

On Unravelling Opinions of Issue Specific-Silent Users in Social Media

Wei Gong, Ee-Peng Lim, Feida Zhu, Pei Hua Cher

School of Information Systems
Singapore Management University
{wei.gong.2011, eplim, fdzhu, phcher}@smu.edu.sg

Abstract

Social media has become a popular platform for people to share opinions. Among the social media mining research projects that study user opinions and issues, most focus on analyzing posted and shared content. They could run into the danger of non-representative findings as the opinions of users who do not post content are overlooked, which often happens in today's marketing, recommendation, and social sensing research. For a more complete and representative profiling of user opinions on various topical issues, we need to investigate the opinions of the users even when they stay silent on these issues. We call these users the *issue specific-silent* users (*i*-silent users). To study them and their opinions, we conduct an opinion survey on a set of users for two popular social media platforms, Twitter and Facebook. We further analyze their contributed personal social media data. Our main findings are that more than half of our users who are interested in issue *i* are *i*-silent users in Twitter. The same has been observed for our Facebook users. *i*-silent users are likely to have different opinion distribution from the users who post about *i*. With the ground truth user opinions from the survey, we further develop and apply opinion prediction methods to *i*-silent users in Twitter and Facebook using their social media data and their opinions on issues other than *i*.

Introduction

Motivation Nowadays, millions of users share content in social media. This abundant user-generated content provides an unprecedented resource for user opinion analysis. Opinions of users are useful in many real world applications (Maynard, Bontcheva, and Rout 2012). Retailers are keen to know how well consumer think of new products and in what product aspects. Political parties and analysts want to predict election outcome based on public opinions. Universities also rely on public ratings on their academic and research programs to secure good ranking. User opinion insights are important to organizations and governments. They allow decision makers to fine tune customer relationship policies and government policies, and to help individuals' decision-making process (e.g., which products to buy, which movies to watch or which politicians to vote).

While many social media users share their opinions online, many more others prefer to stay silent. A user may

choose to keep silent on an issue even when she is interested in it, or when she has opinions on it. This can be caused by different reasons. Previous studies suggest that social media users generate content selectively (Hampton et al. 2014; Das and Kramer 2013; Sleeper et al. 2013). As user-generated content is often visible to others, users may practise self-censorship when deciding what content to share (Das and Kramer 2013; Sleeper et al. 2013). For example, a user may not share her opinion online because she does not want to start an argument with others, she thinks the opinion is not appropriate to share in public, or she is afraid that many of her friends have different opinions (Preece, Nonnecke, and Andrews 2004; Hampton et al. 2014; Sleeper et al. 2013).

User-generated content therefore include opinions on an issue from only those who post about the issue. When we conduct opinion analysis on these content, we will likely derive a biased conclusion of what the public think about the issue. The main question here is then how can we obtain opinions on topical issues from a set of users who are interested in the issues but do not share their opinions in social media. We call these users the *issue specific-silent* users or *i*-silent users. For example, if a user is interested in issue "Healthcare Cost" but never posts about it, she is then considered a Healthcare Cost-silent user. We call the users who post about an issue the *issue specific-active* users or *i*-active users. It is important to note that *i*-silent users may still generate content unrelated to issue *i*. Hence, they may not be *overall silent users* who do not post anything or post only a little in a long period time (Tagarelli and Interdonato 2013; Gong, Lim, and Zhu 2015). On the other hand, an overall silent user is one who is *i*-silent for all issues.

Research Objectives In this work, we study the opinions of *i*-silent users in social media with two research goals. The first goal is to examine to what extent *i*-silent users exist for different issues and whether their opinion distribution is similar or different from that of *i*-active users. Achieving this goal is non trivial as ground truth opinions on issues are not in the observed social media data.

To obtain users' ground truth opinions, we conduct a user survey on Singapore social media users on Twitter and Facebook. In this survey, participants share their interests and opinions on seven Singapore related issues, and declare whether they discuss the issues in Twitter and Facebook. The

issues include *Healthcare Cost, Retirement, Public Housing, Public Transport, Jobs, Education, and Population Growth*. They are long-standing topical issues which are well aware of by people in Singapore. The users are thus expected to have opinions on them. Short term issues (e.g., events, news) are not included as they normally do not attract long lasting public interests. Opinions on these short term issues are likely to be confined to only very small number of users.

We have derived a number of interesting findings from our survey results. We found that in both Twitter and Facebook, more than half of the users who are interested in issue i are i -silent users across all issues. i -silent users are more likely to be neutral than i -active users and i -active users are more likely to feel positive than i -silent users. These findings suggest the number of i -silent users can be large and they are likely to have opinion distribution different from that of i -active users. It is therefore necessary to consider i -silent users when profiling opinions of a user population.

The second goal of this work is to predict the opinions of i -silent users in Twitter and Facebook. Addressing this goal enables us to profile i -silent users even when they have no posted content about the issue. This opens up new opportunities to engage the i -silent users in various applications including product recommendation, personalized content filtering, and social media marketing. We propose two types of features for the prediction: (a) *sentiment features* extracted from users' content, and (b) *opinion features* extracted from users' predicted opinions or ground truth opinions on other issues. We demonstrate the effectiveness of our features and show that predicting i -silent users' opinions can achieve reasonably good accuracy from user posted content that is *not* related to issue i , and achieve better accuracy when we make use of user opinions on other issues.

Related Work

In this section, we review the related works on opinion mining techniques, opinion analysis in social media, and opinions of i -silent users.

Opinion Mining on Social Media Data

Opinion mining is a classical text mining task to identify and extract "what people think" from textual content such as customer feedback emails, discussion forums, reviews and other social media postings (Pang and Lee 2008; Liu 2012). Understanding what people think using opinion mining is useful in product recommendation, product design, customer relationship management, and political sensing (Pang and Lee 2008). For example, users may buy products after reading opinions in product reviews. Companies improve product design and service delivery based on opinions in customers' feedback. Opinion mining has been intensively studied by the computational linguistics research community. The main focus is to determine whether a phrase, a sentence or a document is positive or negative, or to determine a user's view on certain issue, event or product (Pang, Lee, and Vaithyanathan 2002; Dave, Lawrence, and Pennock 2003; Popescu and Etzioni 2005; Wilson, Wiebe, and Hoffmann 2005; Socher et al. 2013). In this project, we adopt *Recurrent Neural Tensor Network* (Socher et al. 2013) trained with

the 5000+ issue related labeled tweets so as to derive the sentiment polarity (i.e., positive, neutral and negative) of short text on the selected topical issues.

