптимизацией (SEO). Таким образом достигается высокая конкуренция как среди владельцев веб-ресурсов, так и среди разработчиков SEO-программ. В настоящее время, большинство SEO-программ способны обнаруживать проблемы веб-ресурса, но, в тоже время, они не способны дать ответ, находится ли его поисковая оптимизация на желаемом уровне. Этим и обусловлена необходимость создания модели расчета интегральной оценки качества поисковой оптимизации. Объектом расчета оценки была выбрана внутренняя поисковая оптимизация, так как она лучше поддается анализу.

МОДЕЛЬ РАСЧЕТА ИНТЕГРАЛЬНОЙ ОЦЕНКИ КАЧЕСТВА ВНУТРЕННЕЙ ПОИСКОВОЙ ОПТИМИЗАЦИИ

Для расчета интегральной оценки качества внутренней поисковой оптимизации Q , автором статьи была предложена формула:

$$Q = \sum_{i=1}^n (f_i * k_i),$$

где f_i — это вес фактора оптимизации, а k_i — коэффициент этого фактора.

Значение всех коэффициентов изначально равно нулю и увеличивается при соблюдении некоторых условий. Соответственно несоблюдение этих условий на исследуемой странице негативно скажется на итоговой оценке. Максимальное значение коэффициента — 1, минимальное — 0.

Модель опирается на следующие факторы:

Статус ссылок

Ссылки играют решающую роль при вычислении PageRank — главного алгоритма Google для ранжирования веб-сайтов. Именно поэтому данный фактор является самым важным. За каждую ссылку со статусом ОК, значение коэффициента увеличивается на 1. За ссылку со статусом Redirect — на 0.5, т.к. некоторые (но не все) поисковые системы могут относиться к редиректам негативно. За ссылку со статусом 404, значение коэффициента увеличиваться не будет, т.к. поисковые системы однозначно негативно относятся к битым ссылки. После того, как последняя ссылка была проверена, полученное значение коэффициента делится на общее количество ссылок. Таким образом коэффициент приобретает итоговое значение.

Анкорный текст ссылки

Анкорный текст ссылки способен увеличить ее CTR (скорость клика). Это фактор средней важности. За наличие анкорного текста, значение коэффициента увеличивается на 1, за отсутствие — не увеличивается. После того, как последняя ссылка была проверена, полученное значение коэффициента делится на общее количество ссылок. Таким образом коэффициент приобретает итоговое значение.

Ключевые слова в анкорном тексте ссылок

Этот фактор имеет низкую важность, т.к. нет прямого доказательства того, что для поисковых систем имеет большое значение находятся ли ключевые слова в обычном тексте или в анкорном тексте ссылки. При условии наличия ключевых слов в анкорном тексте хотя бы 2/3 от общего числа ссылок, коэффициент становится равным 1. В ином случае — 0.

Классическая тошнота текста

Обычно рассчитывается для слова, которое используется чаще всего. Для того, чтобы посчитать классическую тошноту нужно подсчитать сколько раз самое популярное слово употребляется на странице и извлечь из этого числа корень:

$$V_{\text{кл.}} = \sqrt{N_{\text{сл.}}},$$

где $N_{\text{сл.}}$ — количество употреблений слова на странице.

Этот фактор имеет низкую, т.к. при расчете классической тошноты не учитывается объем текста. Если тошнота находится в оптимальных пределах, коэффициент становится равным 1, если превышает — 0.5, если ниже — 0.

Академическая тошнота текста

Академическая тошнота — это отношение количества употребления слова к общему количеству слов в тексте, выраженное в процентах. По сути, академическая тошнота — это частота употребления слова на странице.

$$V_{\text{ак.}} = \frac{N_{\text{сл.}}}{N_{\text{общ.}}} * 100$$

Этот фактор является важным, т.к. при проведении работ по оптимизации важно учитывать показатели тошноты. Например, если ключевые слова присутствуют на странице в недостаточном количестве, поисковая система может решить, что они употребляются эпизодически и не несут существенной смысловой нагрузки. В тоже время избыточная тошнота приведет к тому, что текст станет практически нечитаемым.

