The Workshops of the Eleventh International AAAI Conference on Web and Social Media
AAAI Technical Report WS-17-16: Observational Studies through Social Media

Estimating the Effect of Exercising on Users' Online Behavior

Seyed Amin Mirlohi Falavarjani,[1,2] Hawre Hosseini,[2] Zeinab Noorian,[2] Ebrahim Bagheri [2]
[1]University of New Brunswick
[2]Laboratory for Systems, Software and Semantics (LS3), Ryerson University

Abstract

This study aims to estimate the influence of offline activity on users' online behavior, relying on a matching method to reduce the effect of confounding variables. We analyze activities of 850 users who are active on both Twitter and Foursquare social networks. Users' offline activity is extracted from Foursquare posts and users' online behavior is extracted from Twitter posts. Users' interests, representing their online behavior, are extracted with regards to a set of topics in several subsequent time intervals. The shift of users' interests across different time intervals is taken as a measure of user behavior change on the social network. On the other hand, we employ user check-ins at a gym or fitness center as a sign of exercise and consider it to be an offline activity. In order to find the effect of exercise on online behavior, we identify users who did not go to the gym for at least two months but did so at least nine times in the next three months. We show that shift in interest reduces significantly for users after they start exercising, which implies that the offline activity of exercising can influence how users' interests are shaped and change on the social network over time.

Introduction

Randomized Controlled Trials (RCT) are often used to study the causal impact of a given treatment. In RCTs, subjects are randomly divided into treatment and control groups in order to minimize selection bias while other possible factors that might affect the outcome of the treatment are controlled to avoid confounding effects. This type of experiment is powerful and is a reliable technique for inferring cause-and-effect relationships. However, providing completely controlled and random environments is not always possible or feasible. In contrast, observational studies provide an alternative to RCTs that allow for a more relaxed way of studying possible causal relations. Observational studies are useful when an interruption, which is outside the control of the investigators, affects a group of subjects based on which the subjects can be divided into groups for comparative analytics.

Given the potential of observations studies, researchers have already engaged in performing such studies based on observations performed on social networks (Bagroy, Kumaraguru, and De Choudhury 2017: De Choudhury et al. 2017: De Choudhury and Kiciman 2017: De Choudhury, Sharma, and Kiciman 2016). For instance, in a recent research (Althoff, Jindal, and Leskovec 2017), the influence of online social networks on users' online and offline behavior has been studied. In their work, the authors use a dataset extracted from a fitness application that tracks user' steps during the day in conjunction with a social network where the users connect and communicate regarding their fitness activities. The objective of the study was to investigate whether users' online behavior on a social network can impact their offline activity in terms of factors such as the number of steps per day. They show that a social network, under certain circumstances and for a specific period of time, can impact users' offline activity. This finding reinforces the belief that online virtual motivators can influence users' real world decisions especially as it relates to physical activity.

In our work, we are interested in the dual to the same problem by investigating whether users' offline activity can impact their online behavior. To this end, we perform an observational study on the content that users' post on Twitter as well as the offline activities that they report on Foursquare. We view users' posts on Twitter as their online behavior (expression of beliefs) and their Foursquare check-ins as a model of their offline activities. While the work in (Althoff, Jindal, and Leskovec 2017) considers only online social network behavior in terms of activities such as friend requests, we process the content published by the users' on Twitter to build a model of their online behavior. Moreover, while (Althoff, Jindal, and Leskovec 2017) measures the number of steps per day for each user as a sign of offline activity, we measure user check-ins at different location categories on Foursquare as an indication of their offline activity.

More specifically, we consider exercise as an example of an offline activity, which can affect users' online behavior. In other words, we are interested to study whether those users' who exercise more often have different online behavior compared to those users who do not exercise frequently. To this end, we extract and view check-ins at locations that are categorized under 'Gym / Fitness Center' on Foursquare as a sign of exercising for the user. In addition, for the same users, we analyze the content that they have posted on Twitter and analyze the topics that they engage in to build a user interest profile for each of the users. We build user interest profiles based on the set of topics that each individual talks about within a specific time interval. The main research question that we would like to answer is whether engaging in an offline behavior such as exercising by going to the gym impacts how a user's interests shift over time. The method used for this purpose is *interrupted time series design with comparison group* in which a series of observations happens during the time intervals before and after the interruption. In order to find the effect of exercising on users' online behavior (users interest shift over time), we recognize users who did not exercise for at least two months but started to do so by exercising at least nine times in the next three months. In this study, we compare users' interest shift before and after the interruption and show that interest shift reduces after users start their exercise routine. To the best of our knowledge, this is among the first works, which examine the effect of offline activity on online behavior.

