Measuring the effectiveness of different privacy tools that limit online behavioral tracking

Omar AbouelNour [1] **Bilal Munawar** [1]

Abstract

Online behavioral tracking is the practice of analyzing users' online activity to display tailored advertisements to them. This violates users' privacy because third party tracking tools exploit users' confidentiality and integrity. In this paper, we propose and theoretically analyze a methodology to measure the effectiveness of different privacy tools that limit online behavi oral tracking.

1. Problem Statement and Motivation

Online Behavioral targeting places digital tags on the users' web browsers and utilizes these digital tags to track consumer behaviour. In particular, Online Behavioral targeting collects a user's digital footprint, such as the browser history, ads clicked and products purchased. Subsequently, the user's personal information is used to match the advertising content and optimise the advertisement display frequency across the entire online traffic. There are several technologies which enable effective behavioral targeting. These include cookies, deep packet inspection and device fingerprinting (Borgesius, 2013).

Online behavioral targeting collects personal information without user's consent and evades user's confidentiality. A number of users have expressed valid privacy concerns against behavioral advertisement. A recent case study (Dwyer, 2009) of consumer tracking on `Levis.com` demonstrates a violation of the ethics of consumer privacy and illustrates the behavior of online behavioral tracking. This case study shows that `Levis.com` links a total of nine tracking tags to its online traffic which link to eight third party trackers. These third parties are not acknowledged in the privacy policy of `Levis` and violates the trustworthiness of its online consumers.

There are a number of present day tools which limit online behavioral targeting. However, these existing solutions either block all online tracking, such as access to real time location for maps or require significant modifications to the existing tracking infrastructure. These limitations render these technologies as ineffective and impractical (Katz-Bassett, 2014).

Online shoppers form a core aspect of the e-commerce industry, which has seen exponential growth over the last decade. In a national privacy survey, an approximate of ninety percent Americans expressed serious concerns about their online privacy (TRUSTe, 2012). Third party websites can purchase confidential user data from authorized websites that sell a product to online users. An article from Times Magazine (Stein, 2011) reports that behavioral targeting has allowed for complete violation of users' privacy, rendering users demographics details, social media preferences and online purchase history through a combination of user's name and email address. Online behavioral targeting violates the trustworthiness of online consumers on e-commerce sites. Therefore, studying the limitations of existing tools and proposing new methods to measure the effectiveness of privacy tools which limiting online behavioral tracking is essential to the functioning of the e-commerce ecosystem.

In this paper, we will propose and theoretically analyze a methodology to measure the effectiveness of privacy tools which limit online behavioral targeting.

2. Approach

In this section, we propose a method to measure the effectiveness of privacy tools which limit online behavioral tracking and subsequently allow for comparison among these tools.

A. *Data Collection*

In this section, we plan to introduce a methodology to collect data in order to detect online behavioral targeting. We create a number of online dummy profiles to mimic consumer behavior on the internet. We define a unique set of interests, hobbies, demographics, and social media preferences for each of our profiles. Subsequently we

[1] New York University, United States. Correspondence to: Omar AbouelNour <oa767@nyu.edu>, Bilal Munawar <bm2515@nyu.edu>.

