For Your Eyes Only: Consuming vs. Sharing Content

Ram Meshulam
Outbrain Inc.
39 West 13th Street New York
NY 10011, United States
rmeshulam@outbrain.com

Roy Sasson
Outbrain Inc.
39 West 13th Street New York
NY 10011, United States
rsasson@outbrain.com

ABSTRACT

This paper analyzes two types of user interactions with online content: (1) private engagement with content, measured by page-views and click-through rate; and (2) social engagement, measured by the number of shares on Facebook as well as share-rate. Based on more than a billion data points across hundreds of publishers worldwide and two time periods, it is shown that the correlation between these signals is generally low. Potential reasons for the low correlation are discussed, and the notion of private-social dissonance is defined. A more in-depth analysis shows that the dissonance between private engagement and social engagement consistently depends on content category. Categories such as Sex, Crime and Celebrities have higher private engagement than social engagement. On the other hand, categories such as Books, Careers and Music have higher social engagement than private engagement. In addition to the offline analysis, a model which utilizes the different signals was trained and deployed on a live recommendation system. The resulting weights ranked the social signal lower than click-through rate. The results are relevant for publishers, content marketers, architects of recommendation systems and researchers who wish to use social signals in order to measure and predict user engagement.

Keywords

Behavioral Modeling, Facebook, Recommender System, Social Network

1. INTRODUCTION

From the early days of social networks, academic researchers and commercial firms have been trying to leverage social engagement data to improve users' content experiences. Some have tried to incorporate social data to improve personalization [2, 8]. Others have used it in order to predict content popularity [1, 9] or to predict whether content will become viral [3, 5, 6]. Publishers also try to infer from social signals which stories are currently trending and should be covered more extensively. The underlying assumption behind all of these efforts is that users' actions with respect to content on social networks reflect their content consumption preferences.

However, there is evidence that the content-related actions of users on social networks are not fully correlated with their preferences outside the social network [10]. For example, there is work aimed to understand and predict which content has the potential to become viral. Perhaps surprisingly, the answer is not simply 'content that was selected for consumption by many users' i.e., popular content. Instead, only specific content types have such potential. Moreover, studies of user behavior and user identity on social networks show that users do not share every content item they enjoy reading. Many times, the act of content sharing is motivated by self-promotion of the user.

This paper utilizes an extensive dataset from 200 large publishers to analyze the relationship between consuming content and sharing content. The contributions of the paper are as following: (1) we show that the overall correlation between social signals and content consumtion signals is low, and discuss the different biases each signal incorporates (2) show that the dissonance between social engagement and click-through rate varies across content categories (3) show that on a click-prediction model, the weight of the social signal is significantly lower than the weight of click-through rate.

Our findings could be relevant to publishers, content marketers and professionals in the domain of recommendation systems, as it highlights both the value and concerns of utilizing social signals to measure and predict user engagement.

The reminder of this paper is as follows: section 2 presents related work. Section 3 shows the correlations between different signals which were collected at the article level, defines the dissonance between the signals and analyzes it at the content category level. Section 4 demonstrates the value of each signal for click prediction under a live recommendation system. We conclude the paper in section 5.

2. RELATED WORK

The related work comes from three domains, which all use data from social networks while aiming for different goals: (1) understanding (and predicting) viral content; (2) studying user behavior and identity on social networks; and (3) improving the performance of recommendation systems by incorporating data from social networks.

Castillo et al. [1] studied the life cycle of news articles. The study concentrated on 'news' and 'in-depth' sections of

WWW'16 Companion, April 11–15, 2016, Montréal, Québec, Canada.
ACM 978-1-4503-4144-8/16/04.
http://dx.doi.org/10.1145/2872518.2890094.

one publisher. They showed that the life cycle of articles in terms of visits distribution, tweets and shares distribution over time vary across different sections of the publisher. In their work they were able to improve, for some content types, the prediction of traffic to articles using data from social networks.