Если тошнота находится в оптимальных пределах, коэффициент становится равным 1, если превышает — 0.5, если ниже — 0.

Наличие важных тегов

К важным тегам относятся теги: title, meta-title, description, h1. Поисковые системы очень внимательно относятся к этим тегам и показывают их содержимое в результатах поисковой выдачи, поэтому этот фактор является важным. Наличие каждого из этих тегов обязательно. За каждый из

этих тегов, значение коэффициента увеличивается на 1. Далее сумма делится на 4 (количество важных тегов).

Длина важных тегов

Это фактор средней важности, т.к. если длина тегов превышает допустимые пределы, содержимое не поместится в результаты поисковой выдачи целиком. Если длина находится в оптимальных пределах, значение коэффициента увеличивается на 1, если превышает — на 0.5, если ниже — на 0.25, если отсутствует — не увеличивается. Далее сумма делится на 4 (количество важных тегов).

Ключевые слова важных тегов

Этот фактор немного важнее предыдущего. Так как поисковые системы отображают содержимое этих тегов в результатах поисковой выдачи, наличие ключевого слова там может увеличить CTR (скорость клика). За наличие ключевого слова в теге, значение коэффициента увеличивается на 1. Далее сумма делится на 4 (количество важных тегов).

Наличие атрибута alt

Важный фактор, т.к. большинство поисковых систем уделяет внимание данному атрибуту во время индексации, благодаря чему, поиск картинки может привести пользователя на сайт. За наличие атрибута alt, значение коэффициента увеличивается на 1. Далее сумма делится на общее количество изображений.

Соотношение текста к коду

Этот фактор имеет низкую важность. Если соотношение текста к коду больше 10%, коэффициент становится равным 1, если меньше — это будет косвенно говорить о малом количестве контента.

Определим веса факторов с учетом их важности. Для этого воспользуемся аналитической иерархической процедурой Т. Саати [1].

Для начала добавим обозначения факторам:

- 1) A1 — Статус ссылок;
- 2) A2 — Анкорный текст ссылок;
- 3) A3 — Ключевые слова ссылок;
- 4) A4 — Академическая тошнота;
- 5) A5 — Классическая тошнота;
- 6) A6 — Наличие важных тегов;
- 7) A7 — Длина важных тегов;
- 8) A8 — Ключевые слова важных тегов;
- 9) A9 — Наличие атрибута alt;
- 10) A10 — Соотношение текста к коду.

Далее построим матрицу парных сравнений [2,3,4].

Правила заполнения матрицы парных сравнений указаны в таблице 1.

Таблица 1 — Значимость коэффициентов матрицы парных сравнений

X_{ij}	Значимость
1	i-ый и j-ый критерий одинаково важен
3	i-ый критерий незначительно преимущественен j-го
5	i-ый критерий преимущественен j-го
7	i-ый критерий значительно преимущественен j-го
9	i-ый критерий явно преимущественен j-го

Матрица парных сравнений, средние геометрические и веса факторов представлены в таблице 2.