Social media such as Twitter and Facebook has been a popular conduit for opinion mining and data science research (Pak and Paroubek 2010; Tumasjan et al. 2010; Chung and Mustafaraj 2011; Skoric et al. 2012; Barbosa and Feng 2010; Kouloumpis, Wilson, and Moore 2011). Opinion mining on social media data has been used to predict election results such as German Federal Election in 2009 (Tumasjan et al. 2010), US Senate special Election in Massachusetts 2010 (Chung and Mustafaraj 2011), Dutch Senate Election in 2011 (Sang and Bos 2012), Irish General Election in 2011 (Birmingham and Smeaton 2011) and French Presidential and Legislative Elections in 2012 (Ceron et al. 2014). Other examples include the prediction of stock market by analyzing public mood and emotions in Twitter (Bollen, Mao, and Zeng 2011), movie box office prediction using the number and sentiments of movie related tweets (Asur and Huberman 2010), and modeling of opinion shift over time (Lin et al. 2013).

Opinions of i -Silent Users

All the aforementioned studies have shown that social media content can be used to effectively determine users' opinions. However, social media content is generated when users choose to self-report their thoughts (Kiciman 2012; Guerra, Meira, and Cardie 2014). Thus the opinions of issue-specific silent users are not taken into account. As a result, one may obtain a biased opinion profile of the entire user community (Lin et al. 2013; Gayo-Avello 2012). For example, Lin et al. (2013) suggested that because of the self-reporting nature of social media, social media is a relatively poor tool for make population inferences. Gayo-Avello (2012) also pointed out that failure to consider silent users has contributed to poor election prediction accuracy.

There are very little work focusing on i -silent users' opinions. Two big research questions linger around these users, namely: (a) Do the silent users share the same opinions as the active users? and (b) How can one predict the opinion of silent users? Mustafaraj et al. (2011) compared the content generated by Twitter users who post very often and other users who post only once during the US Senate special Election in Massachusetts 2010. They found significant difference between the two groups of users' content. The result suggests that users who post none or only little content may hold opinions very different from very active users. It also suggests the importance of inferring i -silent users' opinions. As i -silent users do not post any content on the issue, inferring their opinions is challenging. To the best of our knowledge, our work is the first addressing this problem.

i -Silent Users in Social Media

To study i -silent users in social media and to obtain their ground truth opinions, we conduct a social media user survey. In this section, we describe the survey procedure and present our findings.

Survey Procedure

The social media survey serves two purposes. It collects the ground truth opinions of users on topical issues. It also allows us to gather complete social media content of each user for opinion prediction. Since Twitter and Facebook are the two popular social media platforms, we focus on their users so as to allow us to compare the findings obtained from their users. We also confine the users to be from Singapore who are expected to be familiar with the same set of topical issues.

Our survey requires each Twitter participant to have created her Twitter account at least three months ago and have at least 10 followees and 5 followers. Similarly, each Facebook participant is required to have created her Facebook account at least three months ago and have at least 20 friends. This ensures that the survey will not involve inexperienced users. We recruited the participants from undergraduate students of three largest universities in Singapore by email and poster. The participants are also incentivized to invite friends to join the survey. Each participant received at least 10 Singapore dollars for completing the survey, inviting friends and sharing their social media data. Both the survey itself and the survey methodology were approved by the Institutional Review Board (IRB) of the authors' university.

The survey has two parts. The first part establishes some basic information and ground truth opinions about the users. The survey requires information about the user's gender and age. Each user also answers multiple choice questions for each of the seven issues (Healthcare Cost, Retirement, Public Housing, Public Transport, Jobs, Education, and Population Growth). They are: (1) Is the user interested in the issue? (i.e., does she have opinion on the issue?) (2) What is the user's opinion on the issue ([0-3]negative/[4-6]neutral/[7-10]positive)? (3) Does the user discuss this issue in Twitter if she is a Twitter user, or in Facebook if she is a Facebook user? And (4) What is her social media friends' opinions on the issue according to her perception?

The second part of the survey collects a complete set of social media data from the participants which includes both content and social connections. The social connections are follower and followee links for Twitter users, and friend links for Facebook users. We asked the Twitter users to provide their Twitter screen names so as to crawl their Twitter data including tweets, social connections and their public followers and followees' tweets using Twitter API. As we also allow protected Twitter users to participate in our survey, for these protected accounts, we created a special Twitter account to follow them for a short time period so as to crawl their data.

To obtain Facebook users' data including friends and posts (i.e., statuses), we directly ask participants to provide us their Facebook data archives. Each Facebook archive includes almost all information in the user's account and we clearly stated this in the survey's informed consent form. Unfortunately, these archives exclude the friends' posts.

The survey was conducted from Sep 14, 2015 to Nov 12, 2015. We finally had 108 Twitter users and 74 Facebook users participated in the survey. Twitter users comprise 75 females and 33 males with an average age of 21.0. Face-

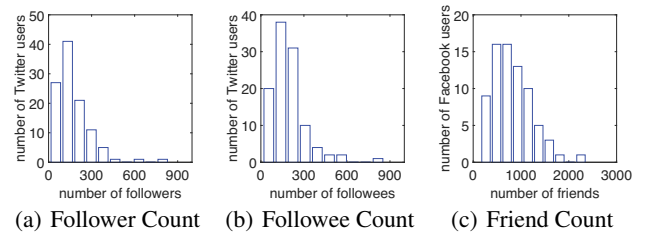


Figure 1: Twitter participants' follower count and followee count distribution and Facebook participants' friend count distribution.

book users comprise 48 females and 26 males with an average age of 21.3. Both users groups share very similar gender and age distributions. Figures 1(a), 1(b) and 1(c) show the Twitter users' follower count, followee count and the Facebook users' friend count distributions respectively. Our survey participants do not have very large number of followers or friends, thus they are "ordinary" users (not celebrities) whom we want to focus on in this research.