Another effort to predict the popularity of content, based on early signals, was presented in [9]. The goal was to predict the number of 'diggs' on Digg[1] and views on YouTube[2] based on data (both popularity and social data) on the first few hours after publication. The authors found that the social data was not a good predictor for popularity, and concluded that it is due to the way that Digg promotes new content.

A different usage of social data is to improve the accuracy of recommendation systems [2, 8]. The underlying assumption is that users with social connections share content preferences, ratings and consumption patterns. It improves the similarity measurement between users, and improves collaborative filtering and other recommendation algorithms which utilize user-user similarity.

Berger & Milkman [3] tried to characterize what makes content viral. Their analysis is based on 'the most e-mailed' article list of the New York Times[3]. The study is focused on the psychological impact of the content as a feature, which can detect viral content. Specifically, content which triggers awe, anger or anxiety is more likely to become viral. Investigating online videos on social networks yielded similar results [5, 6].

Zhao et al. [10] studied the behavior of users on Facebook. They show that users build their online identities carefully, making more implicit identity claims such as wall-posts and pictures, as opposed to explicit claims as filling the 'about me' form. Instead of reflecting the user's (offline) identity, the Facebook profile shows what the users aspire to be.

Das Sarma et al. [7] have focused on predicting shares on an e-commerce platform. The authors found that basing the model on page-views and purchase actions in addition to share action has increased the accuracy of the model. On a related work [4] the authors have identified a gap (called 'the expression gap') between social shares and views or purchases on a commerce platform. In a way, our paper extends their work by analyzing the characteristics of this gap in the domain of online content, comparing it to other signals and showing a case study on a live recommendation system that utilizes social signals and private signals.

3. SOCIAL VS. PRIVATE ENGAGEMENT

This section shows the correlation between various social and private engagement signals and discusses the limitations of each signal.

3.1 Dataset and Background

Outbrain Inc.[4] is a content discovery firm. The 'Outbrain Engage' platform enables publishers to install a widget on their article pages, thereby allowing Outbrain's recommendation system to select and serve the best content of the publisher on it. Outbrain's engine powers recommendations for publishers all over the world, driving tens of millions of clicks per day. In order to determine which recommendations to present, Outbrain's servers collect multiple events for each article page, which are the basis for this work. The analysis presented in this paper is based on newly-published articles on Outbrain's network during the week of November 9th 2015. To validate that our results do not capture merely temporary interest or behavior on certain content types, the results were cross validated with earlier data from the week of February 23rd 2015, more than eight months earlier (denoted as *secondary dataset*). The results were consistent across time.

For each article, we measure the number of page-views during the first three days since its publication. To obtain homogeneous and reliable measurements, the data includes only articles which were published during weekdays, with at least 500 page-views, and within Outbrain's largest 200 publishers. The average age of an article across page-views was around four days, as many of Outbrain's large partners are news oriented publishers. The final data-set for the week in November is comprised of 49K articles, accounting for 825M page-views, out of which 713M are circulated page-views and 112M are Facebook referrals. From Facebook, 47M shares were collected. The figures for February data were within similar ranges.

For each article d, the following data was collected:

Circulated Page-Views: denoted PV_d. This is the number of page-views where the referring page (i.e., the address of the page that linked to the article) is a page of the publisher itself. In many cases, the referring page is the publisher's homepage.

Facebook Referrals: denoted ref_d. The number of page-views where the referring source is Facebook.

Outbrain Clicks: denoted $clicks_d$. The number of times that a link to d was recommended on Outbrain's widget, and the user has clicked on it.

Outbrain Click-Through rate: denoted CTR_d. The number of Outbrain clicks divided by the number of times the article was recommended by Outbrain's widgets.

Facebook shares: Denoted $shares_d$. the number of times the article was shared via Facebook share button, as reported by Facebook[5]

Based on the above measurements the following two signals were calculated:

Facebook share rate: defined as $shareRate_d = \frac{shares_d}{PV_d}$.