Related Works

Related work to this study can be considered from two perspectives, namely those works which have used observational studies and those which have focused on assessing personality aspects of social network users. From the first perspective, (Dos Reis and Culotta 2015) use observational studies with matching to infer causality between exercise and mental health. In (Anderson et al. 2013), the authors use quasi-experimental design to show how badges can influence user behavior on Stack Overflow. Similarly, (Oktay, Taylor, and Jensen 2010) also designed three different quasi-experiments to understand users behavior on a Stack Overflow. (De Choudhury et al. 2016) use an observational study with matching to understand the transition from mental illness to suicide ideation.

Also, there have been studies focused on assessing user personality aspects by examining their online behavior (Farnadi et al. 2016; Roshchina, Cardiff, and Rosso 2011). Most of these works use supervised methods and are based on big five personality traits (Gosling, Rentfrow, and Swann 2003) including openness, conscientiousness, extraversion, agreeableness and neuroticism. (Stephanie and Sklar 2016) use check-ins from Foursquare to design a new algorithm for location recommendation. Personality is used in some papers to improve the recommendation (Hu and Pu 2011; Roshchina, Cardiff, and Rosso 2011) and rating systems (Karumur, Nguyen, and Konstan 2016; Liu, Cao, and Yu 2016).

There are many impressive recent work on observational studies based on social networks such as (Olteanu, Varol, and Kiciman 2017; De Choudhury, and Kiciman 2017; Bagroy, Kumaraguru, and De Choudhury 2017); however, our work distinguishes itself by analyzing the impact of offline activity on online behavior in terms of user interest and interest shifts.

Study Design

In this section, we study the effect of users' offline activities, i.e., engaging in regular exercise on changing their online behavior, i.e., shift in interests. We show that engaging in exercising on a regular basis results in a more persistent user interest profile where the users' interests do not shift as much compared to the users who do not exercise. In other words, the shift of interests in the treatment group drops significantly compared with a matched control group.

Our study consists of two steps. In the initial step, we extract users' online behavior. Our work relies on building a user profile for several subsequent time intervals. Each user profile is a vector consisting of probability distribution over topics in a given time interval for that user. In the second step, and in order to infer a causal relationship, we adopt an interrupted time series design with comparison group. In this model, we identify two groups of subjects that are carefully matched that differ in the treatment, which is regular exercising after a certain point in time. More specifically, we find users that have not checked in at any gym or fitness center for a period of two months; however, they can be separated as the treatment group embarks on exercising at a certain point in time while the control group does not. By analyzing the user interest profile of the users before and after the interruption, we analyze the impact of exercising on user interest shifts. In order to eliminate the effect of confounding variables, we match users in the treatment group with similar users in the control group based on their pre-treatment online behavior.

User Interest Modeling (Online Behavior)

We model users' online behavior based on their user interest profile which is computed based on the active topics on Twitter in each time interval. We employ the LDA topic modeling approach to extract active topics across several subsequent time intervals. For K active topics across all time intervals $Z=\{z_1, z_2, \ldots z_k\}$, for every user u, we model user's interest profile for time interval t, $UP^t(u)$, which is the distribution of u's interests over Z in time interval t.

Definition 1 (User Interest Profile) The user interest profile of user u in time interval t, $UP^t(u)$, is represented by $(f_u^t(z_1), \ldots, f_u^t(z_k))$ which is a vector of weights over K topics where $f_u^t(z_k)$ shows the degree of u's interest in topic z_k at time interval t.

It should be noted that topic and user interest detection methods from microblogging services have already been well studied in the literature and therefore are not the focus of our work. For the sake of our experiments, we have used our own interest modeling work presented in (Zarrinkalam et al. 2016; Zarrinkalam et al. 2017).

