

Информационные технологии

УДК 681.3

РАЗРАБОТКА МЕТОДОВ ОПТИМИЗАЦИИ ИНТЕРНЕТ-РЕКЛАМЫ НА ОСНОВЕ КЛЮЧЕВЫХ СЛОВ

Д.А. Ёлкин¹, И.А. Минаков², С.И. Вольман¹

¹ ООО НПК «Максифаер Девелопмент»
Россия, 443125, г. Самара, ул. Ново-Садовая, 349, 3-й этаж

² Институт проблем управления сложными системами РАН
Россия, 443020, г. Самара, ул. Садовая, 61

E-mail: dmitry.yolkin@maxifier.com

Рассматривается одна из последних тенденций в области предложения рекламного контента – расширенные возможности использования ключевых слов для целевого показа рекламных объявлений пользователям интернет-порталов. Описываются различные способы использования ключевых слов и сочетаний, преимущества, обеспечиваемые использованием данного подхода, механизм и логика доставки, возможности оптимизации по ключевым словам, примеры из практики и достигнутые результаты. В рамках исследования, запущенного с целью выявления потенциального прироста CTR кампаний в результате применения оптимизации по ключевым словам, было выявлено, что в среднем ожидаемый прирост варьируется от 10 до 20 %. Это в совокупности с другими методами оптимизации может существенно увеличить эффективность хода рекламных кампаний.

Ключевые слова: интернет-реклама, рекламная кампания, ключевые слова, оптимизация по ключевым словам, оптимизация медийной рекламы, пара ключ/значение.

В настоящее время в области оптимизации интернет-рекламы наблюдаются новые тенденции, не проявлявшиеся еще несколько лет назад. Среди таких трендов можно отметить стремление рекламодателей оценивать эффективность рекламных кампаний [1] не по количеству доставленных показов или кликов, а по эффективности действий, которые конечный пользователь совершил с каждой конкретной кампанией, например уровню интерактивного взаимодействия с рекламным баннером или регистрации на сайте рекламодателя, не говоря уже о приобретении товара или услуги.

Кроме того, все чаще и чаще в последнее время рекламодатели стремятся показать рекламу не просто на конкретной странице, в конкретной стране или городе, и именно конкретному пользователю или хотя бы целевой аудитории – т. е. тем, кто, по предварительной оценке, должен быть наиболее восприимчив к конкретной рекламе. Например, если мы говорим о рекламе автомобилей, то данную

Дмитрий Александрович Ёлкин, руководитель департамента оптимизации и управления.

Игорь Александрович Минаков (д.т.н.), старший научный сотрудник лаборатории анализа и моделирования сложных систем.

Семен Игоревич Вольман, технический директор.

рекламу целесообразно показывать пользователям, которые, к примеру, посещали сайт автомобильных новостей больше чем несколько раз за последнюю неделю.

С технической точки зрения для реализации этого стремления используются ключевые слова. Рекламодатели и медийные агентства конструируют сложные семантические конструкции путем формирования булевого выражения, описывающего на основе ключевых слов целевую аудиторию, что дает им большую гибкость при управлении своими рекламными кампаниями. Какие преимущества предоставляет использование ключевых слов?

1. С помощью ключевых слов пользователю может показываться в первую очередь та реклама, которая наиболее точно соответствует запросу, введенному им в поисковую систему (google, yandex и т. д.), и, тем самым, которая наиболее релевантна текущим потребностям пользователя;

2. С помощью ключевых слов и выражений, заданных для рекламных кампаний, можно установить семантическое соответствие со страницами, семантическими дескрипторами или иными способами описания семантики страниц, на которых может быть показана данная кампания. Семантические дескрипторы могут строиться автоматически в фоновом режиме с помощью специальных утилит и приложений. В простейшем же случае веб-мастер при верстке каждой страницы может вставлять в запрос на получение рекламы от рекламного сервера некоторый семантический дескриптор (например перечень ключевых слов, описывающих семантику страницы), на основе которого и возвращаются кампании, наиболее релевантные данной странице;

3. Можно более гибко управлять собственным инвентарем (т. е. перечнем сайтов, страниц, которые доступны пользователю и на которых может быть показана реклама). Так, например, если на некотором сайте все страницы динамические и каждая новая статья имеет свой уникальный адрес в интернете в пределах данного сайта, то гораздо проще не вносить каждую статью в общий регистр доступных страниц, а просто разметить ее необходимыми ключевыми словами, описывающими содержание статьи, и уже привязывать рекламу к странице на основе этих ключевых слов.

