

Real-time Visualization and Targeting of Online Visitors

Deepak Pai
Adobe Research
Adobe
dpai@adobe.com

Sandeep Zechariah George
Adobe Research
Adobe
sgeorge@adobe.com

ABSTRACT

Identifying and targeting visitors on an e-commerce website with personalized content in real-time is extremely important to marketers. Although such targeting exists today, it is based on demographic attributes of the visitors. We show that dynamic visitor attributes extracted from their click-stream provide much better predictive capabilities of visitor intent. In this demonstration, we showcase an interactive real-time user interface for marketers to visualize and target visitor segments. Our dashboard not only provides the marketers understanding of their visitor click patterns, but also lets them target individual or group of visitors with offers and promotions.

1. INTRODUCTION

Organizations are embracing online platforms to increase their reach and reduce operational costs of brick and mortar stores. Of late even software companies are moving towards online purchase and subscription models. However the online conversion rates are typically in the range of 1-3 % [4], which is significantly less compared to conversions in offline world. Hence they are interested in analysing online visits to understand low conversion rates and influence customers to make purchases. Techniques like web-page personalization, campaigns, offers etc. are routinely adopted to persuade visitors. These techniques increase conversions at the cost of significant marketing budget. As marketing to non-purchasers amounts to wasted dollars, marketers would like to identify visitors who would best respond to a campaign. Such techniques today are based on demographic attributes of the visitors like city, state, age, income etc. Our approach to identify this target group is based on visitor click-streams.

Click-stream refers to the sequence of URLs that are visited by the customer in addition to meta-data such as geo-location, purchase information, page categories, referrer etc. This information is used to gain insight into the user intent, and subsequently predict potential leads, churn or any such desired metric. Clustering of visitor sessions by appropriate features generates market segments for targeted campaigns.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).

SIGIR'14, July 6–11, 2014, Gold Coast, Queensland, Australia.

ACM 978-1-4503-2257-7/14/07.

<http://dx.doi.org/10.1145/2600428.2611180>.

Look-a-like modelling, anomaly detection, time series prediction are commonly used to identify out of sequence observations at individual or aggregate levels. Applications of such modelling techniques include fraud detection, web page personalization, market segmentation, offline and online targeting etc. The techniques are used across retailers, publishers, finance and other domains. Our experiments show that methods based on click-stream analysis provide significantly better predictive capabilities than other techniques based on demographic attributes.

In this demonstration we showcase a real-time dashboard for marketers, which helps them visualize browsing patterns of visitors on their website. Visitors could be segmented based on various attributes like their geo-location, device type etc. Moreover we also predict the likelihood of visitor to make a purchase based on supervised learning algorithms. Marketers have the ability to select a visitor or group of similar visitors to form a market segment. Offers or promotional campaigns can be pushed to these selected users, increasing their chances of making a purchase.

2. RELATED WORK

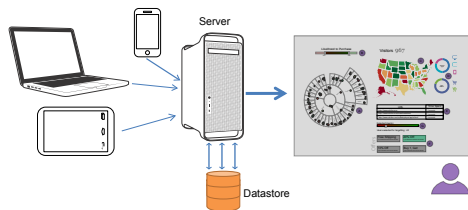
Primarily analytics involved exploration or generating insights from historic data and using it for future decisions. Of late industry has started moving towards the direction of providing real-time analytics. Companies like Chartbeat¹ provide real-time dashboard to marketers showing various attributes of visitors on the website. However they lack predictive capabilities to distinguish purchasers from non-purchasers. Further, certain companies like Netmining² provide purchase propensity score of visitors. Although such dashboards are useful, there is no action-ability that a marketer could take, to increase conversions on his website. Our technology overcomes these shortcomings by not only providing information about visitor attributes in real-time, but also the ability to segment visitors by their on-site behaviour and appropriately target them. Such personalization leads to higher conversion and more revenue to marketers.

3. ARCHITECTURE

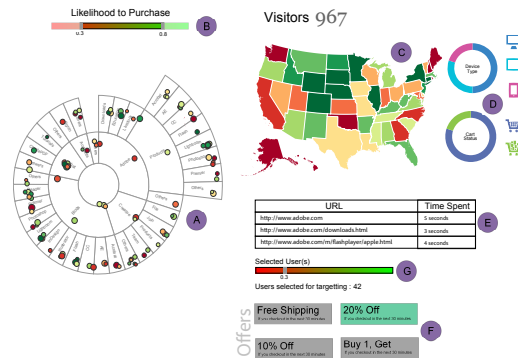
This section describes overall architecture of the system, input data, dashboard, propensity scoring mechanism and targeting. Figure 1a shows the overall architecture of our system. When visitor clicks on a web-page, click information is fed into our processing engine. The processing modules store the data in persistent storage as well as compute propensity score for the visitor.

