Why Do You Write This? Prediction of Influencers from Word Use

Jalal Mahmud
IBM Research - Almaden
jumahmud@us.ibm.com

Abstract

With the widespread usage of social media, there has been much effort on detecting influential users for different information propagation applications. An inherent limitation of such methods is that they can only detect influential users after such users show observable signals of influence. However, in many real world applications including a counter-campaign, an organization needs a way to detect influencers early, so that they can take appropriate measure before it is too late to intervene. In this work, we present a method to detect such would-be influencers from their prior word usage in social media. We compute psycholinguistic category scores from word usage, and investigate how people with different scores exhibited different influence behaviors on Twitter. We also found psycholinguistic categories that show significant correlations with such behaviors, and built predictive models of influence from such category based features. Our experiments using a real world dataset validates that such predictions can be done with reasonable accuracy.

Introduction

Recent years have seen a rapid growth in micro-blogging and the rise of popular micro-blogging services such as Twitter. With the widespread use of such services, and the ever growing number of users, there has been much effort on identifying influential users (Agarwal et al. 2008, Bakshi et al. 2011, Huang et al. 2012, Weng et al. 2010). Such influential users can propagate information (e.g., spreading emergency alerts) and change opinions in campaigns ranging from politics and government to social issues. Existing methods of influencer identification are based on a social media users' social network information (Huang et al. 2012, Cha et al. 2010), content of posts (Agarwal et al. 2008), information forwarding/propagating activity (Romero et al. 2011), topic-specific activity (Weng et al. 2010), strength of reposting (Bakshi et al. 2011), etc.

An inherent limitation of these approaches is that due to their reliance on observable activity, content and social network based metrics of influencers, they can only identify influencers after the fact. As a result, they fail to detect would-be influencers since these individuals have not established strong observable signals of influence, such as high number of followers or high number of reposts (retweets). Meanwhile, in many real world applications including marketing and messaging campaigns ranging from politics to government and social issues, it is desirable to identify influencers early so that targeted messages can be sent to them for information propagation. In addition, many organizations including government would also want such ability to detect influencers early so that they can take appropriate measure before it is too late to intervene. For example, when anti-government messages are spread in social media, government would want to identify influencers of anti-government campaign early to intervene and minimize the spread of their campaign.

In this work, we describe a method to predict such would-be influencers in social media. Bakshi et al. described one such approach to predict influencers early from their past activity (Bakshi et al. 2010). Our research is complementary to their effort, and specifically focuses on whether prior word usage can predict influencers in social media.

Previous research reported correlations of people's word usage with personality, and predicting personality from word usage (Yarkoni et el. 2010, Gill et al. 2009, Golbeck et al. 2011, Mairesse et al. 2006). In addition, algorithms that use word-based features to predict other attributes, such as sentiment (Pang et al. 2008) and political polarization (Cohen et al. 2013) also exist. However, word use has not been studied in the context of influence in social media. We measured word use in a number of psycholinguistic categories as defined by the Linguistic Inquiry and Word Count (LIWC) dictionary (Pennebaker et al. 2001). Then, we correlated such psycholinguistic category scores with user's influence computed from social media. Based on our analysis with a real-world dataset collected from

Twitter, we have found statistically significant correlations of several psycholinguistic categories with influence. Furthermore, we have built predictive models from psycholinguistic category features for predicting influence score of a Twitter user. Our experiments validate that such predictions can be done with reasonable accuracy.

This paper is organized as below. We present a brief overview of influence measurement in previous research, and our approach in the next section. Then, we describe our dataset for this work. Next, we describe how we conduct psycholinguistic analysis with LIWC dictionary. The following sections describe the correlation of influence behavior with LIWC categories, and predictive model of influence from LIWC categories. Finally, we discuss the findings and conclude the paper.