Так как использование ключевых слов – мощный инструмент, который относительно недавно стал активно применяться, на текущий момент задача оптимизации по ключевым словам приобретает всю большую и большую актуальность.

Примеры ключевых слов и выражений

Прежде чем приступить к описанию конкретных примеров ключевых слов, следует отметить, что общий термин «ключевые слова» можно разделить на несколько категорий.

1. Непосредственно *ключевые слова*, которые представляют собой отдельные слова или фразы, написанные на понятном человеку языке и наделенные некоторым семантическим смыслом. Как правило, когда рекламная кампания описывается с помощью ключевых слов, то оператор может указать, что именно он хочет – чтобы все или хотя бы одно из указанных ключевых слов или фраз присутствовало /отсутствовало на странице, на которой должна быть показана реклама.

В качестве примеров можно привести следующие выражения:

– *makeup, eye+makeup, eyeshadow, eye+shadow, lip+gloss, mascara* – данный набор ключевых слов объединен через логическое ИЛИ и значит, что

кампания может быть показана на странице, если хотя бы одно из этих слов/фраз встретится на этой странице;

- *bmw, audi, !porsche, !opel* – данный набор ключевых слов означает, что реклама будет показана на страницах, которые описывают такие марки, как BMW и Audi, но не будет показана на страницах, посвященных описанию Porsche и Opel.

Заметим, что синтаксис, в котором представляются конкретные ключевые слова и выражения, зависит от выбранного рекламного сервера, который и занимается физической доставкой рекламы на страницы для отображения ее конкретному пользователю.

2. Пара «ключ – значение» и булевы выражения, составленные из различных ключей и их значений. Основное отличие от ключевых слов заключается в том, что в данном случае у нас один какой-то ключ может принимать множество всевозможных значений. В качестве примеров можно привести следующие выражения:

- $((CAR_MAKE=nissan) AND (CAR_LOC~lo))$ – данное выражение обозначает, что рекламная кампания может быть показана только на страницах, имеющих отношение к автомобилям марки Nissan, которые продаются/сдаются в аренду в городах, содержащих в своем названии символы lo, например London. Стоит отметить, что это реальное ключевое выражение, взятое от кампании, рекламирующей сдачу автомобилей в аренду;

- $((((CAR_MAKE=hyundai))AND((CAR_AGE=1)OR(CAR_AGE=2)OR(CAR_AGE=3)OR(CAR_AGE=4)OR(CAR_AGE=5)OR(CAR_MIN=5000)OR(CAR_MIN=6000)OR(CAR_MIN=7000)OR(CAR_MIN=8000)OR(CAR_MIN=9000)OR(CAR_MIN=10000)OR(CAR_MIN=11000)OR(CAR_MIN=12000)OR(CAR_MIN=13000))))))$. Не все ключевые выражения тривиальны и легко понимаются пользователем, существуют выражения и такого вида. В данном выражении используется уже больше ключей и их значений, и при показе рекламы пользователю принимается во внимание не только марка автомобиля, но и его возраст и минимальная стоимость. Зачастую для того, чтобы пользователю был более понятен смысл данных выражений, они требуют упрощения за счет выполнения над ними правил булевой алгебры.