Based on Definition 1, we formally define users' interest shift, as depicted in Figure 1, as follows:

Definition 2 (Shift in Interest) The shift in interest of user u in time interval t+1 is the sum of absolute differences of distributions for each topic between time intervals t and t+1:

$$\boldsymbol{SII_u^{t+1}} = \sum_{i=1}^{k}(\boldsymbol{f_u^{t+1}(z_i)} - \boldsymbol{f_u^{t}(z_i)}) \quad (1)$$

Shift in Interest shows how much a given user changes her interests between topics across subsequent time intervals. Users with lower shift in interest demonstrate a more consistent interest profile that does not change between time intervals. However, in contrast, users with a higher shift in interest have a higher likelihood of changing their interests frequently. In our study, the length of each time interval has been considered to be one month.

Matching

In order to form the treatment and control groups, we identified users that had not checked in at any gym or fitness center for a period of two months. These users were selected to be studied. Furthermore, we then selected users for the treatment group based on whether they started checking in at a gym or fitness center at least nine times in the next three months. In other words, we consider those users who go to the gym at least once every ten days to form the treatment group. We match each user in the treatment group to a similar user in the control group based on their pre-treatment behavior defined based on their SII in the leading two month period of the study. We will also later show that external variables are balanced in our study.

Experiments and Dataset

Dataset Description

We collected data for 850 users as presented in (Han Veiga. and Eickhoff 2016). These users are active across two popular online social networks: Twitter and Foursquare. The reason for using these two networks is that the available data is very comprehensive and representative of different use

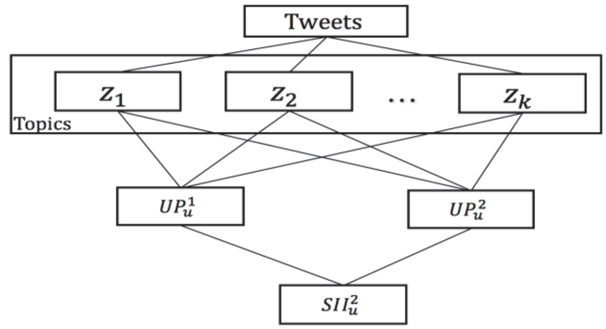

Figure 1. Shift in Interest for user u between time intervals 1 and 2.

cases of social networks. In our work, we use Twitter data to analyze users' online behavior and Foursquare data to understand users' offline activities. In (Han Veiga and Eickhoff 2016), the authors found and matched users who are active on Twitter, Instagram, and Foursquare. They found 850 users who fulfill three criteria of actively using the three mentioned social networks, posting in English, and sharing their content publicly. Their data was collected in January and February 2016. For our work, we collected data for these users for a period of five months from both Twitter and Foursquare. Our dataset is specifically useful for the analysis of the impact of users' offline activities on online behavior since it contains over 2.5M tweets and 450K check-ins. Tweets are representative of user online behavior and check-ins are representative of user offline activities. The timestamp information of the tweets and check-ins enables us to analyze tweets and check-ins in the same time intervals.

Results and Evaluation

In our experiment, we extracted the recent tweets of the 850 users presented in (Han Veiga and Eickhoff 2016). These users are active on twitter and they all use the Swarm application, which is an application from Foursquare in which the users can share their check-ins via Twitter. We look up the Swarm ID of each location mentioned in a tweet and extract the details of the location using Foursquare API. In order to calculate the shift in interest values for each user, we run the Mallet implementation of LDA with 100 topics on all the tweets gathered in the five month time period. We then divided the tweets of each user into one-month time intervals producing five sets of tweets per user. Based on these five sets of tweets, we derive the interest profile of each user in that time interval based on the learnt topic model. We find the sum of differences in topics distribution between two sequential time intervals for every user to allow for the calculation of the shift in interest based on Equation 1. As the treatment group, we include users who do not go to the gym

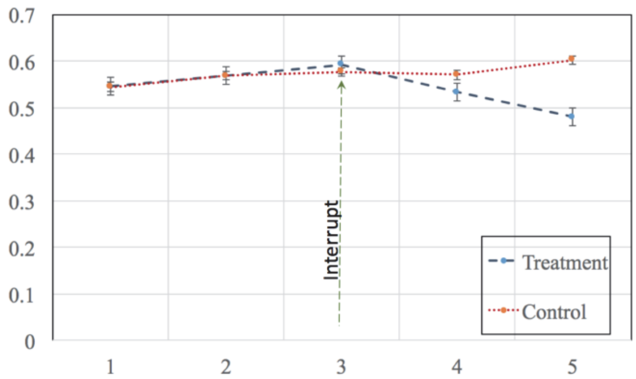

Figure 2. Results of the shift in interest experiment comparing the treatment and control groups.

