

УДК 004.67

ОПТИМИЗАЦИЯ МАРКЕТИНГОВЫХ ПОКАЗАТЕЛЕЙ РЕКЛАМНЫХ КАМПАНИЙ RTB С ПОМОЩЬЮ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

ВАХОВИЧ И.О., аспирант

Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь

Ключевые слова: Интернет, рекламные продажи, Real-time bidding (RTB), Big Data.

Реферат. В последние годы происходит заметный рост рекламных технологий на основе RTB. Заметную роль в таком росте играет то, что различные методы машинного обучения показали свою эффективность в оптимизации кампаний, начиная от предсказания кликов и конверсий, распределения рекламного бюджета, фильтрации мошеннических показов. Одним из интересных примеров такой оптимизации является построение модели предсказания CTR на основе логистической регрессии.

RTB (Real Time Bidding) — технология закупки медийной рекламы посредством программируемых онлайн-аукционов. RTB фокусируется непосредственно на показах целевым посетителям, а не планированию резервов рекламных площадей на определенных сайтах. Каждый показ выкупается за доли секунды — во время загрузки страницы — система RTB мгновенно проводит аукцион. В результате лучшее предложение от рекламодателей появляется на глазах пользователя, которому оно наиболее интересно.

Участники системы:

Supply Side Platform (SSP)— рекламная сеть, предоставляющая в аукцион места для показов рекламы (продавец);

Demand Side Platform (DSP)— рекламная сеть, предоставляющая в аукцион рекламу— видео, баннеры и пр. (покупатель);

Data Management Platform (DMP)— поставщик данных о пользователях;

Agency Trading Desk— платформа, которая помогает рекламодателям более эффективно покупать аудиторию в большом объеме в цифровых средствах массовой информации;

Publisher - владелец сайта;

RTB аукцион работает следующим образом(рисунок 1):

Пользователь ввел в строку браузера название сайта, подключенного к рекламной RTB сети.

Сайт перед загрузкой страницы отправляет сети, к которой он подключен, запрос на показ баннера, а также данные о пользователе (не ФИО и адрес проживания, а доступные для таргетинга данные).

Рекламная сеть передает данные о пользователе в SSP.

SSP классифицирует запрос по максимальному числу параметров: площадка на которую зашел пользователь, данные о пользователе, время захода, данные по ретаргетингу.

SSP выставляет торг — информацию о сайте и пользователе системам рекламодателей. При желании SSP закупает дополнительные данные о пользователе у DMP.

Trading Desk, используя DSP, передает информацию о ставке SSP.

SSP обрабатывает информацию от всех рекламодателей, сделавших ставку. Выбирает самую высокую.

Снижает ее до второй цены и осуществляет показ по цене, превышающей вторую по величине ставку на 1 цент.

Получает объявление от победителя и передает его браузеру пользователя в ответ на запрос рекламного объявления.

Загружается страничка с полученным объявлением.

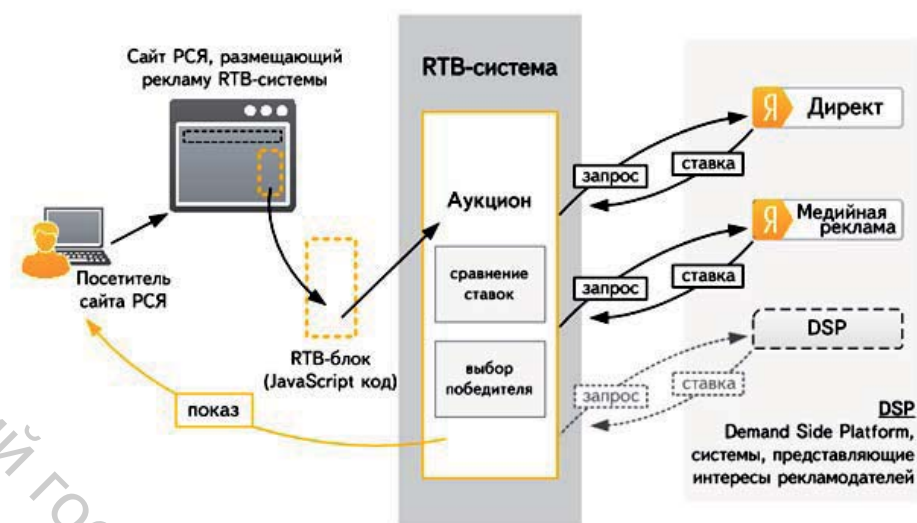


Рисунок 1 – Схема работы RTB аукциона

Следует обратить внимание на то, что все эти операции происходят между получением запроса и ответом на него, то есть в процессе загрузки страницы в браузер пользователя.