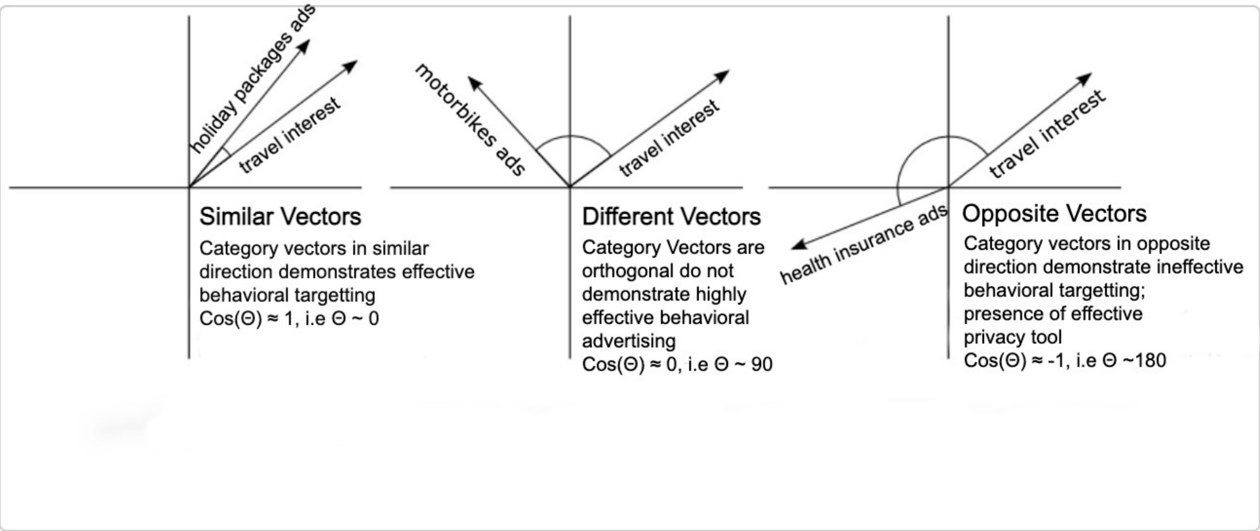
holiday packages ads
travel interest
Similar Vectors
Category vectors in similar direction demonstrates effective behavioral targetting
Cos(Θ) ≈ 1, i.e Θ ~ 0
motorbikes ads
travel interest
Different Vectors
Category Vectors are orthogonal do not demonstrate highly effective behavioral advertising
Cos(Θ) ≈ 0, i.e Θ ~ 90
health insurance ads
travel interest
Opposite Vectors
Category vectors in opposite direction demonstrate ineffective behavioral targetting; presence of effective privacy tool
Cos(Θ) ≈ -1, i.e Θ ~180

Figure 1. This figure demonstrates three different scenarios, which exhibit a decrease in cosine similarity as the privacy tools become more effective

follow the protocols of training and testing in order to measure behavioral targeting for each of our profiles.

B. *Training and Testing*

We create unique browsing behavior and purchase history on particular advertisement topics for each of our profiles. We surf on a total of fifteen sites which offer particular products or services, for example travel and furniture. In this process, we store the cookies on the web browser because it allows third party tracking websites to enable tracking. This process is known as training (**Algorithm 1**) because it trains the tracking algorithm of these tags to learn about the consumer behavior, such as the user's interests and preferences about a particular product.

After storing consumer behavior data through cookies on training web pages, we measure online behavioral advertising through visiting test sites, such as articles, news, and sports channels, which typically display advertisements to the users based on their past consumer behavior. We measure behavioral targeting through analyzing the similarity between the displayed advertisement and consumer interests. This process is called Testing and is demonstrated in **Algorithm 2**.

We repeat the complete process of training and testing in presence and absence of privacy tools. Subsequently, we compare the displayed advertisements on test web pages in order to measure and compare the effectiveness of different privacy tools through cosine similarity.

Algorithm 1 (Balenbako, 2012): This Algorithm is used to train tracking tags from third party sites to learn about consumer behavior. This process is known as **Training**.

```
for each product in (travel, sports, furniture, electronics, winter apparel) do
  for each training page in product (levis.com, target, amazon, wayfair) do
    repeat
      visit the training web page
      click on items of interest
      close the browser and save cookies
    until 10 times
  end for
end for
```

Algorithm 2 (Balenbako, 2012): This Algorithm is used to test online behavior targeting by visiting test pages. This process is known as **Testing**.

```
for each product in (travel, sports, furniture, electronics, winter apparel) do
  for each test web page in product (Times magazine, Atlantic, economist) do
    repeat
      visit the test web page
      save the first five ads
      close the browser and save cookies
    until 5 times
  end for
end for
```

C. *Methodology*

We categorize each online product offering on online advertisements into different sets. Each dummy profile is assigned an interest category and measured for behavioral targeting. We compare and mathematically compute the similarity between the displayed advertisements and consumer interests for each of our dummy profile. As a result, we are able to statistically compare if the targeted advertisement matches the user's browsing history, analyzing the effectiveness of privacy tools to limit the online behavioral tracking.

