



Web Effector



НЮАНСЫ АЛГОРИТМОВ РАНЖИРОВАНИЯ ПОИСКОВЫХ СИСТЕМ

Сергей ЛЮДКЕВИЧ, начальник отдела аналитики и исследований



ФОРМУЛЫ РЕЛЕВАНТНОСТИ



ЯНДЕКС

Формулы на основе машинного обучения (технология MatrixNet).

420+ факторов (апрель 2010 г.), их комбинации (мономы)

Отдельные формулы для регионов (Москва, Санкт-Петербург, 16 крупных городов России, Украина, Белоруссия, Казахстан).

ФОРМУЛЫ РЕЛЕВАНТНОСТИ



GOOGLE

Формулы – ручная настройка или машинное обучение –
?

Peter Norvig, директор по исследованиям Google, май 2008:

1. Люди верят, что могут сделать лучше машины
2. «Ручная» модель менее подвержена катастрофическим ошибкам

200+ факторов (signals)

<http://www.google.com/corporate/tech.html>

Отдельные формулы для стран.

КЛАССИФИКАЦИЯ ФАКТОРОВ



1. Внутренние
 - Текстовые
 - Качество контента
 - Свойства сайта/страницы
2. Внешние (ссылочные)
 - Ссылочные статические (не зависящие от запроса)
 - Ссылочные динамические (зависящие от запроса)
3. Пользовательские (поведенческие)
4. Региональные
5. Запросные

ПЕССИМИЗАЦИЯ / PENALTIES



- Пред-фильтры

дисконтирование или обнуление значений некоторых факторов

- Пост-фильтры (штрафы)

дисконтирование значения функции релевантности

- **Аффилированность сайтов**

группировка в результатах поиска с показом наиболее релевантного аффилиата (Яндекс)

- **Фильтр на одинаковые сниппеты**

- **Негативные факторы ранжирования**

Факторы за качество контента

Факторы за вторичность контента

Факторы за качество входящих ссылок

Портяночный фильтр (текстовый переспам)

- направлен против чрезмерной концентрации ключевых слов в контенте («портянок»)
- запросозависим
- зависит от характеристик документа

Рекомендации:

- Избегать многократного употребления ключевых слов из запроса в одном предложении.
- Употреблять ключевые слова как можно ближе к началу документа.
- Избегать участков текста с плотными скоплениями ключевых слов

Штраф за переоптимизацию

- направлен против манипуляций текстовыми факторами ранжирования
- запросонезависим
- зависит от характеристик документа

Рекомендации:

- Избегать использования тегов физической и логической разметки ``, ``, ``, `<i>`, а также тегов заголовков `<h1>`, ...`<h6>`.
- Использовать «человеко-ориентированные» варианты тега `<title>`.



Качество контента

- Уникальность контента
- Вторичность контента
- Естественность контента
- Порнографическая лексика

СВОЙСТВА САЙТА/СТРАНИЦЫ



- **Возраст сайта/страницы**

фактор ранжирования в Яндексе, «песочница» в Google

- **Наличие попандеров/кликандеров**

пессимизация в Яндексе

- **Формат документа**

HTML ранжируется лучше остальных (pdf, doc, xls и пр.)

- **Употребление ключевых слов в URL**

положительный фактор ранжирования

- **«Плохая наследственность»**

IP, данные whois – принадлежность к «плохим ребятам»

ПЕССИМИЗАЦИЯ:

- «Ссылочный взрыв»
- Плохие ссылки
- Коммерческие ссылки
- Участие в линкобменниках (link-farms)

РЕГИОНАЛЬНОСТЬ



- Выбор соответствующего домена 1-го уровня для сайта (напр., для России – .ru)
- Присвоение сайту требуемого региона
- Употребление названия требуемого региона в тексте документа
- Получение входящих ссылок с сайтов требуемого региона
- Употребление названия требуемого региона в текстах входящих ссылок на документ
- Добавление адреса организации в требуемом регионе

ПОЛЬЗОВАТЕЛЬСКИЕ ФАКТОРЫ → → → → →



КЛИКАБЕЛЬНОСТЬ САЙТА/СТРАНИЦЫ В ВЫДАЧЕ

- CTR
- временные показатели сессии
- количественные показатели сессии

ПОСЕЩАЕМОСТЬ САЙТА/СТРАНИЦЫ

- общая
- по видам источников трафика – поисковый, ссылочный, закладочный и т.п.

ПОЛЬЗОВАТЕЛЬСКИЕ ФАКТОРЫ → → → → →



ПОВЕДЕНИЕ ПОЛЬЗОВАТЕЛЯ НА САЙТЕ/СТРАНИЦЕ

- активные действия пользователя
- проведенное время
- внутренние переходы
- переходы с сайта на другие сайты

СРЕЗЫ ФАКТОРОВ ПО ТИПУ ПОЛЬЗОВАТЕЛЕЙ

- пришедшие по данному по запросу
- из данного региона

УЛУЧШЕНИЕ ПОЛЬЗОВАТЕЛЬСКИХ ФАКТОРОВ РАНЖИРОВАНИЯ



- Целевая страница должна быть лучшим ответом на запрос
- Целевая страница должна побуждать пользователя к выполнению действий
- Сайт должен внушать доверие пользователю
- Сайт должен быть удобным для пользователя
- Привлекательность сниппетов
- Кросс-браузерность
- Корректная кодировка
- Стабильная работа сайта

ПРИМЕСИ К ОРГАНИЧЕСКИМ РЕЗУЛЬТАТАМ ПОИСКА



ЯНДЕКС

- «Спектральная» примесь
- Быстроботовская примесь

GOOGLE

- Примесь из «Адресов»

→ СПАСИБО ЗА
ВНИМАНИЕ! → → → →



Для продолжения темы посетите

www.webeffector.ru

115280, Россия, Москва,
ул. Ленинская Слобода, д. 19, стр. 1, бизнес-
центр «Омега Плаза»

Телефон: (495) 772-97-91 (многоканальный)