Физический механизм сопоставления рекламной кампании, описанной с помощью ключевых слов и выражений, со страницей, на которой может быть показана реклама, происходит следующим образом (см. рисунок):

1) когда пользователь заходит на какую-то страницу, то загружается содержимое этой страницы, в том числе рекламному серверу посылаются запросы на получение рекламных баннеров;

2) в момент формирования запроса к рекламному серверу определяется список ключевых слов или пар «ключ – значение», которые описывают семантическое содержимое данной страницы, и уже этот список вставляется в запрос к рекламному серверу. Что именно посылать рекламному серверу – отдельные ключевые слова или пары «ключ – значение», зависит от того, как именно они ищутся на странице:

- если в качестве ключевых слов используется содержимое cookie пользователя, то целесообразнее посылать ключевое выражение, составленное из пар «ключ – значение»;
- если список ключевых слов определяется с помощью частотного анализа вхождения тех или иных слов на странице (например этим занимается

компания Grapeshot), то проще сформировать ключевое выражение, состоящее из отдельных слов;

3) после того как запрос отправлен рекламному серверу, он определяет, какому именно ключевому выражению, которое используется для описания рекламной кампании, больше соответствует этот запрос. И уже на основе этой информации выбранный рекламный баннер отдается на страницу пользователя.

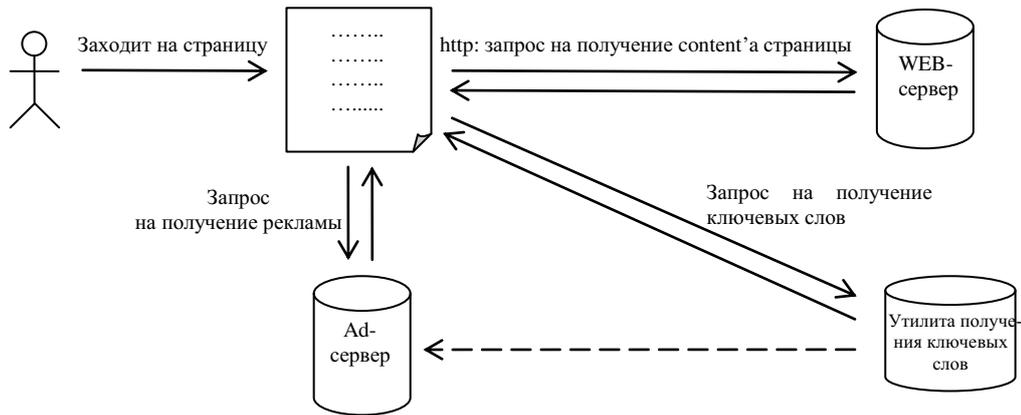


Диаграмма процесса отображения пользователю html-страницы, содержащей рекламу с ключевыми словами

Основные шаги оптимизации по ключевым словам

Оптимизацию по ключевым словам можно разделить на несколько этапов.

1. Преобразование исходного булевого выражения, составленного с помощью ключевых слов и связанного с некоторой рекламной кампанией, к некоторому выбранному базису, в котором может работать система. Приведение исходного выражения к выбранному базису необходимо, потому что разные клиенты могут составлять исходные булевы ключевые выражения разной степени сложности и оптимальности. Преобразование к выбранному базису позволит иметь внутри системы ключевое выражение в одном постоянном формате, что облегчит работу и понимание данного выражения в дальнейшем.

2. Получение статистической информации по каждому ключевому слову, по значению каждого отдельного интересующего ключа либо их совокупности.

3. Вывод на основе имеющейся статистической информации о том, как именно будет происходить оптимизация и какие подходы будут использоваться [2]:

- отдельно по каждому ключу и его значениям (например рассматриваем только все значения ключа key_1 независимо от статистики, набранной другими ключами и значениями);
- не выделяя отдельный ключ либо семантическую категорию и оптимизируя по всем ключам и значениям сразу (например, за раз можно исключить от кампании наименее эффективные значения ключей $key_1, key_2, \dots, key_n$);
- рассматривая и удаляя не отдельные значения отдельных ключей, а наименее эффективную комбинацию тех или иных значений. Например, если значения $key_1=value_1$ и $key_2=value_2$ часто используются сов-

местно, т. е. по ним обоим есть одинаковая статистика, и при этом эта комбинация значений наихудшая среди прочих аналогичных комбинаций, то предлагать исключать за раз не отдельную пару $key_1=value_1$ или $key_2=value_2$, а всю комбинацию $key_1=value_1$ и $key_2=value_2$ целиком.