We propose the cosine similarity methodology to measure the similarity between the displayed advertisements and consumer interests. The cosine similarity is a measure of similarity between two non-zero vectors projected in a multi-dimensional space. The cosine similarity is calculated by computing the cosine of the angle between two vectors. As the angle increases from zero to 180, the similarity between the two vectors decreases, whereas the two vectors are identical if the angle between these vectors is zero.

Mathematically, cosine similarity is defined as:

$$\frac{\tilde{A}.\tilde{B}}{||\tilde{A}||.||\tilde{B}||}$$

In our study, the two non-zero vectors are the product category displayed in the advertisement and the interest category of the dummy profile. We propose to model a cosine curve for the first ten advertisements each dummy profile is displayed on test web pages. If the behavioral targeting were to be highly accurate, the categories displayed on the advertisements will match the interest categories, producing identical vectors, with an angle of zero between them. Subsequently, we will use the notion that since the cosine of 0 (angle between the two vectors) is 1, there should be some deviation along the cosine curve if the privacy tool were to limit behavioral advertising. Each time the cosine curve deviates from 1, the targeted ads become less accurate and the privacy tools become more effective. Therefore, we propose to study how the cosine curve deviates from the origin based on the privacy tool for each dummy profile. We compute the average horizontal distance of the cosine curve from the origin across all the profiles to measure an average cosine similarity for a privacy tool. Subsequently, we repeat this process for different privacy tools to allow for comparison among them.

We propose the Word2Vector framework to represent advertisement topics and consumer interests as vectors in a multi-dimensional space. Word2Vector embodies a neural network which learns to represent key words surrounding an advertisement topic as vectors based on word similarity. Thus, advertisement topics which align with consumer interests are displayed as similar vectors. This frameworks allows us to mathematically represent advertisement categories and consumer interests and compute the respective cosine similarity. As a result, we are able to measure the effectiveness of privacy tools using cosine similarity.

3. Related Work and Novelty

In this section, we will discuss other papers that proposed methods to measure the effectiveness of different privacy tools to limit online behavioral tracking.

Pedro Leon et al. evaluates the effectiveness of privacy tools that limit behavioral tracking by comparing how easy they can be installed and how successful they are at blocking advertisements. The results determined that the most common privacy tools available were ineffective since they not only failed to protect users' privacy but also hindered their online experience by blocking several websites. Although the research is helpful, it only measures the simplicity of installing and using the different privacy tools instead of mathematically comparing their effectiveness.

Ajdari et al. uses a different technique to analyze the usefulness of different privacy tools. The author compares the top one hundred ranked websites in the U.S. depending on the resulting number of cookies and different page load times using various privacy tools. The results ranks website in order of protection against behavioral tracking but does not propose a method to measure the effectiveness of privacy tools.

Balebako et al. proposes two methods to measure the effectiveness of privacy tools. Balebako analyzes and compares the displayed advertisements with user interests. These two proposed methods include (a) comparing the advertisements' display URLs and (b) comparing the words that appear in the title and content of the advertisement. These comparisons are made using cosine similarities. Figure 5 shows the different measures of cosine similarities for each privacy tool, depending on the advertisement topics. "No tool" represents the cosine similarities measured in the absence of a privacy tool.

The results found in Balebako's paper are helpful and give us an insight in the efficiency of different privacy tools. However, there are certain limitations in the proposed methodology to measure the performance of privacy tools. In this research study, certain advertisement topics were excluded from the study because there was an insufficient evidence of behavioral tracking. This highlights discrepancies in online behavioral targeting and raises questions about the methodology used to measure behavioral targeting on advertisement topics. In addition, the study is only limited to testing Google advertisements due to several constraints in the proposed methodology, such as comparison using displayed URLs and key search words. Therefore, the results of