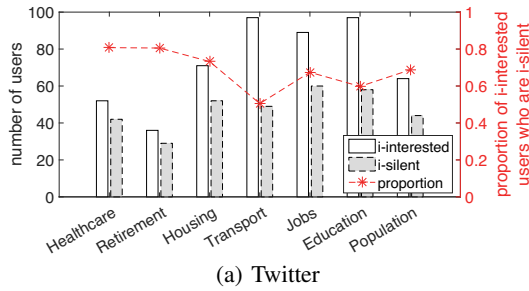
Survey Results and Findings

We analyze the survey results to answer the following questions:

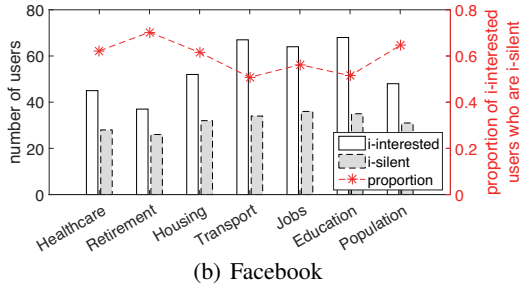
1. To what extent do *i*-silent users exist in social media? Are females or males more likely to be *i*-silent users?
2. Do *i*-silent users have opinions different from *i*-active users?
3. Do *i*-silent users believe that they have the same or opposite opinions with their friends? And how is it compared with *i*-active users' and their friends' opinions? Homophily is often observed among connected users. When a user's friends hold opinions (or perceived opinions) different from the user, it may prevent the user from expressing her opinion. We want to see if the effect exists in our survey and can explain the silent behavior.

Existence of *i*-Silent Users Firstly, we examine to what extent *i*-silent users exist in social media for different issue *i*. Based on the survey results, *i*-silent users in Twitter are the users who declare their interest in issue *i*, but never post issue *i* content in Twitter. Similarly, *i*-silent users in Facebook are defined similarly. Figures 2(a) and 2(b) show, for each issue *i*, the number of *i*-silent users, the number of users who are interested in *i* (i.e., *i*-interested users), and the proportion of *i*-interested users who are silent on *i* in Twitter and Facebook respectively.

We observe that a significant proportion of *i*-interested users are *i*-silent users across all issues in both Twitter and Facebook. The proportion of *i*-interested users who are silent is above 0.5 for all issues. It suggests that many people do not speak up even when they are interested in an issue. We also observe that different issues attract different amount of people's interest. For example, many more participants are interested in Public Transport, Jobs and Education than Healthcare Cost and Retirement. This may be due to the



(a) Twitter



(b) Facebook

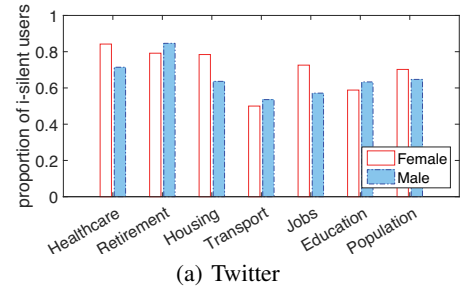
Figure 2: The proportion of i -interested users who are silent on i .

young participants (with average age less than 22) who may not worry about healthcare and retirement. We may expect a different distribution for more senior people.

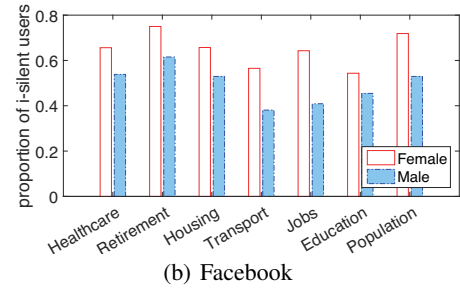
Gender difference among i -silent users. To answer whether females or males are more likely to be i -silent users, we compare the proportion of interested females who are i -silent users and likewise for the male users. Figure 3 shows that in Facebook, females are more likely to be silent on all issues than males (see Figure 3(b)). This result is consistent with findings in (Wang, Burke, and Kraut 2013) which show that female users in Facebook share more personal topics (e.g., family and personal health) while male users share more public topics (e.g., politics, sports, etc.). On the other hand, females in Twitter are more likely to be silent than males on healthcare, housing, jobs and population issues. For other three issues, the females in Twitter are only marginally less silent than males.

i -Silent Users' and i -Active Users' Opinions Next, we compare i -silent users and i -active users' opinions. i -active users are the users who are interested in issue i and post content about it. To ensure the significance of our results, we consider only the issues that have at least 20 i -silent users and 20 i -active users. Figures 4(a) and 4(b) show the proportion of i -silent users feeling negative, neutral and positive about issue i compared with the proportion of i -active users feeling negative, neutral and positive about i in Twitter and Facebook respectively. In each figure, the i -silent and i -active users are denoted by 'S' and 'A' respectively.

We observe that firstly, the proportion of i -silent users being positive is less than the proportion of i -active users being positive across all issues in both Twitter and Facebook (see the green bars on the right). For example, in Twitter,



(a) Twitter



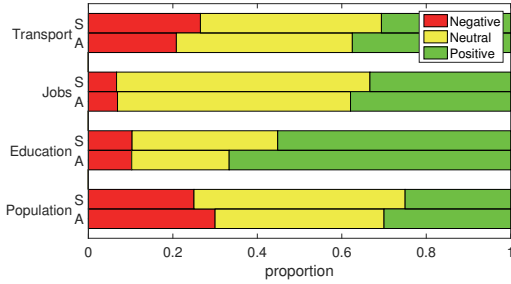
(b) Facebook

Figure 3: The proportion of interested females (males) who are i -silent users.