or fitness center for at least two subsequent months and start to go to the gym or fitness center after an interruption in the third month and continue going to the gym at least nine times for a period of three months after the interruption. There are 80 users with such conditions in our dataset. As the control group, we find users who have the same behavior in interest shift as users in the treatment group before the interruption and matched every user in the treatment group with a user in control group. As our results show, the amount of interest shift of the treatment group reduces after the interruption.

A very important issue that should be considered in inferring causal relations is the effect of external variables. The solution to reduce the effect of external variables is to randomly divide the users into treatment and control groups. If users in the treatment and control group were random, we would expect them to be balanced on all external variables. We use Standard Mean Differences (SMD) in order to ensure that users in the treatment and control groups are balanced on external variables. SMD is defined as the difference in means of external variables of treated and control groups divided by the standard deviation within the treated group. This number should not be more than 0.25 for all external variables, showing that the users in both groups are balanced based on external variables and there is no selection threat in the experiment (Althoff, Jindal, and Leskovec 2017). We report the balancing of the external variables summarized in Table 1.

From Figure 2, we observe that the shift of interest, SII, of treatment and control groups increases before the interruption. However, the shift in interest of the treatment group decreases significantly for months 4 and 5 while the SII of control group increases for the same time intervals (15% decrease in SII for the treatment group; significantly larger than 0 according to Wilcoxon signed rank test; $P<0.001$). Thus, our results show that engaging in exercise by going to the gym or fitness center can help users become more focused in their online behavior and have a more focused set of interests across different time intervals.

Variable	SMD
First tweet date	0.090
First Foursquare check-in date	0.039
Average number of posts per month	0.006
Gender	0.034
Follower	0.016

Table 1. Standard Mean Differences (SMD) on external variables.

Conclusions and Future Work

In this paper, we report on an observational study that shows how exercising as a form of offline activity can influence users' online behavior. Based on the interpolation of data from Twitter and Foursquare, we have been able to identify those users who frequently go to a gym or fitness center and match them against a control group with matching external variables and similar pre-treatment measurements. In order to see whether offline activity has impact on the users' online behavior, we measure users' shift in interest across different time intervals. Lower shift in interest shows users that are more focused on a specific set of topics while a higher shift in interest points to users that frequently change their interests. Our study shows that the users who frequently exercise (at least once every ten days) have a more consistent focus towards topics across subsequent time intervals.

As a part of our ongoing work, we are extending this work in three directions: 1) We are exploring whether other types of offline activity such as frequent shopping, going to a bar, among others also have impact on the users' online behavior. 2) We are looking into formalizing other behavior traits or personal characteristics that can be extracted from online content. This would include measures such as interest diversity, opinion conformity, and convincibility, among others. The definition of these measures will allow us to investigate the impact of users' offline activity on such online behavior. 3) Last, in the extended work, we match users not only based on their online behavior but also based on external variables to reduce their effects. To do so, we can use propensity score matching. In this method, we find logistic regression for users based on their external variables and match users based on their propensity score.

References

Althoff, T.; Jindal, P.; and Leskovec, J. 2017. Online Actions with Offline Impact: How Online Social Networks Influence Online and Offline User Behavior. In *Proceedings of the Tenth international conference on web search and data mining (WSDM*), 537-546. ACM.

Anderson, A.; Huttenlocher, D.; Kleinberg, J.; and Leskovec, J. 2013. Steering User Behavior with Badges. In *Proceedings of the 22nd international conference on World Wide Web (WWW)*, 95-106. ACM.

Bagroy, S.; Kumaraguru, P.; and De Choudhury, M. 2017. A Social Media Based Index of Mental Well-Being in College Campuses. In *Proceedings of the 34th Annual Conference on Human Factors in Computing Systems (CHI)*, to appear. ACM.

De Choudhury, and M. and Kiciman, E. 2017. The Language of Social Support in Social Media and its Effect on Suicidal Ideation Risk. In *Proceedings of the 11th International AAAI Conference on Web and Social Media (ICWSM)*, to appear. AAAI.