4. Получение на основе статистической информации и выбранного оптимизационного подхода нового булевого выражения, которое по сравнению с первоначальным выражением будет содержать меньше неэффективных элементов, либо не будет содержать их вообще, либо, наоборот, будет содержать новые, более эффективные с точки зрения выбранного оптимизационного подхода ключевые слова и их значения.

Выбор и построение ключевого базиса

В зависимости от задач и потребностей могут использоваться различные базисы и алгоритмы приведения к ним. В рамках тех задач, которые нам пришлось решать, использовался следующий подход для построения базиса.

1. По исходному ключевому выражению первоначально определяется совокупность базовых форм, где базовая форма представляет собой ключ, значение и операцию между ними.

2. Формируется нормированное выражение, которое представляет собой дизъюнктивную нормальную форму исходного выражения (дизъюнкцию конъюнкций базовых выражений).

3. Строится «осевая» форма выражения – представление нормированного выражения в виде конъюнкций, причем аргумент любой конъюнкции зависит только от одного ключа. Именно с этой осевой формой, если она могла быть построена, осуществлялась дальнейшая работа с точки зрения оптимизации. Алгоритм построения «осевой» формы на основе нормальной формы следующий.

Берем все ключи и для каждого выполняем следующие шаги:

- для каждой конъюнкции базовых выражений исключаем выбранный ключ, получаем набор выражений без заданного ключа;
- если все полученные выражения эквивалентны, то выносим все исключенные базовые выражения в отдельный операнд нового конъюнктивного выражения и выполняем предыдущий шаг для другого ключа, иначе считаем выражение неприводимым к осевой форме.

Для того чтобы описанный алгоритм стал более понятен и очевиден, приведем пример. Допустим, у нас есть исходное ключевое выражение вида

*car=bmw and type=sedan, car=bmw and type=hatchback, car=bmw and type=cuf,
car=bmw and type=offroad, car=bmw and type=cabriolet, car=bmw and type=moto,
car=bmw and type=premium.*

В данном случае исходное выражение представляет собой нормированное выражение (дизъюнкцию конъюнкций). По нормированному выражению строим «осевую» форму, используя алгоритм, описанный выше. Для данного выражения она будет равна

car=bmw and type in (sedan, hatchback, cuf, offroad, cabriolet, moto, premium).

Именно с этой «осевой» формой будет осуществляться дальнейшая работа. Приведение к «осевой» форме осуществляется по нескольким причинам:

- она проще для понимания конечным пользователем;
- если мы говорим об оптимизации, то основные приемы, используемые в оптимизации для улучшения CTR или доставки кампании, – это удаление неэффективных или добавление более эффективных ключей и их

значений. Но прежде чем что-то удалить из выражения, необходимо убедиться, что кампания идет на данном элементе и что удалить желаемый элемент из исходного булевого выражения в принципе возможно. Аналогичная проверка характерна и для добавления элементов. А сделать проверки, описанные выше, на основе «осевой» формы гораздо проще, чем на основе исходного выражения.

Методы и результаты оптимизации по ключевым словам

В качестве методов оптимизации по ключевым словам могут использоваться статистические методы оптимизации, применяемые для оптимизации и других параметров рекламной кампании (например страниц, часов, дней недели, браузеров, операционных систем, GEO-информации и т. д.). Более детально данные методы описаны в статьях [3, 4]. Если говорить кратко, то для оптимизации по CTR используются статистические методы, работающие на основе исторической статистики за последние 2-3 недели и предлагающие удалить наименее эффективные ключи и их значения, т. е. элементы, на которые приходится достаточное количество показов, но реакция пользователей в виде кликов невелика. Стратегия поиска неэффективных элементов определяется конкретным выбранным оптимизационным алгоритмом.

Если мы говорим об улучшении доставки кампании, то в первую очередь наша задача – расширить таргетинг по ключам, т. е. сделать так, чтобы наша рекламная кампания показывалась на максимально возможном количестве ключей и их значений. Чтобы сделать это, в первую очередь надо добавить те значения ключей, на которых кампания еще не идет и которые семантически совместимы с данной рекламой. Во-вторых, если в ключевом выражении есть ограничения, чтобы кампания не шла на каких-то ключах, надо постараться убрать эти ограничения, если это не противоречит контракту с рекламодателем.