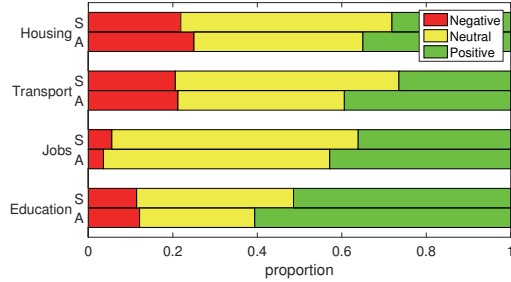
30.6% of Public Transport-silent users are positive, and a larger proportion (37.5%) of Public Transport-active users are positive. It implies that i -active users are more likely to be positive. Secondly, the proportion of i -silent users being neutral is greater than the proportion of i -active users being neutral across all issues in both Twitter and Facebook (see the yellow bars in the middle), which shows that i -silent users are more likely to be neutral. It suggests that users who actively post about an issue are likely to have some positive or negative opinion on it. Thirdly, the difference between the proportion of i -silent users who are negative and the proportion of i -active users who are negative is not consistent across the issues and platforms. The above findings show that i -silent users are likely to have different opinion distribution from i -active users. It is therefore important to predict i -silent users' opinions separately from that of i -active users.

i -Silent Users' and Social Media Friends' Opinions Finally, we examine if i -silent users believe that they have the same or opposite opinions with their social media friends, and how it is compared with i -active users and their friends' opinions. We note that social media friends are not all real friends of a user. Nevertheless, in the context of social media content sharing, it is reasonable to assume social media friends as an important social factor that affects content sharing decision, i.e., silent or active. In this analysis, the friends of a Twitter user refer to her followees from whom the user receives content.

For each issue i , we compute the proportions of i -silent users who believe having the same, moderate different and opposite opinions with their friends respectively. Suppose a user u 's opinion on an issue is O_u , and she perceives that her friends' opinion is O_f , then u believes that she has the same opinion with her friends if O_u and O_f are both nega-



(a) Twitter Participants



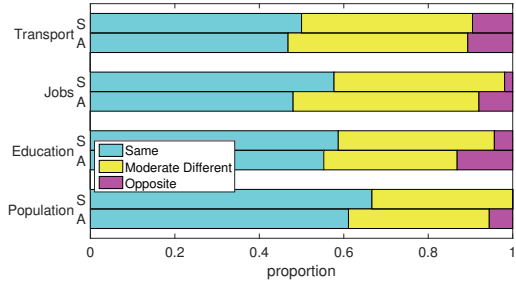
(b) Facebook Participants

Figure 4: Comparison of *i*-silent users and *i*-active users’ opinions. (S represents *i*-silent users and A represents *i*-active users.)

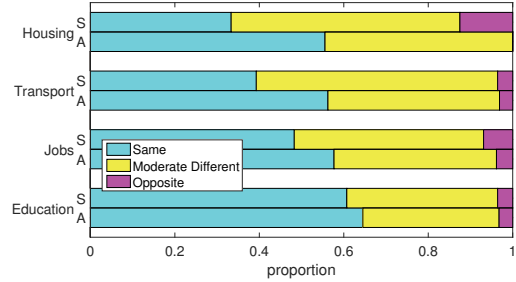
tive, neutral or positive, has moderate different opinion with her friends if one of O_u and O_f is neutral, and has opposite opinion with her friends if one of O_u and O_f is positive and the other is negative. We also compute the similar proportions for *i*-active users. Again, to ensure the significance of our results, we consider only the issues that have at least 20 *i*-silent users and 20 *i*-active users. Figures 5(a) and 5(b) depict the results among Twitter and Facebook participants respectively.

Firstly, we observe that both *i*-silent and *i*-active users believe some moderate difference existing between them and their online friends (see the yellow bars in the middle in Figure 5), but they are not likely to have opposite opinions with their friends (see the magenta bars on the right). The probability of users having opposite opinions with their social media friends is less than 0.13 for all issues in Twitter and Facebook. Thus, no matter users are silent or active on an issue, they perceive that the opinion differences with their social media friends are usually small.

Secondly, Figure 5(b) shows that compared with *i*-silent Facebook users, larger proportion of *i*-active Facebook users believe their having the same opinion with their online friends. For example, among Facebook users, 56.3% of Public Transport-active users believe their having the same opinion with their online friends, and the proportion is 39.3% for Public Transport-silent users. This phenomenon could be explained by that users are more likely to speak up when they believe their friends have similar opinions with them (Hampton et al. 2014). However, we have different observation from Twitter users. Compared with *i*-silent Twitter



(a) Twitter Participants



(b) Facebook Participants

Figure 5: *i*-silent users and *i*-active users’ opinions with their friends’ opinions. (S represents *i*-silent users and A represents *i*-active users.)

users, smaller proportion of *i*-active Twitter users perceive having the same opinion with their online friends. For example, among Twitter users, 46.8% of Public Transport-active users believe their having the same opinion with their online friends, and the proportion is 50.0% for Public Transport-silent users. The findings suggest that Facebook users are less interested to speak up when they have different opinions from their online friends, whereas Twitter users are more interested to speak up when they observe different opinions with their friends.

Why do *i*-silent users behave differently in the two platforms? A possible explanation is that although in both Twitter and Facebook, users can form connections and then get information from others, Facebook is used more as a private account for maintaining social connections with real life friends and family members (Ellison, Steinfeld, and Lampe 2007). People may not want to have arguments with their real life friends and family members online (i.e., in Facebook). On the other hand, Twitter is used more as an information channel where people connect with one another to get information that interests them (Kwak et al. 2010). Twitter users therefore have less personal connections with their friends, and thus more likely to express their differing opinions than Facebook users. Another possible explanation is that in general, our Facebook users have much more social connections than our Twitter participants (see Figure 1). Facebook users may want to be more “discreet” in sharing opinions with these friends.