De Choudhury, M.; Kiciman, E.; Dredze, M., Coppersmith, G.; and Kumar, M. 2016. Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media. In *Proceedings of the 2016 Conference on Human Factors in Computing Systems (CHI)*, 2098-2110. ACM.

De Choudhury, M.; Sharma, S.; and Kiciman, E. 2016. Characterizing Dietary Choices, Nutrition, and Language in Food Deserts via Social Media. In *Proceedings of the 19th Conference on Computer-Supported Cooperative Work & Social Computing*, 1157-1170). ACM.

De Choudhury, M.; Sharma, S.; Logar, T.; Eekhout, W.; and Nielsen, R.C. 2017. Gender and Cross-Cultural Differences in Social Media Disclosures of Mental Illness. In *Proceedings of the 20th Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)*. ACM.

Dos Reis, V.L., and Culotta, A. 2015. Using Matched Samples to Estimate the Effects of Exercise on Mental Health from Twitter. In *Proceedings of the Twenty-Ninth Conference on Artificial Intelligence*, 182-188. AAAI.

Elahi, M.; Braunhofer, M.; Ricci, F.; and Tkalcic, M. 2013. Personality-Based Active Learning for Collaborative Filtering Recommender Systems. In *Congress of the Italian Association for Artificial Intelligence.* 360-371. Springer.

Farnadi, G.; Sitaraman, G.; Sushmita, S.; Celli, F.; Kosinski, M.; Stillwell, D.; Davalos, S.; Moens, M.F.; and De Cock, M. 2016. Computational Personality Recognition in Social Media. *User Modeling and User-Adapted Interaction*, 109-142. Springer.

Gosling, S.D.; Rentfrow, P.J.; and Swann, W.B. 2003. A Very Brief Measure of the Big-Five Personality Domains. *Journal of Research in personality*. 504-528.

Han Veiga, M., and Eickhoff, C. 2016. A Cross-Platform Collection of Social Network Profiles. In *Proceedings of the 39th International conference on Research and Development in Information Retrieval (SIGIR)*, 665-668. ACM.

Hu, R., and Pu, P. 2011. Enhancing Collaborative Filtering Systems with Personality Information. In *Proceedings of the fifth conference on Recommender systems (RECSYS)*, 197-204. ACM.

Karumur, R.P.; Nguyen, T.T.; and Konstan, J.A. 2016. Exploring the Value of Personality in Predicting Rating Behaviors: A Study of Category Preferences on Movielens. In *Proceedings of the 10th Conference on Recommender Systems (RECSYS)*, 139-142. ACM.

Liu, Y.; Cao, X.; and Yu, Y. 2016. Are You Influenced by Others When Rating: Improve Rating Prediction by Conformity Modeling. In *Proceedings of the 10th Conference on Recommender Systems (RECSYS),* 269-272. ACM.

Oktay, H.; Taylor, B.J.; and Jensen, D.D. 2010. Causal Discovery in Social Media Using Quasi-Experimental Designs. *In Proceedings of the First Workshop on Social Media Analytics*, 1-9. ACM.

Olteanu, A.; Varol, O.; and Kiciman, E. 2017. Distilling the Outcomes of Personal Experiences: A Propensity-scored Analysis of Social Media. In *Proceedings of the 20th Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)*, 370-386. ACM.

Roshchina, A.; Cardiff, J.; and Rosso, P. 2011. User Profile Construction in the TWIN Personality-based Recommender System. Sentiment Analysis where AI meets Psychology. In *Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP)*. 73–79. ACL.

Stephanie, Y., and Sklar, M. 2016. Detecting Trending Venues Using Foursquare's Data. *RECSYS Posters*. ACM.

Zarrinkalam, F.; Fani, H.; Bagheri, E.; and Kahani, M., 2016. Inferring Implicit Topical Interests on Twitter. In *Proceedings of the 38th European Conference on IR Research (ECIR)*, 479–491. Springer.

Zarrinkalam, F.; Fani, H.; Bagheri, E; and Kahani, M. 2017. Predicting Users' Future Interests on Twitter. In *Proceedings of the 39th European Conference on IR Research (ECIR)*, 8–13. Springer.