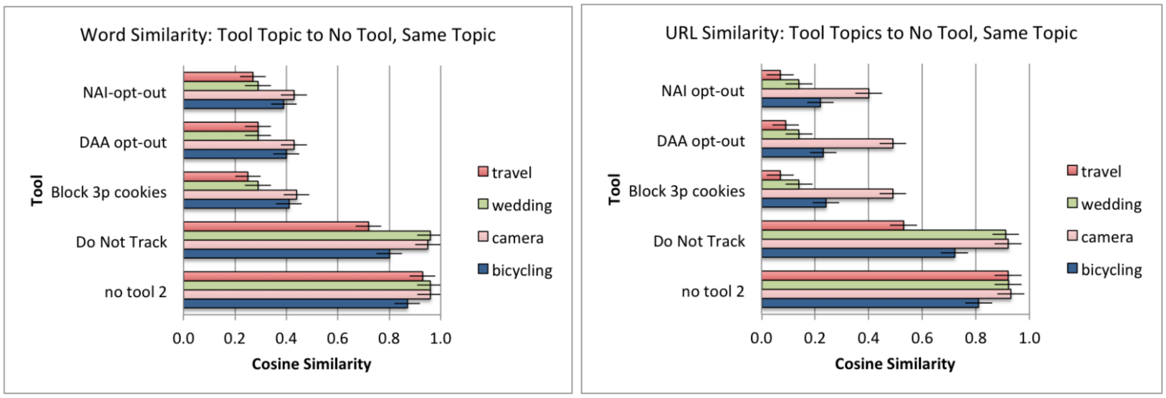

Fig. 5. Results from each tool compared to ads from no-tool on the same topic. Effective tools have low cosine similarities when compared with no-tool as they reduce the number of behavioral ads. Abine Taco and Ghostery eliminated the Google text ads, so they are not shown.

Fig: 5 (Balebako, 2012)

this study are not representative of the entire advertisement infrastructure and its domain. We propose a similar testing and training framework but tweak Balebako's methodology to overcomes these limitations and produce representative results on the efficacy of privacy tools that limit behavioral tracking.

4. Evaluation Approach

To determine whether the methodology proposed effectively measures the efficacy of different privacy tools, multiple baseline measurements must be created.

First, without using any of the algorithms discussed, the advertisements displayed to a user from the different test pages will be saved and categorized depending on different topics, e.g. occupation, age, hobbies, etc. Then, the training algorithm will be applied, and the advertisements displayed to the user would be saved and categorized once again. Comparing both sets of advertisements, from before and after applying the algorithm, will help verify whether behavioral tracking is in fact taking place.

If behavioral tracking is present, the baseline measurements used to determine the effectiveness of the different privacy tools will be computed. For each different advertisement category, a baseline measurement will be calculated by comparing that category's set of advertisements with the user's interests using cosine similarities. This process will be repeated one hundred times without any privacy tools used, under the same conditions, to eliminate any external factors from affecting the guidelines used to compare the different privacy tools. For each category, the mean of the values from the identical tests performed will be that category's baseline measurement.

The various privacy tools' effectiveness will be analyzed by performing both algorithms and comparing results of the cosine similarities with the baseline measurement found. If the results are similar, meaning the difference between the two cosine similarities is small, then it can be concluded that the specific privacy tool being tested was not effective at limiting behavioral tracking. If the results are different, meaning the difference between the two cosine similarities is large, then the privacy tool tested was effectively limiting or eliminating behavioral tracking.

Overall, as the difference between the two cosine similarities being compared increases, the privacy tool being tested is more effective.

References

1. Ajdari, D., Hoofnagle, C., Stocksdale, T., Good, N. (2013). Web Privacy Tools and Their Effect on Tracking and User Experience on the Internet.

2. Balebako, R., Leon, P., Shay, R., Ur, B., Wang, Y., Cranor, L. (2012, May). Measuring the effectiveness of privacy tools for limiting behavioral advertising. Web.

3. Borgesius, F. Z. (2013). Behavioral targeting: A European legal perspective. IEEE security privacy, 11(1), 82-85.

4. Dwyer, C. A. (2009). Behavioral targeting: A case study of consumer tracking on levis. com. Available at SSRN 1508496.

5. Katz-Bassett, E., Heidemann, J. S., Godfrey, B., Feldmann, A. (2014). Proceedings of the 13th ACM Workshop on Hot Topics in Networks. In 13th ACM Workshop on Hot Topics in Networks. ACM.

6. Leon, Pedro, et al. ”Why Johnny can't opt out: a usability evaluation of tools to limit online behavioral

advertising.” Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012.

7. TRUSTe (2012). Consumer privacy index-Q1. Retrieved from http://www.truste.com/consumerprivacy-index-Q1-2012/