Measuring Influence

Researchers have proposed various measures of influence. Bakshi et al. measured influence by a user's ability to seed content containing URLs that generate large cascades of reposts (retweets) (Bakshi et al. 2011). To identify consistently influential users, they aggregated all URL posts by user and computed individual-level influence as the logarithm of the average size of all cascades for which that user was a seed. Cha et al. proposed three measures of user influence in Twitter: the number of followers of a user, the number of retweets containing user's name and number of mentions containing user's name (Cha et al. 2010). Weng et al. measured user's influence using topic-specific page rank approach (Weng et al. 2010). Romero et al. designed an algorithm which considers pair-wise influence relationships among users and user's passivity (a measure of how difficult for other users to influence the user) to assign influence score of Twitter users (Romero et al. 2011). Motivated by Bakshi et al. and Cha et al., we measure *influence score* by the average number of retweets generated from a user's tweets.

Dataset

We used Twitter's streaming API from Nov 1, 2013 to Nov 14, 2013 to randomly sample 1000 users. We crawled their tweets from last one month (Oct 2013) and denoted them as *recent tweets*. We also crawled users' historical tweets (max 200) before last one month (until Oct 2013), and denoted them as *historical tweets*. There were 134857 historical tweets in total. We obtained total number of retweets generated by each tweet in users' recent tweets using Twitter's REST API. Next, we computed influence score of each user from their *recent tweets*, where influence score of a user is the average number of retweets generated by his recent tweets from Oct 1, 2013 until Nov 14, 2013. Average influence score was 0.213, and standard deviation was 0.098.

Psycholinguistic Analysis from text

Motivated by prior works on personality prediction from text (Golbeck et al. 2011, Yarkoni et al. 2010, Gill et al. 2009, Mairesse et al. 2006), we measured word uses in users' *historical tweets* with the Linguistic Inquiry and Word Count (LIWC) 2001 dictionary (Pennebaker et al. 2001). LIWC is the most commonly used language analysis tool for investigating the relation between word use and psychological variables (Pennebaker et al. 2001). LIWC 2001 defines over 70 different categories, each of which contains several dozens to hundreds of words (Pennebaker et al. 2001). We excluded the categories that are non-semantic (e.g., proportion of long words) or relevant to speech (e.g. fillers) and considered the remaining 66 LIWC categories. For each person, we computed his/her LIWC-based scores in each category as the ratio of the number of occurrences of words in that category in one's tweets and the total number of words in his/her tweets. We excluded retweets when computing LIWC-based category scores, because retweets are content generated by others.

Influence and Word Use

In this analysis, we correlate the LIWC category scores with influence score, and use the correlation significance to measure reliability. This analytical approach was used to identify associations between personality and word use (Golbeck et al. 2011, Yarkoni et al. 2010, Gill et al. 2009). Thus, we begin by running a two-tailed Pearson correlation analysis between users' LIWC category scores and users *influence score*. The statistically significant correlations are shown in Table 1.

LIWC categories	Correlation	Significance Level
Anger	-0.213	**
Anxiety	-0.114	*
Swear Words	-0.198	*
Communication	0.164	**
Cognition	0.103	*
Perception	0.152	*
Social Process	0.167	**
Friends	0.171	*
Humans	0.118	*
Positive Feelings	0.215	**
Positive Emotions	0.179	*
Negative emotion	-0.191	**
Discrepancy	0.221	*
Inclusive	0.172	*
Physical States	-0.153	*
Inhibition	-0.184	*
Insight	0.193	*
Certainty	0.167	*

Table 1. Statistically significant correlations (* means $p < 0.05$, ** means $p < 0.01$) of LIWC categories with influence

Independent variables used	Mean Absolute Error (MAE)
All LIWC categories	0.269
LIWC categories with significant correlations	0.223

Table 2. Regression results over a 10-fold cross validation

Features Used	Recall	Precision	F1
All LIWC categories	0.82	0.79	0.804
LIWC categories with significant correlations	0.88	0.91	0.895

Table 3. Binary Classification Results over 10-fold cross validation

Features Used	Recall	Precision	F1
All LIWC categories	0.721	0.744	0.732
LIWC categories with significant correlations	0.775	0.813	0.793

Table 4. Classification Results over 10-fold cross validation when 10-influence categories are considered

LIWC category anger is negatively correlated with influence score which indicates that people who use more angry words (such as *anger*, *angry*) are less likely to be influencer. An anxious person, exhibited by the LIWC category anxiety and consisting of words such as *afraid* and *alarm*, is less likely to be an influential. Similarly, people who use swear words (such as damn, piss) are also less likely to be influencer. These are quite intuitive.