Facebook referrals rate: defined as $refRate_d = \frac{ref_d}{PV_d}$

Measuring private user-engagement is not trivial. It is not possible to show the users the whole inventory of articles and measure their willingness to read each document. The subset of articles which are presented to the user is likely to have much more views and engagement signals than the rest of the inventory. For example, a very large portion of page-views is referred from the publisher's homepage. Thus, measuring engagement in terms of page-views is heavily biased by the editor's choice for the homepage. Moreover, the position of the article on the homepage has also a large impact on the number of page-views it gets. Measuring the absolute number of Facebook shares also suffers from a bias. Even without the bias from the publisher's side, in most cases the sharing of content is done from the content page itself, which means it is heavily influenced by the number of page-views.

[1]digg.com
[2]youtube.com
[3]nytimes.com
[4]outbrain.com

[5]using FB api available on: developers.facebook.com/docs/graph-api

Table 1: Correlation between signals

	page-views	shares	share-rate	referrals	referral-rate	clicks	CTR
page-views	1	.52	**-.09**	.57	.18	.47	.21
shares		1	.75	.82	.69	.44	.23
share-rate			1	.53	.71	.17	.11
referrals				1	.86	.54	.33
referral-rate					1	.39	.28
clicks						1	.59
CTR							1

Facebook referrals might also be prone to biases, which are similar to those of page-views. Some publishers have a solid presence on social networks, including many followers. The publisher's selection of articles to include on its Facebook page will then translate to larger number of referrals as well.

To partially solve this bias, some signals can be normalized: we can measure CTR and share-rate instead of clicks and shares respectively. However, this section shows that share-rate might not be a good proxy for user engagement as it reflects other considerations of the user.

3.2 Correlation between Signals

We would like to analyze the correlation between the different signals. To compute the correlation between a pair of signals across publishers, we first calculated the Spearman correlation between signals for each publisher, and then calculated the average correlation, weighted by the number of page-views per publisher. Table 1 shows the correlation between each pair of the measured signals. Cross validating the results with the secondary data set yielded a negligible difference of up to 10% in overall correlations. The first row shows the correlation between page-views and all other signals. Page-views are highly correlated with shares (0.52). This can be explained by the selection bias - a share action is in most cases a result of a page-view. An article with a high number of page-views has many opportunities to be shared. In light of this, one would expect an even higher correlation between page-views and shares than observed. Furthermore, the correlation between page-views and share-rate is small and negative (-0.09). These two observations suggest that the number of shares is not a constant fraction of page-views, which depends on the general willingness or ability of the user to share content. Indeed, in Section 3.3 we show that users apply different sharing policies for different content types.

Columns 5 and 7 on the first row show there is relatively higher correlation between page-views and CTR (0.21) as well as referral rate (0.18). This can be explained as all three signals are a reflection of a user's choice to click: the metric of page-views indicates the choice to click from what is available on the home page, CTR indicates the tendency to click on a recommendation in a widget and referral rate indicates a choice to click on a shared article when observing it on the user's newsfeed on Facebook. Note that unlike share rates, all of these choices are private and the user does not expose them publicly.

The last column shows the correlation between CTR and all other signals. Again, the correlation between CTR and share-rate is the lowest of all signals (0.11), while the referrals and referrals-rate are the highest. This also suggests that private actions, such as clicking on a suggested link in order to read an article (either suggested by a recommendation system or by another Facebook user) are different than sharing content with other users.

3.3 Private-Social Dissonance

The correlations presented in Table 1 suggest that the decision to consume content is not always correlated with the decision to share it. We call this the *private-social dissonance*. It is the dissonance between private engagement and social engagement. Formally, for a document d, it is defined as:

$$dis_d = log(\frac{r(ctr_d)}{r(shareRate_d)})$$

where $r(ctr_d)$ and $r(shareRate_d)$ are the CTR and share-rate percentiles of d in the publisher's inventory of documents. To reduce the impact of outliers, the log of the value is used. Note that if the rank of share-rate is higher than click-through rate, the dissonance is negative. Dissonance around zero means that the rank of the document in terms of CTR and share-rate is roughly the same.