Как правило, рекламодатель работает напрямую через DSP или Trading Desk. Цель рекламодателя использовать весь доступный на кампанию бюджет и достигнуть нужных показателей:

CPM (cost-per-mile) – определенное количество показов

CTR (click through rate)/ CPC (cost-per-click) – отношение числа кликов и показов не менее ctr_goal и цена клика не более cpc_goal соответственно

CPI/CPA (cost-per-install/action) – цена установки мобильного приложения или любого вида конверсии

CPV (cost-per-video view) – цена просмотра видео

DSP от лица рекламодателя занимается оптимизации таких показателей. За каждый выигранный показ DSP платит деньги SSP, используя бюджет кампании, выделенный рекламодателем. Соответственно чтобы достигнуть целевых показателей, DSP нужно покупать как можно больше потенциально выгодных показов.

До того как купить показ DSP должна оценить вероятность того, что пользователь кликнет по баннеру. Исходя из того, что есть исторические данные о показах и кликах, DSP может использовать их как тренировочные данные для построения модели (supervised learning) на основе методов логистической регрессии, деревьев, SVM и прочих.

Мы будем рассматривать данную задачу как бинарную классификацию, зная для каждого показа результат (был ли клик по баннеру или нет), и построим бинарный классификатор. Каждый показ содержит информации разного рода: про пользователя, контекст (сайт или мобильное приложение, где будет размещен баннер, их категория, среднее число активных пользователей) и параметры самого баннера. Информация про пользователей обычно содержит тип и модель устройства (смартфон, десктоп), операционная система, гео-положение, история откликов на предыдущие показы. Параметры баннера – это размер и формат, место показа, цена ставки (исторический показ), рекламодатель, кампания и ее тип (CPC/CPV/CPI). Также доступна информация о времени показа (час/день недели/месяц).

Существует множество вариантов оптимизации ставок в RTB-аукционе (см. [1] о других подходах). Далее мы рассмотрим классификатор, построенный на основе логистической регрессии, которая используется для предсказания вероятностей событий. Рассмотрим показ Y , который задается набором признаков $X(1)$

$$X = \{x_1, x_2, \dots, x_n\}, \quad (1)$$

Определим логистическую функцию

$$f(z) = P(Y = 1|X = x). \quad (2)$$

Заметим, что $f(z) = 1$, когда за показом следует клик, и 0 в противном случае. Стандартный вид:

$$f(z) = \frac{1}{1 + \exp(-z)} \quad (3)$$

где

$$z = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \dots + \beta_n * x_n, \quad (4)$$

где β_i коэффициенты регрессии для каждого признака x_i . Каждый коэффициент отражает размер вклада отдельного признака на финальную вероятность. Положительный коэффициент означает, что этот признак увеличивает значение вероятности в сторону 1, т.е. показ, в котором присутствует этот признак, имеет большие шансы быть “кликновым”.

Преимущества логистической регрессии при предсказании кликов:

достаточно простая, но эффективная модель. Построение модели на больших данных от 100 Gb занимает до 3-5 минут на одной машине(AWS t2.medium),

методы регуляризации позволяют выбрать наиболее важные признаки,

быстрое вычисление вероятности по формулам(ссылки для z и $f(z)$) в реальном времени для каждого запроса на рекламу – важно, т.к. все должно происходить в момент загрузки веб-страницы или приложения,

четкая интерпретация коэффициентов.

Полученные вероятности далее используются для вычисления ставок в аукционе. Основной принцип заключается в том, что чем больше вероятность клика, тем большую ставку должна поставить DSP, чтобы выиграть этот показ. Разработано много вариантов стратегий ценообразования [2].

Литература:

1. Chen, Y., Berkhin, P., Anderson, B., & Devanur, N. R. (2011, August). Real-time bidding algorithms for performance-based display ad allocation. In Proceedings of the 17th ACM SIGKDD
2. Zhang, W., Yuan, S., & Wang, J. (2014, August). Optimal real-time bidding for display advertising. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1077-1086). ACM.

УДК658.152:519.7

ПРИМЕНЕНИЕ КИБЕРНЕТИЧЕСКОГО ПОДХОДА В ИНФОРМАЦИОННОЙ ЛОГИСТИКЕ

ВЕРЕТЕННИКОВА Е.С., аспирант, КРАСНОВА И.И., доцент

Белорусский национальный технический университет,

г. Минск, Беларусь

Ключевые слова: информационная логистика, кибернетический подход, кибернетическая модель.

Реферат. Исследована возможность использования и обоснована целесообразность применения кибернетического подхода в информационной логистике. Построены кибернетические модели логистической системы и интегрированной цепочки поставок.

Кибернетика – наука об общих законах управления в природе, обществе, живых организмах и машинах, изучающая информационные процессы, связанные с управлением динамических систем [1].

Логистическая система может быть рассмотрена как управляемая система, то есть она может быть отнесена к категории кибернетических систем. Поэтому видится возможным изучать, исследовать и анализировать логистическую систему на основе кибернетического подхода.