¹<https://chartbeat.com/>

²<http://www.netmining.com/>



(a) Architecture of the system showing various components and the data flow



(b) Dashboard displaying visitors and their characteristics. Controls for marketer to select, filter and target visitors

Figure 1: Real-time online visitor monitoring and targeting system

Data: We use data of visitors on the Adobe website³. Every visitor click captured includes information such as visitor identifier (obtained from cookies), IP, page URL, referrer, meta-data associated with the page such as category, product price, whether the click resulted in purchase and so on. The above information is collected from standard methods such as javascript code, that request a single pixel images from the server, resulting in a hit. Additional information such as visitor location is derived from the IP and appended to the click-stream. The data is streamed to the processing engine using streaming technologies such as Storm⁴.

Back-end system and Data Store: Our processing module is implemented in python. We use Apache HBase⁵ as our data store. HBase is a columnar, nosql database which has fast response time for storage and retrieval.

4. VISUALIZATION

Figure 1b shows our real-time dashboard, accessible to the marketer. Item 'a' shows zoom-able sundial representing site-sections of the website. For instance we see regions representing adobe.com, blogs and help, creative cloud among others. Each visitor is represented as a bubble on the dial, with size and color representing the time spent by the visitor on the website and their likelihood of making a purchase respectively. One could zoom into specific site-sections by clicking on the respective slice and understand visitor browsing patterns. For instance visitors could be spending more time on certain site-sections than others or certain site-sections could have higher conversions than others. Such information could help the website owner optimize the page content, link structure to improve conversion. Marketer could use the slider as shown in item 'b' to filter and view visitors with propensity scores in a given range. Purchase patterns from different geo-locations are visualized on a map shown in item 'c'. States are coloured based on their conversion rates on a scale ranging from red to green, with higher conversions shown in green and lower conversions coloured red. This information could then be used to provide offers to visitors from areas with lower conversion. We also show donut visuals in item 'd' for two sample visitor attributes like device type or cart status. Marketer could choose which attributes he would like to view among a large number of available attributes. Clicking on a given visitor reveals additional information about the visitor such as the URLs visited and time spent, shown in item 'e'. Slider provided in item 'f' provides the marketer ability to include as

many or as few visitors similar to the selected visitor. Finally the marketer can target the selected visitor segment in real-time with appropriate offers (item 'g'). The promotional offers are pushed on the website for the selected users. **Propensity Scoring:** Here we briefly describe the mechanism used for computing the propensity score of the user. Various techniques have been explored in literature for classifying purchasers and non-purchasers on website [2, 1]. Markov chains and HMMs have been popular approaches for modelling click-stream data. Our approach is based on building features from the click-stream data, computing similarity score between visitor pairs and using kNN to classify visitors. The details of methodology for classifying visitors is described in [3], where we show that our methods outperform prior-art systems. The same methodology is used to identify visitors similar to any selected visitor, thus creating a visitor segment. **Targeting:** Finally we provide the marketer with the ability to push promotions, discounts or other such offers to selected visitor segments. Our system creates rules based on the selected attributes and offers. The rules are used to dynamically deliver personalized content to the visitor. We also provide predicted lift from the offer based on historic data.

5. CONCLUSION

In this demonstration we have created a real-time dashboard for marketer to visualize on-site traffic patterns and target selected user segments with offers. Marketers have the ability to create appropriate visitor segments for targeting based on various attributes. Our work combines visualization, predictive capabilities and targeting in real-time, outperforming existing solutions.

6. REFERENCES

- [1] W. W. Moe. Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. *Journal of Consumer Psychology*. Elsevier, 2003.
- [2] A. L. Montgomery, S. Li, K. Srinivasan, and J. C. Liechty. Modeling online browsing and path analysis using clickstream data. *Marketing Science*. INFORMS, 2004.
- [3] D. Pai, A. Sharang, M. M. Yadagir, and S. Agrawal. Modelling visit similarity using click-stream data: A supervised approach. In *submission, SIGKDD*. ACM, 2014.
- [4] N. Poggi, D. Carrera, R. Gavalda, E. Ayguadé, and J. Torres. A methodology for the evaluation of high response time on e-commerce users and sales.

³<http://www.adobe.com>

⁴<http://storm-project.net/>

⁵<http://hbase.apache.org/>