We now turn to examine whether the dissonance between private and social engagements varies with the content of the article. To answer this question, each document was classified to a single category, according to the document's content. Outbrain's classifier is based on the open source project Vowpal Wabbit[6]. The ontology includes approximately one hundred categories. Figure 1 shows the average dissonance per category in a descending order (the numbers attached to each category are explained in the next paragraph). Cross validating the results with the secondary data set yielded a Spearman correlation of 0.8 with respect to the order of categories. The higher group of categories (full dots), from Adult to Relationships are the ones with the highest dissonance. Articles on these categories get clicked-on and read, but after they are read, they are not shared. The lower group of categories (empty dots), from Home Improvement to Music, are the categories with the highest *opposite* dissonance. Some of these categories are less popular, but once users read an article in one of these categories, they tend to share it. The rest of the categories, which are somewhere in the middle, are omitted for lack of space. Figure 2 shows some of the missing categories.

The dissonance is positive and very high on categories such as Adult, Sex, Crime and Celebrities. The content in these categories might be interesting to read, but might not be flatter nor serve the average user's social identity. On the opposite side, the dissonance is negatively high on categories such as Music, Wine, Books, Arts and Education - categories

[6]http://hunch.net/ vw/

Figure 1: Content categories with extreme private-social dissonance

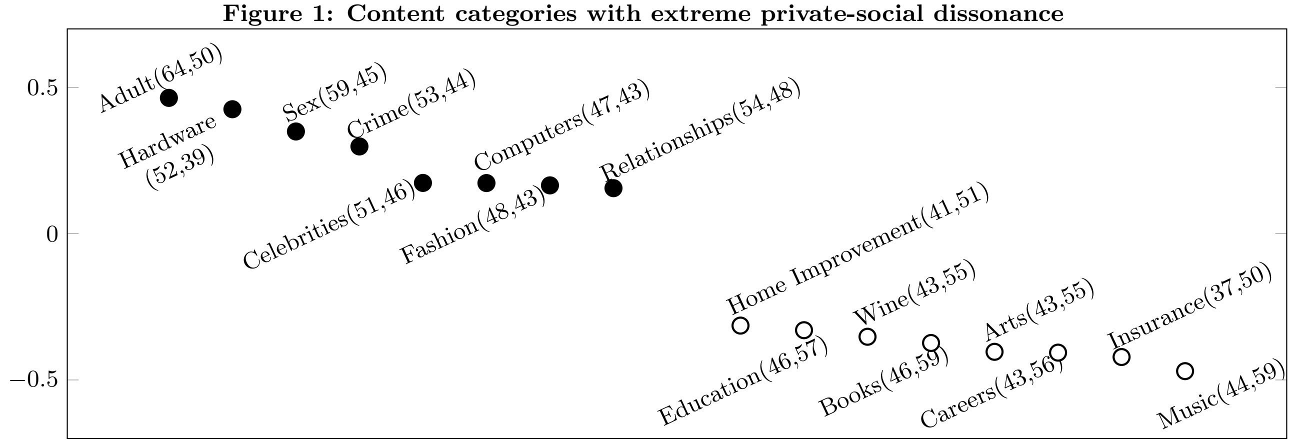
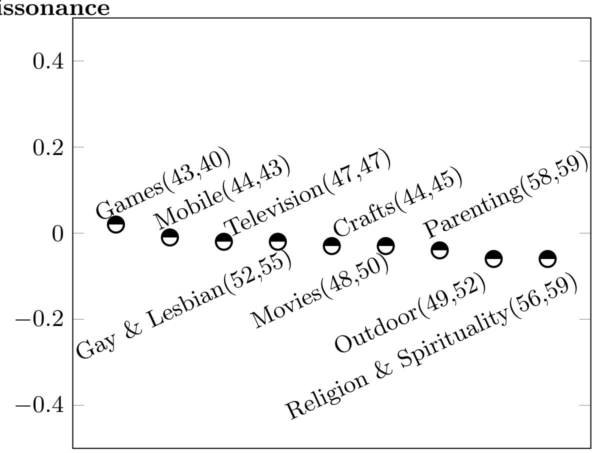


Figure 2: Content categories with no private-social dissonance



which might be considered to flatter and serve the average user's social goals.