Основное отличие ключевых слов от других параметров рекламной кампании заключается в том, что ключевые слова могут быть наделены семантическим смыслом. Т. е. даже не имея никакой статистической информации по ключевым словам и значениям, потенциально, используя методы контекстной оптимизации [2], можно показывать пользователю релевантную рекламу и увеличивать показатели эффективности этой рекламы, такие как доставка и CTR. Авторами статьи ранее были разработаны методы контекстной оптимизации по ключевым словам, суть которых заключалась в том, что для каждой страницы, на которой уже идет или потенциально может идти рекламная кампания, мы должны получить дескриптор, описывающий семантику данной страницы:

- данный дескриптор в простейшем случае может представлять собой набор ключевых слов, наиболее часто встречающихся на данной странице (за исключением всевозможных стоп-слов и слов, не несущих семантической нагрузки);
- в более сложном варианте для каждой страницы строится семантический дескриптор, который представляет собой некоторую онтологическую сеть, с помощью методов морфологического, синтаксического и семантического анализа [5]. По сравнению с методами частотного анализа данный подход позволяет не просто определить наиболее часто встречающиеся слова и фразы, но и выстроить семантические отношения между ними. Это позволяет осуществлять сравнение таких семантических дескрипторов между собой гораздо эффективнее.

Построенный семантический дескриптор может быть использован несколькими способами.

1. Если семантический дескриптор представляет собой набор наиболее часто встречающихся ключевых слов на странице, на которой может идти рекламная кампания, то необходимо добавить эти ключевые слова в параметры таргетинга данной рекламной кампании. Т. е. кроме того, что кампания будет непосредственно запланирована на конкретные страницы, она будет запланирована и на ключевые слова. Это позволит показывать данную рекламу более эффективно. Когда пользователь в некотором поисковике введет ключевые слова, на основе которых ему будет показана одна из страниц, где идет наша реклама, то именно наша рекламная кампания будет показана в первую очередь по сравнению с другими рекламными кампаниями, также запланированными на данную страницу, но не имеющими ключевых слов в параметрах таргетинга. Это происходит потому, что у большинства рекламных серверов рекламные кампании с таргетингом по ключевым словам имеют более высокий приоритет доставки по сравнению с кампаниями, не имеющими ключевого таргетинга. И в результате данного подхода мы сможем улучшить доставку и эффективность нашей кампании.

2. Когда семантический дескриптор представляет собой онтологическую сеть, то:

- необходимо построить аналогичные семантические дескрипторы для всех страниц, на которых в настоящий момент не идет выбранная рекламная кампания;
- необходимо для данной рекламной кампании на основе описания, предоставляемого для данной рекламы в виде понятного человеку текста, построить аналогичный семантический дескриптор в виде онтологической сети;
- после того как дескрипторы для кампании и для всех страниц построены, необходимо произвести сравнение – какие именно семантические дескрипторы страниц наиболее близки дескриптору рекламной кампании, на основе результатов сравнения выбрать наиболее подходящие страницы и запланировать на них рекламную кампанию.

Сравнение эффективности различных оптимизационных подходов

Метод оптимизации	Прирост по доставке, %	Прирост по CTR, %
Статистическая оптимизация по страницам	5–7	5–10
Статистическая оптимизация по страницам + эвристические методы (Frequency Capping, распределение в течение дня)	7–10	10–13
Оптимизация по страницам + оптимизация по ГЕО-информации, браузеру, операционной системе	10–13	13–18
Оптимизация по страницам + ГЕО (включая браузеры, ОС) + статистическая оптимизация по ключевым словам	13–18	18–23
Оптимизация по страницам + ГЕО (включая браузеры, ОС) + статистическая и контекстная оптимизация по ключевым словам	18–22	23–28

В рамках ООО НПК «Максифаер Девелопмент» был разработан собственный программный продукт, который представляет собой систему поддержки принятия решений, направленную на оптимизацию рекламных кампаний по ключевым словам.