Таблица 2 — Матрица парных сравнений

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	Ср. геом.	Веса крит.
A1	1.00	7.00	9.00	5.00	9.00	5.00	9.00	7.00	5.00	9.00	5.76	0.36
A2	0.14	1.00	5.00	0.33	5.00	0.33	3.00	1.00	0.33	5.00	1.07	0.07
A3	0.11	0.20	1.00	0.20	3.00	0.20	0.33	0.20	0.20	3.00	0.40	0.03
A4	0.20	3.00	5.00	1.00	7.00	1.00	5.00	3.00	0.33	7.00	1.93	0.12
A5	0.11	0.20	0.33	0.14	1.00	0.14	0.20	0.20	0.20	1.00	0.26	0.02
A6	0.20	3.00	5.00	1.00	7.00	1.00	5.00	3.00	0.33	7.00	1.93	0.12
A7	0.11	0.33	3.00	0.20	5.00	0.20	1.00	0.33	0.20	5.00	0.61	0.04
A8	0.14	1.00	5.00	0.33	5.00	0.33	3.00	1.00	0.33	5.00	1.07	0.07
A9	0.20	3.00	5.00	3.00	5.00	3.00	5.00	3.00	1.00	5.00	2.52	0.16
A10	0.11	0.20	0.33	0.14	1.00	0.14	0.20	0.20	0.20	1.00	0.26	0.02
Сумма											15.81	1.00

Диаграмма весовых коэффициентов для факторов A1, A2, A3, A4, A5, A6, A7, A8, A9, A10 представлена на рисунке 1.

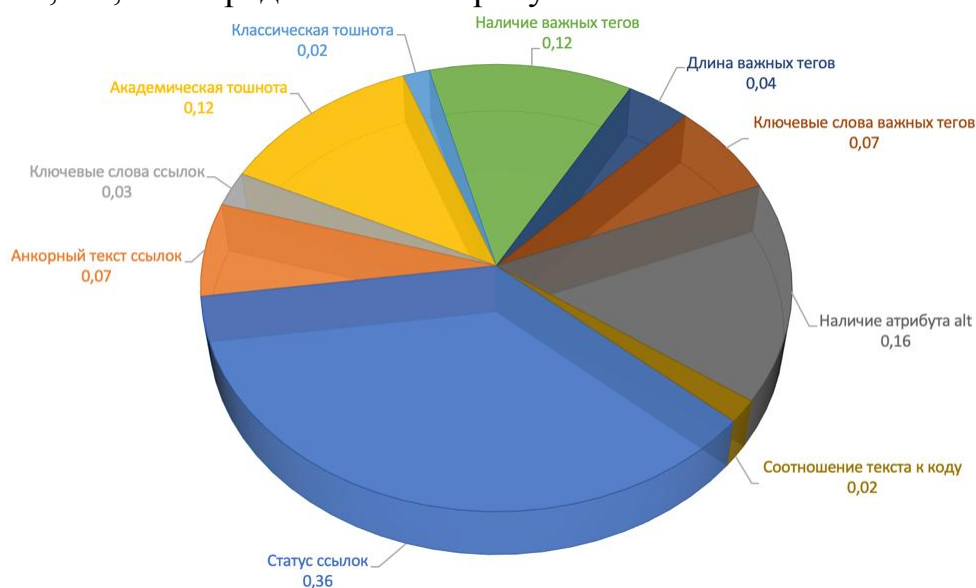


Рисунок 1 — Диаграмма весовых коэффициентов

Проведем проверку матрицы парных сравнений на непротиворечность [2,3,4].

Суммы столбцов матрицы парных сравнений:

$R1 = 2.33; R2 = 18.93; R3 = 38.67; R4 = 11.35; R5 = 48.00; R6 = 11.35; R7 = 31.73; R8 = 18.93; R9 = 8.13; R10 = 48.00.$

Рассчитаем вспомогательную величину L , просуммировав произведения сумм столбцов матрицы на весовые коэффициенты: $L = 11.25$.

Индекс согласованности: $ИС = (L-N)/(N-1) = 0.14$.

Величина случайной согласованности: $СлС = 1.12$.

Отношение согласованности $ОС = ИС/СлС = 0.12$, т.е. не превышает 0.2, из чего следует, что уточнение матрицы парных сравнений не требуется.

