

УДК 659.1:004.738.5

ЦЕЛЕВАЯ ИНТЕРНЕТ-РЕКЛАМА



В.В. Юдин

Магистрат кафедры информатики БГУИР



Н.А. Волорова

Заведующая кафедрой информатики БГУИР, кандидат технических наук, доцент

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь
E-mail: vasily@vypage.com, volorova@bsuir.by*

Аннотация. Основная цель работы в предоставлении гибкой, легкой в использовании платформы целевой рекламы с высокой степенью таргетинга. Платформа позволяет бизнесу в рамках рекламной кампании нацеливать рекламу на малую группу пользователей, определенную точными критериями. В результате бизнес может значительно снизить затраты на проведение рекламной кампании, в то же время получая более высокую конверсию, а пользователи получают более релевантную рекламу.

Ключевые слова: Целевая реклама, Amazon Web Services, Обработка больших объемов информации, Apache Hadoop, Apache Spark, Распределенные вычисления, AWS EMR

Реклама играет очень важную роль как для бизнеса, так и для потребителя. Реклама формирует имидж брэнда и товара, стимулирует спрос. До нашего времени, когда интернет-реклама, как и сама интернет-сеть, не была настолько распространена, реклама была, по большому счету, неперсонифицирована: в основном, это были бумажные носители, телевидение, объявления. Данные виды рекламы позволяли только условно определить целевую аудиторию, что имело отрицательный эффект как для потребителя — множество навязчивой, но, при этом, не интересующей и, как следствие, раздражающей рекламы, — так и для бизнеса — соотношение количества действительно заинтересованных потребителей к количеству показов рекламы было сравнительно мало.

Интернет-реклама — самый прогрессивный и эффективный вид рекламы в настоящее время. Однако, с появлением интернет-рекламы ситуация изменилась не существенно: реклама стала ещё более агрессивной и навязчивой, но не релевантной.

Со временем, интернет-компании, имеющие огромную аудиторию, такие как Google, Facebook, Apple, стали предоставлять наиболее релевантную рекламу на основе данных о пользователях для построения сегментов: использование истории поисковых запросов для определения сферы интересов и предпочтений, использование данных профиля, таких как пол, возраст, часто посещаемые страницы и сообщества. Подавая собственную рекламу через такие платформы, бизнес может достаточно тонко установить критерии для определения сегмента, однако, запрос, определённо точно определяющий целевую аудиторию, например “только пользователи, приобретавшие тёмно-синие джинсовые куртки брэнда X”, до сих пор невозможен.

Платформа целевой интернет-рекламы (далее — Платформа) призвана решить описанную выше проблему, тем самым, не только предоставив наиболее релевантную рекламу пользователям интернет-сети, а также минимизировав расходы на рекламу для бизнеса, показывая

рекламу значительно меньшему сегменту, также увеличивая конверсию, так как люди, увидевшие рекламу, будут с высокой долей вероятности в ней заинтересованы.

Основа, на которой базируется работа Платформы, — сбор данных о пользователях, их поведении и предпочтениях. Для этого данные о пользователе передаются Платформе с привязкой к идентификатору пользователя. Данные могут иметь различную степень подлинности: от заполненной самим пользователем формы с указанием его пола, возраста и других данных до предиктивных данных, вычисленных с помощью нейронной сети, предварительно натренированной на поведении пользователей различных сегментов на странице, которые, конечно, имеют меньшую степень подлинности.

Помимо данных о статических значениях, Платформа также собирает данные о поведении пользователей на ресурсах — события, такие как покупка, помещение товара в корзину, удаление из корзины, а также информация о событиях, такие как купленный предмет, его цвет, брэнд и т.д. События представляют собой важную информацию о поведении пользователя и его предпочтениях.

Кроме статических значений и событий, Платформа также поддерживает внешние сегменты. Согласно политике конфиденциальности некоторых источников данных, предоставляющих информацию о пользователях Платформы, они не могут передать данные в открытом виде, а предоставляют уже сегментированные данные по определенным критериям. Таким образом, внешний сегмент — это уникальный идентификатор сегмента и список идентификаторов пользователей, относящихся к данному сегменту.

Сбор данных о пользователях реализован двумя способами: сбор информации в режиме реального времени путем размещения JavaScript-кода на сайтах-партнёрах, выполняющего запросы к RESTful API, являющемся компонентом Платформы, а также сбор информации в режиме offline — импорт данных, предоставляемых источником данных, на Платформу раз в сутки.

После сбора информации о пользователях, партнёр, заинтересованный в проведении целевой рекламной кампании, с помощью веб-интерфейса, предоставленного Платформой, может создать собственный сегмент, используя предоставленный набор фильтров, а также простые логические операторы (AND, OR, NOT), которые могут иметь неограниченный уровень вложенности. Фильтры генерируются на основе данных, полученных от источников, а также в качестве фильтров могут выступать внешние сегменты с уже предопределённым списком идентификаторов. В итоге, разрабатывая собственную рекламную кампанию, можно создать, например, фильтр “парни от 22 до 25 лет, которые покупали красные куртки или синие штаны в последние два месяца, но не дороже 100 рублей, а также относящиеся к внешнему сегменту X.”

Сайты-партнёры также размещают так называемый “пиксель” (JavaScript-код) Платформы, который, выполняя запрос к API, либо выдаёт Реклама играет очень важную роль как для бизнеса, так и для потребителя. Реклама формирует имидж брэнда и товара, стимулирует спрос. До нашего времени, когда интернет-реклама, как и сама интернет-сеть, не была настолько распространена, реклама была, по большому счету, неперсонифицирована: в основном, это были бумажные носители, телевидение, объявления. Данные виды рекламы позволяли только условно определить целевую аудиторию, что имело отрицательный эффект как для потребителя — множество навязчивой, но, при этом, не интересующей и, как следствие, раздражающей рекламы, — так и для бизнеса — соотношение количества действительно заинтересованных потребителей к количеству показов рекламы было сравнительно мало.