The LIWC category communication is significantly positively correlated with influence, which indicates that more communicative people are more likely to be influencer. The LIWC category cognition, exemplified by words such as *accept*, *acknowledge*, *admit*, and *agree*, has a significant positive correlation with influence behavior. The LIWC category perception represents words such as *ask*, *call*, and *contact*. These words are often used by people who are more interactive, and the positive correlation indicates that interactive people may be more influential. A social (exemplified by words such as *interact, involve*) and friendly (example words are friends, partner, buddy) person are likely to be influencer. Similarly, LIWC category humans (example words are people, guy, man, girl) is also positively correlated with influence. An individual with positive feelings (exemplified by words such as *care*, *cheer, attachment*) and positive emotions (exemplified by words such as *accept* and *admit)* are more likely to be influencer. These are quite intuitive. On the other hand, people who use negative emotion words (such as hurt, ugly, nasty) are less likely to be an influential.

The LIWC category discrepancy has significant positive correlations with influence behavior. This represents words such as expect, hope, need, must, should which may indicate determination and desires for the future. In addition, we found that a person who scores high in the LIWC category inclusive, containing words such as along and also, is also more likely to be influencer. This category is intuitively related with people who are more social and friendly, and it follows that people who scores high also seem to be influencer. The LIWC category physical states has significant negative correlations with influence behavior. This category represents words such as diabetes, disease, dizziness, and sleep. These words often indicate someone's sickness or inactivity. This may indicate apathy on the part of the user, which would make them less likely to be influencer. LIWC category inhibition (represents words such as shy, unease, retard) also negatively correlated with influence behavior.

The LIWC category insight is represented by words such as think, know, consider. We found a significant positive correlation of this category with influence behavior. Similarly, LIWC category certainty (exemplified by words such as fact, confidence, always) is positively correlated with influence.

Prediction Models

We attempted to build predictive models for influence score based on the psycholinguistic category features to understand their predictive power to predict influence score. We performed both regression analysis and a classification study using WEKA (Hall et al. 2009), a widely used machine learning toolkit.

For regression analysis, we formulated linear regressions to predict influence score using LIWC measures. Thus, LIWC attributes were independent variables of the regression model and users' influence score was the dependent variable. We tried a number of regression approaches, including simple linear regression, multiple linear regression, pace regression, logistic regression and SVM regression. We performed 10-fold cross validation. SVM regression slightly outperformed other algorithms.

Table 2 presents the result of the regression analysis in terms of mean absolute error (MAE). We find that influence score can be predicted within 22.3%-26.9% MAE. The best result is obtained when LIWC categories with significant correlations were used as independent variables in the regression model.

In the classification study, we tried two different settings. In the first setting, we tried supervised binary machine learning algorithms to classify users with above-median levels of *influence score*. In the second setting, we divide influence scores into 10 equal sized bins, and trained supervised classifiers with 10 classes.

We experimented with a number of classifiers from WEKA, including naive Bayes, SMO (an SVM implementation), J48 (a decision-tree-based classifier), Random Forest (an ensemble method that combines multiple decision trees). Table 3 and 4 show the classification result of the best WEKA classifier in terms of Recall, Precision and F1 under 10-fold cross validation. We see that classifying high (above median) or low (below median) influencers from LIWC category based features can be done quite accurately (89% F1 when significant LIWC categories are used as features). Accuracy reduces when we considered 10 influencer categories. Prediction difficulty increases with more class and hence prediction accuracy drops. However, even with 10 influence categories, our model performs reasonably well (79% F1 with significant LIWC features). For each setting, best classifier is obtained when LIWC categories with significant correlations are used as features.