To summarize, our survey results show that *i*-silent users exist across all the seven issues in Twitter and Facebook, and

female Facebook users are more likely to be silent on these issues. We also show that in both Twitter and Facebook, i -silent users are more likely to be neutral than i -active users and i -active users are more likely to be positive than i -silent users, and both i -silent and i -active users think they do not have much opinion conflicts with their social media friends.

Opinion Prediction

In this section, we predict opinions on the seven issues for i -silent users as well as i -active users using their contributed social media posts. To predict users' opinion on an issue i , we need to find the posts (i.e., tweets in Twitter and statuses in Facebook) that are related to i , i.e., i -related posts. For example, the post "Train is so crowded :(" is a related to Public Transport issue. For each i -related post, we further need to determine the opinion polarity about the issue i . In other words, we need to obtain the post's sentiment, i.e., positive, neutral, or negative. In the following, we first describe how to extract i -related posts. We then classify the sentiments of posts. Next, we derive the features for opinion prediction modeled as a classification problem. Finally, we present the opinion prediction results.

Issue Related Posts Extraction

To extract i -related posts from a large pool of social media posts, a straightforward way is to manually label a number of i -related posts and i -unrelated posts, and train a classifier to find all i -related posts. However, as i -related posts are likely to only constitute a very small proportion, directly labeling posts will incur too much manual effort before we can assemble a reasonably sized i -related posts. For this reason, we focus on identifying highly issue specific-keywords (i.e., i -keywords) to distinguish i -related posts from other posts.

We obtained these keywords from a set of issue related news articles. In The Straits Times (www.straitstimes.com, the most widely circulated Singapore newspaper), news are categorized into local topical issues including five of our seven selected issues. They are Education, Public Housing, Public Transport, Healthcare Cost and Jobs¹. We then crawled articles under these issues. By searching on The Straits Times website, we found the sets of news articles about the remaining two issues, Retirement and Population Growth². All the collected articles have URL with prefix www.straitstimes.com/singapore to ensure that they are Singapore based. In this way, we collect at most 200 articles for each issue. We call the articles about issue i the i -articles.

From the i -articles, we extract discriminative phrases as keywords for issue i . To compute the discriminative power of a phrase p (unigram or bigram) for issue i denoted by $d_i^{(p)}$, we first define the relative frequency of p in i -articles, i.e., $f_i^{(p)} = \frac{\text{number of } i\text{-articles containing } p}{\text{number of } i\text{-articles}}$. We then define p 's relative frequency in all articles, i.e., $f^{(p)} = \frac{\text{number of articles containing } p}{\text{number of articles}}$.

¹www.straitstimes.com/singapore/education (housing, transport, health, manpower)

²www.straitstimes.com/search?searchkey=retirement (population+growth)



Figure 6: Emoji examples in Twitter

The phrase p is discriminative in issue i if its relative frequency in i -articles is significantly larger than its relative frequency in all articles. Thus, we define the *discriminative power* of p by the difference of p 's relative frequency in i -articles and all articles, i.e., $d_i^{(p)} = f_i^{(p)} - f^{(p)}$. We subsequently rank the phrases according to their $d_i^{(p)}$ in descending order, choose the top 30 phrases, and manually remove some duplicated phrases (for example, we remove 'a school' as we already have 'school' as a keyword for Education). Table 1 shows i -keyword examples for the seven issues.

We then obtain the candidate i -related posts by selecting those containing any of the i -keywords. As our i -keywords are English words, the candidate i -related posts are likely written in English. We therefore do not perform further filtering to remove non-English posts. We call the set of candidate i -related posts S_i . However, not all posts in S_i are related to issue i . For example, post 'Let's train harder!' is not a Public Transport-related post although it contains keyword 'train'. Therefore, to further filter out the unrelated posts, we manually labeled 1000 randomly selected posts in S_i with 'related' and 'unrelated' labels, and then we use these labeled posts to build a Naive Bayes classifier and classify all the posts in S_i so as to get the final set of i -related posts. We can achieve at least 0.82 F-score for i -related posts across all the seven issues.

Sentiment Classification for Posts

To understand the sentiment of a post, we adopt the state-of-the-art Stanford sentiment analysis system proposed by Socher et al. (2013). This system uses a deep learning model, *Recursive Neural Tensor Network (RNTN)*, trained on Stanford sentiment treebank. The Stanford sentiment treebank is a dataset with 11,844 sentences from movie reviews and each sentence is a fully labeled parse tree. This dataset of trees (i.e., treebank) can be used to analyze the compositional effects of sentiment.

Although this Stanford sentiment analysis system achieves good results on movie reviews, it cannot be directly used on our problem. The first reason is that the system is trained using labeled movie reviews which are written in more formal way than posts in social media. Furthermore, the posts we have are posted by Singapore users, who use some regional slangs that do not appear in the Stanford sentiment treebank. For example, word 'sian' is used to express how bored and frustrated a person feels. Another reason is that the Stanford sentiment treebank does not include emojis (see Figure 6) and many emoticons (e.g., -.-, :D, :P, ^^), which are frequently found in posts and are useful for predicting sentiment (Agarwal et al. 2011). Emojis are represented using unicode (The Unicode Consortium 2015). For example, `\U0001F60A` is the unicode representation of a smile face.

Issue	Keywords
Healthcare Cost	health, patients, hospital, medical, treatment, amp, mind amp, disease, dr, body, blood, cancer, general hospital
Retirement	retirement, cpf, savings, provident, provident fund, central provident, fund, age, retire, payouts, income, cpf savings
Public Housing	housing, housing board, flat, flats, hdb, unites, room, order, resale, build, buyers, national development, property
Public Transport	transport, bus, lta, smrt, commuters, mrt, services, stations, train, transit, trains, buses, cars, sbs, passengers
Jobs	manpower, employers, companies, workers, skills, jobs, work, employment, mom, employees, hiring, career, job
Education	education, school, students, schools, student, learning, parents, children, university, teachers, programmes, academic
Population Growth	immigration, population, population growth, economic, foreign, immigrants, foreigners, ageing, birth rate

Table 1: Some keywords for each issue.