Значения весовых коэффициентов:

- 1) Статус ссылок: $A1 = 0.36$;
- 2) Анкорный текст ссылок: $A2 = 0.07$;
- 3) Ключевые слова ссылок: $A3 = 0.03$;
- 4) Академическая тошнота: $A4 = 0.12$;
- 5) Классическая тошнота: $A5 = 0.02$;
- 6) Наличие важных тегов: $A6 = 0.12$;
- 7) Длина важных тегов: $A7 = 0.04$;
- 8) Ключевые слова важных тегов: $A8 = 0.07$;
- 9) Наличие атрибута alt: $A9 = 0.16$;
- 10) Соотношение текста к коду: $A10 = 0.02$.

, где $\sum a_i = 1$.

Немного скорректируем значения и умножим их на 100, для большего удобства при работе с ними.

Итоговые веса факторов указаны в таблице 3.

Таблица 3 — Факторы для расчета оценки качества внутренней поисковой оптимизации

№	Название фактора	Вес
1	Статус ссылок	35
2	Анкорный текст ссылок	7
3	Ключевые слова ссылок	3
4	Академическая тошнота	12
5	Классическая тошнота	2
6	Наличие важных тегов	12
7	Длина важных тегов	5
8	Ключевые слова важных тегов	7
9	Наличие атрибута alt	15
10	Соотношение текста к коду	2
Сумма		100

В соответствии с моделью, максимальная оценка качества внутренней поисковой оптимизации составляет 100 баллов, а минимальная — 0.

В связи с этим выделим уровни оптимизации:

- 1) Плохая оптимизация (от 0 до 29 баллов);
- 2) Слабая оптимизация (от 30 до 49 баллов);
- 3) Нормальная оптимизация (от 50 до 69);
- 4) Хорошая оптимизация (от 70 до 89 баллов);
- 5) Отличная оптимизация (от 90 до 100 баллов).

Для тестирования модели было разработано программное средство.

Тестирование проводилось на специально подготовленном сайте, обладающем следующими характеристиками:

общее количество ссылок: 42; ссылок со статусом ОК: 35; ссылок со статусом Redirect: 3; ссылок со статусом 404: 4; ссылок с анкорным текстом: 38; ссылок с ключевыми словами больше чем 2/3; классическая тошнота превышает оптимальные пределы; академическая тошнота находится в оптимальных пределах; важных тегов: 3; 1 из важных тегов имеет оптимальную длину, а 2 превышают; 2 важных тега содержат ключевые слова; общее количество изображений: 15; количество изображений с атрибутом alt: 7; соотношение текста к коду является оптимальным.

Результаты тестирования представлены на рисунке 2.

