ABSTRACT

The digital advertising world is under pressure from legislators, regulators, consumers and privacy advocates – among other stakeholders - who are reacting to the surveillance crisis that is leading to what will be the eventual deprecation of third-party cookies.

At the same time, marketers are finding that changes in Apple’s privacy standards requiring consumers to opt into tracking, is driving the cost of acquisition well beyond the economic models that many DTC brands are built on.

To dimensionalize this, 96% of Apple users have opted out, and the result has dramatically changed the model for brands across all industries, seeing up to 60% increases in acquisition costs.

The SORT™ Whitepaper will describe our cookieless, AI based technology and approach, which leverages “Smart Groups” and non-Euclidian distancing, to target users based on a combination of “in-the-moment” behaviors and the mindsets they have at the time of the advertising encounter, without access to any aspects of their personal identities. This service is currently running in production at scale, across over 100 advertisers spanning all verticals including brands such as, Ambetter, Colorado Tourism and Owl Labs.

The accuracy and effectiveness tests performed on the machine learning based model, show outstanding and ever-improving results for ad-based engagement metrics and has been comprehensively reviewed by Neutronian, a pioneer of independent data quality certification, and was awarded with Neutronian’s Cookieless Certification.
PROBLEM DESCRIPTION

There are three key trends which are rising in the digital advertising domain:

**USER PRIVACY** has become an important issue. Ad platforms and browsers alike are taking steps to improve transparency and privacy. Much of the digital marketing discussions today focus on data privacy, specifically managing what consumers consent to share and for what purposes. The fact is, consumers are demanding privacy and as part of that, marketers should align with GDPR/CCPA legislation and privacy-first solutions. In a recent AdWeek session titled "The New Blueprint for Consumer Privacy is taking Shape", marketing executives Christine Cox, IKEA U.S.'VP of Digital Data and Analytics along with Mozilla’s CMO Lindsey Shepard, talk about meeting consumers where they are in a cookieless future to protect their privacy and ensure their trust*.

**3rd PARTY COOKIES** are disappearing. The use of third-party cookies has been an important element of the marketing mix for many organizations. Advertisers in the ad tech industry rely heavily on identifying users and targeting them based on 3rd party cookies. However, leading browsers have made public announcements and technical deployments to reduce the digital advertising industry’s access to third-party cookies for data collection, storage, and sharing. As a result, there has been growing momentum to find alternatives for understanding customer behavior and target accordingly, as many browsers have already stopped allowing cookie-based tracking such as Firefox and Safari.

**INACCURACY OF 3rd PARTY COOKIES.** Ironically, despite its broad adoption, cookie-based targeting has never been 100% accurate. That’s because tracking user’s internet behavior to categorize them into a targetable persona or demo groups is inherently flawed, as it makes assumptions on the user based on their internet behavior (i.e., just because someone visited a make-up site, they must be a female). In other words, its very surveillance methodology which infuriates users, is the Achilles heel of the cookie.

In general, competitive realities limit third-party cookies as to what they can track when it comes to cross-device and cross-browser user behaviors (for example, Google Chrome will never know what browsing behaviors you have when you are on Safari, and vice versa). Additionally, as users are becoming hyper-aware of their privacy, they are also frequently clearing their cache & disabling cookie tracking on certain websites- which limits the amount & frequency of data collection. So, while cookies are currently still available, they might not be the best/ most accurate way to reach an audience.
To overcome these obstacles, we undertook extensive research to develop a scalable, cookieless, AI based strategy that will allow for the delivery of the most relevant and most accurate marketing message to each user, driven by identifying the general “moment and mindset” at which we encounter them (as opposed to identifying the specific person), leading to documented brand performance that is superior to cookie-based targeting.

**TECHNOLOGY SOLUTION AND RESEARCH DESCRIPTION**

We achieved successful results by processing millions of data points, to classify users into addressable SmartGroups. Using clustering methodology & analysis of shared traits, users are assigned into a single SmartGroup, as we have found users who fall into the same SmartGroup, react in a similar way to ads.

The technology that creates SmartGroups, uses real-time signals to capture the defining details of the moment when a user lands on a page, including time of day, weather, context of page, browser, device, etc.

This moment-based model means that users move fluidly between SmartGroups every time they land on a different page in our network; as such they will belong to different SmartGroups at different times – and we have found understanding how certain SmartGroups will respond to an ad, outperforms traditional tactics based on user identity.

None of the signals used are cookie-based or utilize personally identifiable information, which makes this solution compliant with GDPR & CCPA regulations.
All the features used to assemble the SmartGroups are moment-based and rely on request-level attributes only. The fact that the model does not rely on campaign specific attributes allows us to use the inferred Smart Groups to optimize various business KPI’s, without building a specific model for each.

Most classic machine learning clustering models are built from numeric data as they use Euclidean distance to find the groups’ centroids and decide which group a data point should be associated to. SORT™, on the other hand, uses an algorithm which can be trained on a multitude of flexible data signals. Instead of measuring distance between points via the Euclidean model, it calculates the similarities in groups versus the differences.

The SmartGroup training of the model is iterative until convergence. At the end of the process, we have a model at hand which divides our historical traffic and signals to a predefined number of SmartGroups, and we can use it to predict the most relevant group for new observations as well.

This clustering model is an unsupervised and autonomous one, meaning that there isn’t a clear target feature that we are trying to predict, so we cannot measure retroactively how far the model results are from the real outcomes. Therefore, we had to use creative ways for measuring the accuracy of the model and whether it can fulfill its purposes.