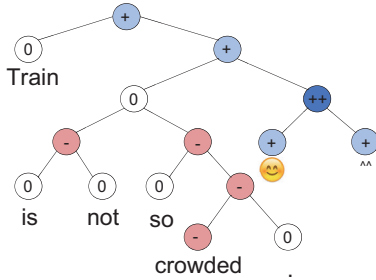


Figure 7: Labeled parse tree for “Train is not so crowded. \U0001F60A ^”

For the aforementioned reasons, we create our own sentiment treebank by manually labeling 5,291 randomly selected issue related posts. We use post “Train is not so crowded. \U0001F60A ^” as an example to explain how we label posts. This post contains emoji \U0001F60A and emoticon ^. First, we encode emojis and emoticons³ using unique codes. For example, we replace \U0001F60A to code ‘U0001F60A’ and ^ to code ‘emoticon0001’. The updated post is thus “Train is not so crowded. U0001F60A emoticon0001”. Next, we use the Stanford Parser (Klein and Manning 2003) to generate a parse tree for the updated post. The Stanford Parser considers our unique codes as noun words in the parse tree. We then replace the unique codes in the parse tree to the corresponding emojis and emoticons before the tree is manually assigned sentiment labels. Figure 7 shows the fully labeled parse tree for our example post. Note that each node in the parse tree is assigned one of the sentiment labels from very negative to very positive (—, —, 0, +, ++). To label a node, we consider only the part of the sentence covered by the subtree rooted at the node. For example, to label the first node on the third level, we examine the phrase “is not so crowded.” and assign a neutral label. Our labeling tool is built based on Stanford sentiment treebank labeling tool⁴.

Our labeled posts include 1,421 negative (labeled as — or —), 3,408 neutral (labeled as 0), and 462 positive (labeled as + or ++) posts. Less than 10% of our issue related posts are positive. This is in stark contrast with the surveyed user opinions in Figure 4 where we observe more positive users than negative users. It shows that although people may ex-

³The emoticons and emojis are found at: https://en.wikipedia.org/wiki/List_of_emoticons

⁴<http://nlp.stanford.edu:8080/sentiment/labeling.html>

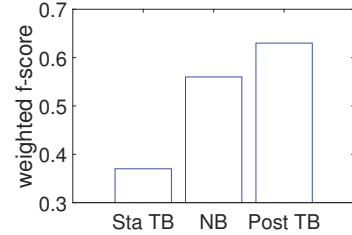


Figure 8: Classification results (weighted f-score) for issue related posts.

press many negative posts about an issue, their overall opinions on the issue can still be positive. For example, a user posting many times about crowded train may still feels positive about the overall Public Transport service in Singapore.

We then train a RNTN model on our post sentiment treebank (Post TB) and compare with the same model trained on the Stanford sentiment treebank (Sta TB) and Naive Bayes (NB, which considers emojis and emoticons). We evaluate the three models using weighted f-score (i.e., sum of f-score of each sentiment class weighted by the proportion of posts of each class) and the results are obtained using 5-fold cross validation. According to Figure 8, Post TB outperforms Stanford TB and NB by 26% and 7% respectively. We subsequently use Post TB to predict sentiments of social media posts.

Sentiment of Words When labeling parse trees, we need to label sentiments of all individual words (including English words, emojis and emoticons) which appear as leaf nodes of the trees. A word may appear multiple times during the labeling process. From our post sentiment Treebank, we obtain the sentiment of a word by taking the majority vote of its sentiment labels. In total, we obtain 662 positive words, 905 negative words and 10,763 neutral words which will be used in deriving sentiment features for user opinion prediction.

Features for Opinion Prediction

With our extracted *i*-related posts, the post sentiment classifier and the sentiment of words, we extract two types of features for predicting opinion on issue *i*: (a) *sentiment features* extracted from posts, and (b) *opinion features* extracted from user opinions on other issues.

Sentiment Features (SF) To construct sentiment features of a user, we use her statuses if she is a Facebook user, or her tweets and all her public followers’ and followees’ tweets if she is a Twitter user. Given a set of posts *P* (*P* can be sta-

tuses, or tweets), we define three sets of features for predicting user’s opinion on issue i . Let P_i be the set of i -related posts, W be the set of words in all posts, W_i be the set of words in P_i , P^+ be the positive posts in P , P^0 be the neutral posts in P , P^- be the negative posts in P , W^+ be the positive words in W , W^0 be the neutral words in W , and W^- be the negative words in W . If a word never appeared in our sentiment treebank, we consider it neutral.

The first set of features are: the proportion of i -related posts that are positive, neutral, and negative, i.e., $\frac{|P_i^+|}{|P_i|}$, $\frac{|P_i^0|}{|P_i|}$, and $\frac{|P_i^-|}{|P_i|}$, and the proportion of words that are positive, neutral, and negative in i -related posts, i.e., $\frac{|W_i^+|}{|W_i|}$, $\frac{|W_i^0|}{|W_i|}$, and $\frac{|W_i^-|}{|W_i|}$. These features indicate if the user posts some positive, neutral, or negative content about the issue.

The second set of features are: the proportion of all posts that are positive, neutral and negative, i.e., $\frac{|P^+|}{|P|}$, $\frac{|P^0|}{|P|}$, $\frac{|P^-|}{|P|}$, and the proportion of words that are positive, neutral, and negative in all posts, i.e., $\frac{|W^+|}{|W|}$, $\frac{|W^0|}{|W|}$, $\frac{|W^-|}{|W|}$. This set of features tells if the user posts some positive, neutral, or negative content in general.

The third set of features are: the features from the first set divide by the features from the second set, i.e., $\frac{|P_i^+|}{|P_i|} / \frac{|P^+|}{|P|}$, $\frac{|P_i^0|}{|P_i|} / \frac{|P^0|}{|P|}$, $\frac{|P_i^-|}{|P_i|} / \frac{|P^-|}{|P|}$, $\frac{|W_i^+|}{|W_i|} / \frac{|W^+|}{|W|}$, $\frac{|W_i^0|}{|W_i|} / \frac{|W^0|}{|W|}$ and $\frac{|W_i^-|}{|W_i|} / \frac{|W^-|}{|W|}$. This set of features tells if the user is more positive, neutral, or negative when posting about the issue than when posting general content.