Figure 2 shows a special class of categories, in which the dissonance is the closest zero. On categories such as Movies, Mobile, Parenting and Television, CTR and willingness to share are roughly equal.

Each category in figures 1 and 2 is accompanied by two numbers: the average percentile of CTR and the average percentile of share-rate for the documents in that category. Note that the dissonance is not correlated with popularity. For example, the following categories have documents with roughly the same average CTR levels, though they belong to three different dissonance groups: Fashion and Computers have high *positive* dissonance, Books and Education have high *negative* dissonance and Television and Movies categories have dissonance close to *zero*. Another example for similar CTRs and completely different levels of dissonance are Parenting and 'Religion & Spirituality' categories versus Sex category. All three categories show a very high and almost equal average CTR, but Sex has a very high positive dissonance while the first two have almost no dissonance. In other words, Parenting and 'Religion & Spirituality' show high CTR as Sex, but unlike Sex, show also high share-rate.

4. USE CASE: LEARNING TO RANK

The offline analysis in section 3 showed that there is low correlation between Facebook share-rates, publisher's page-views and widget's CTRs. We have argued that social signals confound what users are interested in reading with what they publicly reveal they are reading. However, as mentioned above, other signals suffer from biases as well, some of which might be unknown. In this section we put these signals to the test: we leverage both private and social engagement signals in one model and use them for making content recommendations in production. As we explain below, the setup in which this model makes recommendations is very unique as it partially overcomes the biases originating from the publisher's selection, making our conclusions more robust.

4.1 About Sphere

Sphere[7] is a web-based application integration platform. It enables developers to add Outbrain's personalized content recommendations as part of their applications. The user-contexts in which Sphere supplies content recommendations depend on the developed application. Contexts vary from standard content recommendations applications on mobile devices, to content recommendations that appear on the side of email services or on the lock-screen of a phone.

Sphere's recommendation modeling ranks articles from various publishers in order to maximize click-probability given a user and application. The supervision data that is used for Sphere's modeling utilizes user clicks within each application. It is ideal for testing true user interest in content, since (1) presentation is homogeneous regardless of the origin publisher; (2) the eligible content for recommendation is from many publishers; (3) no constraints are imposed on the recommendations.

4.2 Results

For Sphere's recommendations engine, we have trained and tested a model that utilizes three engagement signals in prediction of a user-click:

1. The CTR of the article on its original publisher.

2. Facebook referrals. Facebook referrals are used here as a proxy for shares, since it is available in serving

[7]sphere.com

time, unlike Facebook shares. The correlation between shares and referrals is relatively high (0.82), as shown in section 3.

3. Page-views: the number of page-views that the article had on its original publisher.

We used a regression model that combines the three signals with other personalization signals. Formally, the following model was used:

$$click = \alpha + \beta(x) + \gamma(p) + \delta(xp) + \epsilon$$

where α is a constant. x is a vector of observed signals - pageviews, CTR and Facebook referrals. p is a vector of personalization scores (which are out of the scope of this paper). Four applications of different natures participated in the test, where a separate model was trained for each application. Estimating the model on a test set showed that in every application, each of the three engagement metrics had statistically significant predictive power, though they varied in size, as reported below. Although weights varied quite substantially across applications, their order in terms of relative importance was found to be highly consistent. In each individual application, the weight of CTR in the model was the highest. In most cases it was followed by referral-rate which was 28% lower on average (the difference ranged between 9%-73% on individual applications). The page-views signal was ranked last in three out of four applications, with an average decline of 54% in relative weight in comparison to CTR. Reduction in weights between CTR and page-views ranged between 10% and 72% on individual applications. The performance of the trained models, estimated separately a model per application) was put under an A/B test in production and compared to a variant which utilized page-views signal only. During the test, 7.4M recommendations were served. The overall average lift was 15% in terms of CTR compared to the baseline variant.