Discussion

Our findings suggest several implications for social media research. Our work is the first study that reports correlations of word usage with influence behavior. Our work also discovers a set of psycholinguistic categories which have significant correlations with users' influence in Twitter. Such psycholinguistic categories are also shown to be useful to predict the influence score of a user. Our findings also suggest that certain word use can predict such likelihoods. For instance, people who use angry words (such as *anger, hate*, *kill*) are less likely to be influencer, people who write about communication are more likely to be influencer, and people who write about positive feelings are more likely to be influencer. Predictive models of influencer from prior word usage can identify users early to be an influencer. This is useful for various information propagation applications including viral marketing (Brown et al. 1987), and social/government/political campaigns (e.g., a campaign to support "vaccination"). Identifying such influencers early can be also useful to counter their affect. Our research to identify influencers from prior word use is orthogonal to existing research on influence modeling (Agarwal et al. 2008, Bakshy et al. 2011, Cha et al. 2010) for such applications. Our regression analysis confirms that such predictions can be done with reasonably low error (22.3% MAE). Furthermore, our classification study shows that classifying influencers can be done quite accurately (89% binary classification accuracy, and 79% classification accuracy with 10-classes). In this study we did not explore ranking algorithms. However, learning-to-rank algorithms (Liu et al. 2009) may be useful to further harness the predictive information from word use.

Conclusion

We have presented a study on understanding how people with different word usage exhibit different influence behaviors on Twitter. We conducted psycholinguistic analysis from word usage and found certain psycholinguistic categories have significant correlations with such influence behavior. We built predictive models of influencers from the psycholinguistic category features and demonstrated that our models are reasonably accurate. We have identified several directions of future work. First, we will extend our analytics to other social media platforms. Second, we will verify the generality of the findings with other social media datasets (e.g., campaign specific dataset). Third, we will investigate how word use with other activity based features can predict influence. Fourth, we will investigate whether topic modeling such as LDA can be useful to compute features from word use in the prediction model. Fifth, we will conduct a similar analysis for other measures of influence (e.g., number of followers gained in a specific time). Finally, we will integrate our model with a campaign application, and investigate its usage in the real world.

References

Agarwal, N., Liu, H., Tang, L., and Yu, P. S. Identifying the influential bloggers in a community. In *Proc. WSDM*, 2008.

Bakshy, E., Hofman, J. M., Mason, W. A., and Watts, D. J. Everyone's an influencer: quantifying influence on twitter, In *Proc. WSDM*, 2011.

Brown, J., Reinegan, P. Social ties and word-of-mouth referral behavior. Journal of Consumer Research 14:3 (1987), 350-362.

Cha, M., Haddadi, H., Benevenuto, F., and Gummadi, K.P. Measuring user influence in twitter: The million follower fal-lacy. In *Proc. ICWSM*, 2010.

Cohen, R., and Ruths, D. 2013. Classifying Political Orientation on Twitter: It's Not Easy!. In ICWSM'13.

Gill, A.J., Nowson, S., Oberlander, J. 2009. What Are They Blogging About? Personality, Topic and Motivation in Blogs. In ICWSM'09.

Golbeck, J. Robles, C., Edmondson, M., Turner, K. 2011 Predicting Personality from Twitter. In *Proc. IEEE SocialCom. 2011*.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. The WEKA data mining software: an update. In ACM SIGKDD newsletter, Vol 11, Issue 1, 2009.

Huang, J., Cheng, X.-Q., Shen, H.-W, Zhou, T., and Jin, X. Ex ploring social influence via posterior effect of word-of-mouth recommendations. In *Proc*. WSDM, 2012.

Liu T. Y. 2009. Learning to rank for information retrieval. Foundations and Trends in Information Retrieval, 3(3), 225-331.

Mairesse, F., Walker, M. 2006. Words Mark the Nerds: Computational Models of Personality Recognition through language, I*n Proc. of CogSci 2006*.

Pang, B. and Lee, L. 2008. Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval (2).

Pennebaker, J.W., Francis, M.E., and Booth, R.J. Linguistic Inquiry and Word Count. *Erlbaum Publishers*, 2001.

Romero. D. M., Galuba. W., Asur. S, and Huberman, B. A. Influence and passivity in social media. In *Proc. ECML/PKDD*, 2011.

Weng, J. Lim. E.-P., Jiang. J, and He. Q. Twitterrank: Finding topic-sensitive influential twitterers. In *Proc. WSDM*, 2010.

Yarkoni, Tal. Personality in 100,000 words: A large-scale analysis of personality and word usage among bloggers. Journal of Research in Personality.