Towards Obesity Surveillance Using Multifaceted Online Social Relational Factors in Reddit

Albert Park, Yaorong Ge

Software and Information Systems, University of North Carolina at Charlotte, Charlotte, North Carolina, United States

Objective

We aim to better understand online social interactions and environments of individuals interested in weight management from a social media platform called *Reddit*.

Introduction

Overweight and obesity are recognized as one of the greatest modern public health problems [1], yet worldwide prevalence of obesity has nearly doubled over the past 30 years [2]. As part of a strategy to control the obesity pandemic, the WHO recommends an obesity surveillance at the population level [3]. Empirical studies have shown the importance of social networks in obesity [4] and new strategies focusing on social interactions and environments have been proposed [5] to prevent the further increase in obesity prevalence. With the increasing use of the internet, online social networks, interactions, and environments (i.e., online social relational factors) deserve more attention.

Nearly three- quarters of Americans go online daily [6], for functions like connecting with individuals via social network sites [7]. Like face to face interactions, studies have suggested that social interactions and networks on the internet can influence behavior changes [8]. Previous studies examining social networking sites typically examine a few selected social networking sites (example studies [9,10]), although individuals could be members of multiple social networking sites. To better leverage online social relational factors for the purpose of characterizing and monitoring population obesity trends, we investigate weight management community members' other communities and their level of participation, a first step toward utilizing online multifactorial social interactions and environments.

Methods

In this study, we studied Reddit (*http://www.reddit.com*), a popular social interaction site, because Reddit hosts many subreddits (i.e., sub-communities), including weight management communities called r/loseit. First, we use a dataset [11] — made available on Reddit — that had been used in many informatics studies [12-14]. For this study, we used a portion of the dataset from Jan 2015 to May 2015. In the first five months of 2015, 5,006,186 members were active in 96,462 subreddits, while submitting 17,851,561 posts and 266,268,920 associated comments. Second, we identified members with more than 3 posts on r/loseit in that period and removed 'bot' accounts by manually examining the top 20 frequent posters and their account IDs. Third, we extracted these members' entire discussions made on Reddit, regardless of the subreddits. Fourth, we identified these members' overall activities on Reddit and visualized in a network [15].

Results

After removing bot accounts, we identified 7,734 members who had more than 3 posts in r/loseit from Jan 2015 to May 2015. On average, these members participated in 78.5 subreddits (standard error: 0.1; median: 49.0), while participating in 13,649 unique subreddits as a whole. Members' participated subreddits are summarized in Figure 1. The size of the nodes represents the number of participating members and the thickness of edges represents the number of members who participated in both subreddits.

Conclusions

We present preliminary findings towards better understanding the online multifactorial social interactions and environments on a social networking site called Reddit. We provide evidence that members encounter many social interactions that occur outside of the community of our interest, the weight management community. However, what members discuss outside of the weight management community as well as the interactions' influence on weight managements and changes remain unanswered. For example, many members also participate in a subreddit called r/fitness, a community that could share many similar interests with

r/loseit. However, the purpose for participating in both communities is unknown. On the basis of our initial analysis, we suggest leveraging online multifaceted social relational factors for the purpose of characterizing and monitoring population obesity trends.

References

1. Jeffery RW, Utter J. 2003. The changing environment and population obesity in the United States. *Obes Res*. 11(Suppl), 12S-22S. PubMed https://doi.org/10.1038/oby.2003.221

2. World Health Organization. Global Health Observatory (GHO) data: Obesity. (2009). Available at: http://www.who.int/gho/ncd/risk_factors/obesity_text/en/. Archived at: http://www.webcitation.org/6rQICh7Oq.

3. World Health Organization. 2000. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. *World Health Organ Tech Rep Ser*. 894, i-xii, 1-253. PubMed

4. Christakis NA, Fowler JH. 2007. The spread of obesity in a large social network over 32 years. *N Engl J Med*. 357, 370-79. PubMed https://doi.org/10.1056/NEJMsa066082

5. Leroux JS, Moore S, Dubé L. 2013. Beyond the 'I' in the obesity epidemic: a review of social relational and network interventions on obesity. *J Obes*. 2013, 348249. PubMed https://doi.org/10.1155/2013/348249

6. Perrin A. One fifth of Americans report going online 'almost constantly'. *Pew Research Center* (2015). Available at: http://www.pewresearch.org/fact-tank/2015/12/08/one-fifth-of-americans-report-going-online-almost-constantly/
Archived at. http://www.webcitation.org/6s9ZTXRDl.

7. Greenwood S, Perrin A, Duggan M. 2016. Social Media Update 2016. Pew Research Center Internet. 1-9. http://www.pewinternet.org/2016/11/11/social-media-update-2016/.
Archived at. http://www.webcitation.org/6q0FDWNRI.

8. Laranjo L, et al. 2015. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. *J Am Med Inform Assoc*. 22, 243-56. PubMed https://doi.org/10.1136/amiajnl-2014-002841

9. Park A, et al. 2016. 'How Did We Get Here?': Topic Drift in Online Health Discussions. *J Med Internet Res*. 18, e284. PubMed https://doi.org/10.2196/jmir.6297

10. Park A, Conway M, Chen AT. 2018. Examining Thematic Similarity, Difference, and Membership in Three Online Mental Health Communities from Reddit: A Text Mining and Visualization Approach. *Comput Human Behav*. 78, 98-112. PubMed https://doi.org/10.1016/j.chb.2017.09.001

11. Reddit_Member. I have every publicly available Reddit comment for research. ~ 1.7 billion comments @ 250 GB compressed. Any interest in this? (2015). Available at: https://www.reddit.com/r/datasets/comments/3bxlg7/i_have_every_publicly_available_reddit_comment/ Archived at. http://www.webcitation.org/6kgAuNxDE.

12. Park A, Conway M. Tracking Health Related Discussions on Reddit for Public Health Applications. *Annu. Symp. proceedings. AMIA Symp*. **2017,** 1362– 1371 (2017).

13. Park A, Conway M. 2018. Harnessing Reddit to Understand the Written-Communication Challenges Experienced by Individuals With Mental Health Disorders: Analysis of Texts From Mental Health Communities. *J Med Internet Res*. 20, e121. PubMed https://doi.org/10.2196/jmir.8219

14. Park A, Conway M. 2017. Towards Tracking Opium Related Discussions in Social Media. *Online J Public Health Inform*. 9(1), e073. https://doi.org/10.5210/ojphi.v9i1.7652

15. Bastian, M., Heymann, S. & Jacomy, M. Gephi: An Open Source Software for Exploring and Manipulating Networks. (2009).
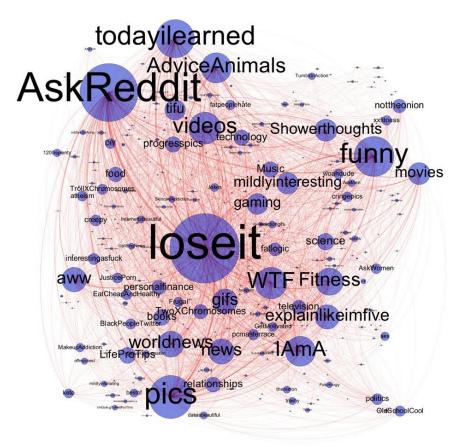

Figure 1. An overview of r/loseit members' participation in other subreddits.