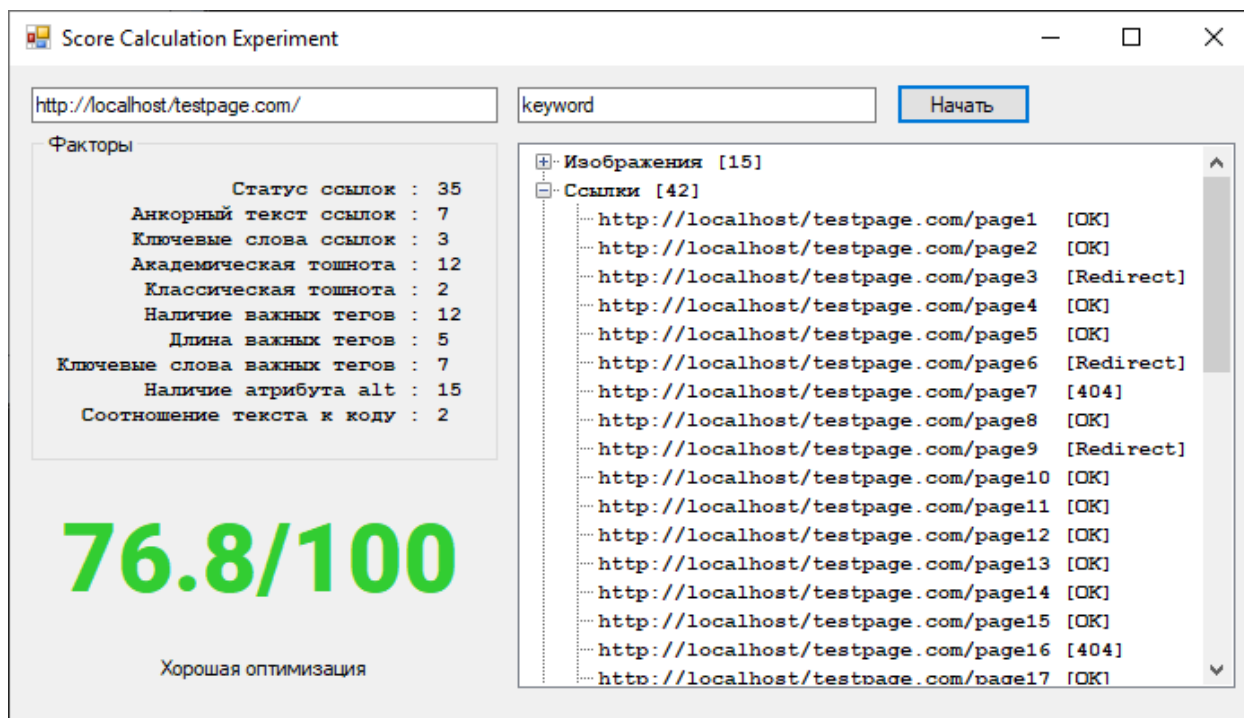


Рисунок 2 — Программное средство для расчета интегральной оценки качества внутренней поисковой оптимизации

Подробные расчеты значений коэффициентов и интегральной оценки представлены ниже.

Расчеты значений коэффициентов:

Статус ссылок: $\frac{35*1+3*0.5+4*0}{42} = 0.87;$

Анкорный текст ссылок: $\frac{38}{42} = 0.9;$

Ключевые слова ссылок: 1;

Академическая тошнота: 1;

Классическая тошнота: 0.5;

Наличие важных тегов: $\frac{3}{4} = 0.75;$

Длина важных тегов: $\frac{1+0.5+0.5+0}{4} = 0.5;$

Ключевые слова важных тегов: $\frac{2}{4} = 0.5;$

Наличие атрибута alt: $\frac{7}{15} = 0.47;$

Соотношение текста к коду: 1.

Расчет интегральной оценки:

$Q = 35 * 0.87 + 7 * 0.9 + 3 * 1 + 12 * 1 + 2 * 0.5 + 12 * 0.75 + 5 * 0.5 + 7 * 0.5 + 15 * 0.47 + 2 * 1 = 76.8,$
что можно классифицировать как хорошую оптимизацию.

ЗАКЛЮЧЕНИЕ

Интегральная оценка качества внутренней поисковой оптимизации позволяет получить наглядное представление о том, находится ли поисковая оптимизация веб-ресурса на желаемом уровне и этим самым повысить эффективность принимаемых владельцем веб-ресурса решений.

Библиографический список

1. Кондрацкий Д.Е., Рыбанов А.А. Исследование методов и алгоритмов автоматизированной системы оценки альтернативных вариантов методом Т.Саати // NovaInfo.Ru. 2016. Т. 3. № 46. С. 107-116.
2. Морозов А.О., Рыбанов А.А. Экспертная оценка программных продуктов для расчета метрических характеристик физической схемы базы данных // Современные научные исследования и инновации. 2015. № 1. Ч.1. URL: <http://web.snauka.ru/issues/2015/01/42101> (дата обращения: 29.09.2017).
3. Рыбанов А.А. Определение весовых коэффициентов сложности тем учебного курса на основе алгоритма Саати // Педагогические измерения. 2014. № 4. С. 21-28.
4. Рыбанов А.А., Макушкина Л.А. Технология определения весовых коэффициентов сложности тем дистанционного курса на основе алгоритма Саати // Открытое и дистанционное образование. 2016. № 1 (61). С. 69-79.