The model’s main purposes are: 1. Predict the SmartGroup for each new request that comes in. 2. For each campaign, find the best performing SmartGroups and adjust its delivery to prioritize them, under the assumption that for any given campaign, users from the same SmartGroups will react similarly.

The desired number of SmartGroups was determined by optimizing the two criteria detailed below (together with some other production and technical considerations), which included experimenting with different SmartGroup numbers. At the end of this research, we determined that an optimal number of SmartGroups to start with, was 30 SmartGroups.

**GOAL #1** Verify that different SmartGroups have a CTR (Click Through Rate) that is significantly different from the total average and is statistically significant for observations. This would mean the model’s predictions are effective enough for distinguishing high vs. low performing Smart groups per campaign.

**VERIFICATION:**
To confirm this criterion, we applied two protocols on same data set.
1. For each delivered impression, predict a user’s SmartGroup using the model, meaning the Smart Group the impression falls into.
2. Randomly assign a number between 0-29 to each delivered impression.
For each group in each one of the cases, a specific KPI (in this case CTR) was measured. In the graph below, you can see the distance of the KPI from its average for each of the SmartGroups for the two protocols. As expected, in the random protocol, all groups have similar KPI values and in the model strategy there is a large diversity in KPI between the groups – this is an advantageous situation as we can predict SmartGroups with extremely high performance (e.g., SmartGroup 24) versus SmartGroups with extremely poor performance (e.g., SmartGroup 17).

Additionally, you can see in the pie chart below that there are enough observations (for statistical inference) for each SmartGroup, based on a data set of about 3MM impressions.
GOAL #2 Correlation between SmartGroup performance by training set and test set.

VERIFICATION:
Data set was divided into train and test:
- Training data was used to train and build the model, which ended in dividing the train set impression data into 30 SmartGroups.
- Performance KPI (distance from average KPI) was measured for each SmartGroup in the train set.
- The model was used to predict SmartGroup per each impression in the test set.
- Performance KPI (distance from average KPI) was measured for each SmartGroup in the test set.
- We calculated the correlation between the performance of SmartGroups in the train set versus performance of the same SmartGroups in test set.

The following graph illustrates the results for a specific campaign. As evidenced, the model’s predictions on the test-set’s population generated similar KPI patterns per cluster as in the train-set population. The methodology managed to identify the same correlation trends among all the campaigns that were tested. This effectively means that SORT can predict, with very high certainty, the audiences that would interact well with the creative versus the ones which wouldn’t.

PRODUCT DESIGN
OFFLINE: The SmartGroups AI model runs periodically based on the latest data. A representation of the model is saved and then loaded to the Undertone ad server memory.
REAL TIME: Two key procedures:

- For each ad request – there is real-time low latency SmartGroup association prediction (<10ms) and logging. This enables SORT to analyze the best performing SmartGroups per campaign and create data-rich dashboards, along with ongoing analysis to determine the performance of SmartGroups for different brands.

- The model functions by ongoing processing of the data, analyzing the best performing SmartGroups for each campaign (according to pre-defined KPIs) and prioritizing the campaign delivery towards those SmartGroups, without compromising the scale of the delivery.

The following flow chart describes Undertone’s Product data flow:

RESULTS FROM PRODUCTION

After a period of testing this product offline and in real environments, across 5 pilots, the technology was released under the name SORT™ - Smart Optimization of Responsive Traits and is now being used in production across a variety of brands and advertising campaigns run by Undertone, a Perion company.

To date, we are seeing SORT™ consistently outperforming traditional 3rd party cookie targeting methods across varying verticals, formats and KPI’s (tested via A/B testing), with up to 112% lift in Click-thru Rates, 60% lift in Interaction Rates and over a 14% decrease in Cost-Per-Site Visit.

KEY BENEFITS OF SORT™ TECHNOLOGY:

- Provides privacy & performance. Outperforming cookie-based methods across all High Impact formats by up to 2X
SUMMARY

In this whitepaper, Undertone described and validated a sophisticated AI-based clustering technique, verified the accuracy of the results offline and built a substantial architecture to deliver at scale. The technology enables advertising and brands to achieve superior results to third-party cookies – with valuable marketing insights – while being privacy compliant and staying ahead of the trends. We are continuously working on strengthening the model to produce even deeper insights and results for advertisers in the future.

- Fully privacy-compliant targeting solution that does not collect any PII and adheres to all industry regulations
- Scalable across all websites, browsers, devices, and ad formats
- Inclusive vs Exclusive approach to targeting - no pre-determined bias
- All targeting is based on real-time data signals and the specific combination of attributes present in the moment someone lands on a page in our network (ie: human-based vs being just a data set)
- Can be used in place of or as a complement to existing targeting strategies such as 1st party targeting.
- Certified by Neutronian, Trademarked and Patent-Pending
- Easy to activate. SORT™ is available Direct & Programmatic Guaranteed with no opt-ins/ integrations required by the advertiser, consumer, or publisher.

“The Colorado Tourism Office has been hard at work discovering targeting solutions that are future-proofed. So, when presented with the opportunity to test out Undertone’s new cookieless technology, SORT™, we jumped at it. In the tests we ran, we were able to reach entirely new audiences by scaling our delivery on previously inaccessible browsers, AND results also significantly exceeded our cookie-based targeting tactics. Scale, privacy & performance? Can’t beat it!”

VIEW DEMO