For i -silent users, the first and the third feature sets will have feature value 0, as they do not have i -related posts.

Opinion Features (OF) We consider the user’s opinions on other issues as the second type of features for opinion prediction. The intuition is that: (a) the user may have certain sentiment bias on all issues. For example, some users are more likely to be negative, but some are more likely to be positive; (b) the user’s opinion on an issue may be correlated with or similar as her opinion on some other issue. For the above reasons, we attempt to predict the user’s opinion on a target issue by making use of her opinions on other issues. To extract the opinion features, we consider two cases. The first case is when we have already acquired a user’s ground truth opinions on other issues. This case could happen in some real applications. For instance, one may want to predict a user’s interests by knowing her other interests, to predict a user’s interests by knowing her gender, or to predict a user’s age by knowing her interests. Another case is that we do not have the user’s ground truth opinions on other issues. This case may be more common. For the first case, we directly use a user’s opinions on other issues as features. For the second case, we first predict the user’s opinions on other issues using only sentiment features from the content. We then use the predicted results as opinion features.

Opinion Prediction Results

With the above features, we train a SVM classifier to predict user opinion in Twitter and another classifier for Face-

	Hea.	Ret.	Hou.	Tra.	Job.	Edu.	Pop.
Negative	6	8	14	23	6	10	17
Neutral	26	19	36	41	52	29	30
Positive	20	9	21	33	31	58	17
i -silent	42	29	52	49	60	58	44
i -active	10	7	19	48	29	39	20

Table 2: Class distribution for issues from Twitter users.

	Hea.	Ret.	Hou.	Tra.	Job.	Edu.	Pop.
Negative	5	8	12	14	3	8	9
Neutral	24	20	24	31	36	22	29
Positive	16	9	16	22	25	38	10
i -silent	28	26	32	34	36	35	31
i -active	17	11	20	33	28	33	17

Table 3: Class distribution for issues from Facebook users.

book. In our evaluation, we use 1000 posts (or less if the user does not post this number of posts) from each Twitter user or Facebook user. For a Twitter user, we also use at most 1000 tweets from each public followee or follower of the user. Tables 2 and 3 show the class distribution and the number of i -silent and i -active users for the seven issues for Twitter and Facebook users respectively. As the number of negative users in all issues are usually very small, to ensure the significance of our results, we show f-score for positive class with at least 20 users. The f-score is obtained with 5-fold cross validation. Again, we consider the issues that have at least 20 i -silent users and 20 i -active users. Finally, only Public Transport, Jobs, and Education issues meet our criteria in both Twitter and Facebook.

Tables 4 and 5 show the opinion prediction results for Twitter and Facebook users respectively. The baseline methods are a random predictor and a SVM classifier using unigrams from users’ posts. Our methods include: (a) the sentiment features (SF) from user content, (b) the sentiment features from users’ posts and opinion features (OF) from predicted user opinions on other issues, and (c) the sentiment features from users’ posts plus the opinion features from ground truth user opinions on other issues. For Twitter users, there are three kinds of user content, namely: (a1) users’ tweets, (a2) user public followees’ tweets, and (a3) user public followers’ tweets. For Facebook users, user content refers to Facebook statuses of the users.

We summarize our findings as follows. Firstly, for both Twitter and Facebook users, all our methods outperform the baseline methods significantly for both i -silent users and i -active users. It suggests that considering the sentiment of posts and words can achieve better performance than considering the words alone. Secondly, for both Twitter and Facebook users, the prediction accuracy of i -active users’ opinions is better than that of i -silent users. This findings is expected as i -active users contribute posts about issue i . Thirdly, we can predict i -silent users’ opinions with reasonable accuracy although they do not post i -related posts. It implies that i -silent users’ i -unrelated tweets can be used to predict their opinions on i . Fourthly, the sentiment features from user tweets, user followers’ tweets and user followees’

		Random	Unigrams- user tweets	SF-user tweets	SF- followee tweets	SF- follower tweets	SF-user tweets + OF- predicted opinion on other issues	SF-user tweets + OF- ground truth opinion on other issues
Public Transport	All users	0.34	0.43	0.51	0.49	0.50	0.51	0.54
	<i>i</i> -silent users	0.31	0.34	0.45	0.40	0.42	0.46	0.50
	<i>i</i> -active users	0.38	0.45	0.52	0.51	0.54	0.56	0.58
Jobs	All users	0.35	0.38	0.50	0.51	0.50	0.52	0.55
	<i>i</i> -silent users	0.33	0.41	0.49	0.51	0.52	0.50	0.53
	<i>i</i> -active users	0.38	0.31	0.50	0.51	0.47	0.52	0.61
Education	All users	0.60	0.66	0.74	0.71	0.66	0.74	0.75
	<i>i</i> -silent users	0.55	0.61	0.72	0.67	0.64	0.71	0.71
	<i>i</i> -active users	0.67	0.68	0.75	0.76	0.69	0.77	0.80

Table 4: Opinion prediction results (f-score for postive class) using SVM for Twitter users

		Random	Unigram- user statuses	SF-user sta- tuses	SF-user statuses + OF-predicted opinion on other issues	SF-user statuses + OF-ground truth opinion on other issues
Public Transport	All users	0.33	0.33	0.48	0.49	0.69
	<i>i</i> -silent users	0.26	0.18	0.41	0.41	0.42
	<i>i</i> -active users	0.39	0.38	0.53	0.55	0.82
Jobs	All users	0.39	0.45	0.46	0.55	0.69
	<i>i</i> -silent users	0.36	0.34	0.40	0.52	0.61
	<i>i</i> -active users	0.43	0.54	0.54	0.60	0.74
Education	All users	0.56	0.63	0.70	0.72	0.72
	<i>i</i> -silent users	0.51	0.53	0.68	0.68	0.71
	<i>i</i> -active users	0.61	0.70	0.73	0.75	0.74

Table 5: Opinion prediction results (f-score for postive class) using SVM for Facebook users

tweets yield similar performance. The findings imply that to predict a *overall* silent user’s opinions, we may consider her neighbors’ posts. Finally, combining the sentiment features and the opinion features from predicted user opinions on other issues usually yields better performance than using the sentiment features only, and furthermore, combining the sentiment features and the opinion features from ground truth user opinions on other issues achieves the best performance. It suggests that a user’s opinions on other issues can help predict the user’s opinion on this issue.