Интернет-реклама — самый прогрессивный и эффективный вид рекламы в настоящее время. Однако, с появлением интернет-рекламы ситуация изменилась не существенно: реклама стала ещё более агрессивной и навязчивой, но не релевантной.

Со временем, интернет-компании, имеющие огромную аудиторию, такие как Google, Fa-

sebook, Apple, стали предоставлять наиболее релевантную рекламу на основе данных о пользователях для построения сегментов: использование истории поисковых запросов для определения сферы интересов и предпочтений, использование данных профиля, таких как пол, возраст, часто посещаемые страницы и сообщества. Подавая собственную рекламу через такие платформы, бизнес может достаточно тонко установить критерии для определения сегмента, однако, запрос, определённо точно определяющий целевую аудиторию, например “только пользователи, приобретавшие тёмно-синие джинсовые куртки бренда X”, до сих пор невозможен.

Целевая реклама призвана решить описанную выше проблему, тем самым, не только предоставив наиболее релевантную рекламу пользователям интернет-сети, а также минимизировав расходы на рекламу для бизнеса, показывая рекламу значительно меньшему сегменту, также увеличивая конверсию, так как пользователи, которым будет продемонстрирована реклама, будут с высокой долей вероятности в ней заинтересованы.

Основа, на которой базируется работа целевой рекламы, — сбор данных о пользователях, их поведении и предпочтениях. Данные могут иметь различную степень подлинности: от заполненной самим пользователем формы с указанием его пола, возраста и других данных до предиктивных данных, вычисленных с помощью нейронных сетей, предварительно натренированных на поведении пользователей различных сегментов на странице, которые, конечно, имеют меньшую степень подлинности.

Помимо данных о статических значениях, также возможен сбор данных о поведении пользователей на ресурсах — событий, таких как покупка, помещение товара в корзину, удаление из корзины, а также информации о событиях, такой как купленный предмет, его цвет, бренд и т.д. События представляют собой важную информацию о поведении пользователя и его предпочтениях.

Сбор данных о пользователях может быть реализован различными способами: сбор информации в режиме реального времени путем размещения JavaScript-кода на сайтах-партнёрах, выполняющего запросы к API, а также периодический импорт данных, предоставляемых различными источниками.

После сбора информации о пользователях, бизнес, заинтересованный в проведении целевой рекламной кампании, с помощью предоставленного, в зависимости от реализации платформы, интерфейса, может создать собственный сегмент, используя предоставленный набор фильтров и простые логические операторы. Набор фильтров определен данными, полученными от источников данных. В итоге, разрабатывая собственную рекламную кампанию, бизнес имеет возможность достаточно тонко настроить сегмент, например “парни от 22 до 25 лет, которые покупали красные куртки или синие штаны в последние два месяца, но не дороже 100 рублей.”

Таким образом, исходными данными являются анонимизированные данные о пользователях с привязкой к уникальным идентификаторам, а также набор сегментов и их правил. После получения данных от различных источников, целесообразно привести их к единому формату для оптимизации дальнейшей обработки данных. Затем необходимо объединить все данные, имеющие отношение к одному и тому же пользователю. Полученные данные можно отфильтровать по заданным критериям, в итоге получив список пользователей, которые и будут являться целевой аудиторией рекламной кампании.

В качестве инфраструктурного решения для реализации платформы целевой интернет-рекламы целесообразно использовать AWS (Amazon Web Services), так как AWS предоставляет все необходимые сервисы “из коробки”, такие как файловые хранилища, EMR-кластеры с предустановленным Hadoop, Athena, позволяющая анализировать данные в S3 стандартными средствами SQL, и др.

EMR предоставляет возможность динамически масштабировать сервера, используемые

для обработки данных, что позволяет как ускорить вычисления, расширяя кластер, так и оптимизировать стоимость, избавляясь от излишних ресурсов, которые не нужны в данный момент времени. EMR также поддерживает HBase — распределенное хранилище данных, — удовлетворяя нужды проекта целевой рекламы. Для распределенной обработки данных целесообразно использовать Apache Spark, обладающий высокой производительностью и позволяющий реализовать приложение на четырех языках: Java, Scala, Python и R.

Таким образом, используя целевую рекламу при проведении рекламных кампаний, бизнес значительно снижает затраты на их проведение, в то же время увеличивая конверсию, а пользователи получают максимально релевантную и, как следствие, не раздражающую рекламу.

Список литературы

- [1]. Tom White “Hadoop: The Definitive Guide, 4th Edition”, O’Reilly Media
- [2]. Matei Zaharia, Bill Chambers “Spark: The Definitive Guide”, O’Reilly Media
- [3]. Jamie Allen “Effective Akka”, O’Reilly Media

TARGETED ADVERTISING

V. YUDIN

Master of of Informatics Department of BSUIR

N.A. VOLARAVA, PhD

Head of the Chair of Informatics of the BSUIR

Belarusian State University of Informatics and Radioelectronics, Republic of Belarus

E-mail: vasily@vypage.com, volarova@bsuir.by

Abstract. The main aim of the work is to provide deep understanding of how the targeted advertising works, and to provide an overview of the approaches and technologies that can be used to create a platform with a high level of targeting. Target advertising allows businesses to target small groups of users by very specific criteria. As a result, businesses reduce costs of advertising campaigns, having better conversion rate at the same time, and users get more relevant and less annoying advertisement.

Key words: Targeted advertising, Amazon Web Services, Big data processing, Apache Hadoop, Apache Spark, Distributed calculations, AWS EMR