В рамках исследования, запущенного с целью выявления потенциального прироста CTR кампаний в результате применения оптимизации по ключевым словам, было выявлено, что в среднем ожидаемый прирост варьируется от 10 до 20 %. Это в совокупности с другими методами оптимизации может существенно увеличить эффективность хода рекламных кампаний. Детальное сравнение приведено в таблице. Данное исследование проводилось на выборке, состоящей из 30 рекламных кампаний, каждая из которых наблюдалась в течение месяца. Каждая рекламная кампания была разделена на 5 подкампаний, имевших изначально одинаковые параметры таргетинга. Далее к первой группе кампаний применялись только методы статистической оптимизации по страницам, ко второй – методы статистической оптимизации по страницам и эвристические методы. Аналогично к третьей, четвертой и пятой группам применялись соответствующие методы, приведенные в таблице. Анализ показывает, что чем большее количество методов применяется при оптимизации, тем выше потенциальный прирост эффективности.

Заключение

В рамках данной статьи рассмотрены основные шаги, которые выполняются в рамках оптимизации по ключевым словам, дано определение ключевого слова, пары ключ/значения и булевых выражений, которые применяются для описания таргетинга рекламной кампании по ключевым словам, рассмотрен физический процесс доставки рекламной кампании на страницу, которую посетил пользователь, с учетом специфики данной страницы с позиции ключевых слов и различные подходы к оптимизации по ключевым словам.

Анализ подходов к оптимизации ключевых слов показал, что их использование позволяет дополнительно повысить CTR рекламных кампаний в среднем на 10-20 %.

Использование таргетинга по ключевым словам дает возможность гибко настраивать рекламные кампании, чтобы добиться наибольшей эффективности и в то же время упростить процесс управления инвентарем владельца сайтов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Plummer Joe, Rappaport Steve, Hall Taddy, Barocci Robert.* The Online advertising playbook // Wiley 2007, John Wiley & Sons, Inc, 2007, p. 28-32.
2. *Минаков И.А., Елкин Д.А., Вольман С.И.* Обзор подходов оптимизации интернет-рекламы // Проблемы управления и моделирования в сложных системах: Тр. XI Междунар. конф., Самара, 22 июня – 24 июня 2009. – Самара: СНЦ РАН, 2009. – С. 634-643.
3. *Елкин Д.А., Минаков И.А., Вольман С.И.* Использование методов статистической оптимизации для улучшения эффективности показа интернет-рекламы // Проблемы управления и моделирования в сложных системах: Тр. XII Междунар. конф. – Самара, 21 – 23 июня 2010. – Самара: СНЦ РАН, 2010. – С. 577-585.
4. *Елкин Д.А., Минаков И.А., Вольман С.И.* Автоматическая оптимизация интернет-рекламы // Вестник Самарского государственного технического университета. Сер. Технические науки. – 2011. – № 3(31). – С. 228-232.
5. *Минаков И.А.* Алгоритм кластеризации семантических дескрипторов документов // Вестник Самарского государственного технического университета. Сер. Технические науки. – 2009. № 1 (23). С. 34-46.

Статья поступила в редакцию 5 июля 2014 г.

METHODS OF ONLINE ADVERTISING OPTIMIZATION BASED ON KEYWORDS AND KEYPAIRS ANALYSIS

D.A. Elkin¹, I.A. Minakov², S.I. Volman¹

¹ OOO Maxifier Development
349, Novo-Sadovaya st., Samara, 443125, Russia

² Institution of the Russian Academy of Sciences Institute for the Control of Complex Systems of RAS
61, Sadovaya st., Samara, 443020, Russia

The paper describes last trends in Internet advertisement area - keyword/key-value terms and optimization methods used based on keyword statistics. Article covers technical process of ad-deliveries, different optimization methods and their comparative analysis on real life scenarios.

Keywords: *Online advertising, ad campaign, ad-servers, keywords, keywords optimization, keywords/keypairs.*

*Dmitry A. Elkin, Director of optimization, Maxifier Development.
Igor A. Minakov, Senior Scientist, ICCS RAS.
Simon I. Volman, Director of Technology, Maxifier Development.*