5. SUMMARY & FUTURE WORK

The engagement level of users with online content is an important decision-making measurement. It can help publishers to decide what to write of next, or what content to promote, either on their own homepages or across other distribution channels such as sponsoring their content on other publishers. It is also an important signal for recommendation systems which aim to maximize user engagement.

Raw signals which measure absolute number of events suffer from various selection biases, and need to be normalized to overcome those. However, we have shown that Facebook share-rate suffers from a different bias, denoted as the private-social dissonance. These findings correlate with earlier work which showed that the incentive to share is not necessarily personal engagement with the content. The findings were based on more than a billion of user signals, and proved consistent over time. When used to predict a click on a live recommendation system, the weight of the social signals was not insignificant, yet it was weaker than CTR which seems to have the highest external validity of content attractiveness.

Moreover, the bias of share-rate is not distributed evenly across categories of content or CTR levels. Some categories have high engagement in terms of CTR and low share-rate, for example, Sex and Celebrities. Some have the opposite dissonance, for example, Wine and Arts. There are also categories for which users' personal engagement matches their willingness to share, such as Parenting and Movies. Publishers, marketers, architects of recommendation-systems and anyone who uses social signals as an engagement metric should be aware of the private-social dissonance.

This work can be generalized in several ways. One direction for future work is to investigate the relation between engagement and events other than Facebook shares. Twitter is an interesting candidate. Another direction is to use refined private-engagement signals, such as time-on-page or scrolling behavior. An interesting question can be - 'do users actually read what they share?'.

Another interesting direction for future work is to utilize the private-social dissonance in a classifier for inappropriate content: articles with high positive dissonance are many times inappropriate to some extent. Such a classifier is based on users' behavior and does not rely on natural language processing or image processing.

6. REFERENCES

[1]

[2] In *Data Mining for Social Network Data*, volume 12 of *Annals of Information Systems*. 2010.

[3] J. Berger and K. L. Milkman. What makes online content viral? *Journal of Marketing Research*, 49(2):192–205, 2012.

[4] A. Das Sarma, S. Si, E. F. Churchill, and N. Sundaresan. The "expression gap": Do you like what you share? In *Proc. of the 23rd Int. Conf on WWW*, WWW '14 Companion, Republic and Canton of Geneva, Switzerland, 2014. International World Wide Web Conferences Steering Committee.

[5] R. E. Guadagno, D. M. Rempala, S. Murphy, and B. M. Okdie. What makes a video go viral? an analysis of emotional contagion and internet memes. *Computers in Human Behavior*, 29(6):2312 – 2319, 2013.

[6] K. Nelson-Field, E. Riebe, and K. Newstead. The emotions that drive viral video. *Australasian Marketing Journal*, 21(4):205 – 211, 2013.

[7] S. Si, A. Das Sarma, E. F. Churchill, and N. Sundaresan. Beyond modeling private actions: Predicting social shares. In *Proc. of the 23rd Int. Conf. on WWW*, WWW '14 Companion, pages 377–378, New York, NY, USA, 2014. ACM.

[8] P. Symeonidis, E. Tiakas, and Y. Manolopoulos. Product recommendation and rating prediction based on multi-modal social networks. In *Proc. of the 5th ACM Conf. on Recommender Systems*, RecSys '11, pages 61–68, New York, NY, USA, 2011. ACM.

[9] G. Szabo and B. A. Huberman. Predicting the popularity of online content. *Commun. ACM*, 53(8):80–88, Aug. 2010.

[10] S. Zhao, S. Grasmuck, and J. Martin. Identity construction on facebook: Digital empowerment in anchored relationships. *Computers in human behavior*, 24(5):1816–1836, 2008.