Discussion and Conclusion

The main contributions of this work is to study the existence of issue-specific silent users and their opinions. We focus on two popular social media platforms, Twitter and Facebook, and conduct a survey to obtain users’ opinions on seven different topical issues (Healthcare Cost, Retirement, Public Housing, Public Transport, Jobs, Education, and Population Growth) and to collect users’ personal social media data. To our best knowledge, similar study was not conducted before. Our study has found that more than half of the users who are interested in issue *i* are *i*-silent users in both Twitter and Facebook. This finding suggests that people not posting about an issue does not imply that they are not interested in this issue. Hence, a large number of *i*-silent users’ opinions will be overlooked if we only consider *i*-active users’ posts only.

We also find that for the selected issues, *i*-silent users are more likely to be neutral than *i*-active users, and *i*-active users are more likely to be positive than *i*-silent users. This finding suggests that *i*-silent and *i*-active users are likely to

hold different opinion distributions. Thus, to profile the public opinion about an issue *i*, it is important to take *i*-silent users’ opinions into account.

Our work also contributes to the opinion prediction for *i*-silent users as well as *i*-active users in Twitter and Facebook. Opinion prediction for social media users is a challenging task, as we notice that people may give negative feedback about an issue but at the same time feeling overall positive about the issue. In other words, people may express unhappiness about one aspect of an issue but still feel positive for most other aspects.

We explore two types of features for opinion prediction task: the sentiment features extracted from users’ content and the opinion features extracted from users’ predicted opinions or ground truth opinions on other issues. We demonstrate the effectiveness of these features and show that although predicting *i*-active users’ opinion yield better performance than that of *i*-silent users, it is still possible to predict *i*-silent users’ opinions by leveraging on their *i*-unrelated content. We can have better performance if we make use of predicted *i*-silent users’ opinions on other issues and achieve the best performance if we acquire the ground truth *i*-silent users’ opinions on other issues. To be able to predict *i*-silent users’ opinions will enable researchers to infer the opinion distribution in population level, and also have a better understanding of *i*-silent users.

As the first attempt to study issue specific-silent users, this work has been confined to a small user community in a single country and the survey conducted can be affected by response bias. We therefore plan to study the opinions of issue specific-silent users for a much larger user commu-

nity, covering possibly larger set of issues and in other countries/regions. More research is clearly required to improve the accuracy of opinion prediction, particularly for the silent users. It is also interesting to find out the reasons for users to stay silent on an issue and for them to post after staying silent for some time.

Acknowledgments

This research is supported by the Singapore National Research Foundation under its International Research Centre@Singapore Funding Initiative and administered by the IDM Programme Office, Media Development Authority (MDA).

References

- Agarwal, A.; Xie, B.; Vovsha, I.; Rambow, O.; and Passonneau, R. 2011. Sentiment Analysis of Twitter Data. In *LSM 2011*.
- Asur, S., and Huberman, B. A. 2010. Predicting the Future with Social Media. In *WI-IAT 2010*.
- Barbosa, L., and Feng, J. 2010. Robust Sentiment Detection on Twitter from Biased and Noisy Data. In *COLING 2010*.
- Birmingham, A., and Smeaton, A. F. 2011. On Using Twitter to Monitor Political Sentiment and Predict Election Results. In *Sentiment Analysis where AI meets Psychology Workshop at IJCNLP 2011*.
- Bollen, J.; Mao, H.; and Zeng, X. 2011. Twitter Mood Predicts the Stock Market. *Journal of Computational Science* 2(1):1–8.
- Ceron, A.; Curini, L.; Iacus, S. M.; and Porro, G. 2014. Every Tweet Counts? How Sentiment Analysis of Social Media Can Improve Our Knowledge of Citizens Political Preferences with an Application to Italy and France. *New Media & Society* 16(2):340–358.
- Chung, J. E., and Mustafaraj, E. 2011. Can Collective Sentiment Expressed on Twitter Predict Political Elections? In *AAAI 2011*.
- Das, S., and Kramer, A. 2013. Self-Censorship on Facebook. In *ICWSM 2013*.
- Dave, K.; Lawrence, S.; and Pennock, D. M. 2003. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In *WWW 2003*.
- Ellison, N. B.; Steinfield, C.; and Lampe, C. 2007. The Benefits of Facebook Friends: Social Capital and College Students Use of Online Social Network Sites. *Journal of Computer-Mediated Communication* 12(4):1143–1168.
- Gayo-Avello, D. 2012. “I Wanted to Predict Elections with Twitter and all I got was this Lousy Paper”—A Balanced Survey on Election Prediction using Twitter Data. *arXiv preprint arXiv:1204.6441*.
- Gong, W.; Lim, E.-P.; and Zhu, F. 2015. Characterizing Silent Users in Social Media Communities. In *ICWSM 2015*.
- Guerra, P. C.; Meira, Jr., W.; and Cardie, C. 2014. Sentiment Analysis on Evolving Social Streams: How Self-report Imbalances Can Help. In *WSDM 2014*.
- Hampton, K.; Rainie, L.; Lu, W.; Dwyer, M.; Shin, I.; and Purcell, K. 2014. Social Media and the “Spiral of Silence”. *Pew Research Center; Published August 26, 2014, available online*.
- Kiciman, E. 2012. OMG, I Have to Tweet that! A Study of Factors that Influence Tweet Rates. In *ICWSM 2012*.
- Klein, D., and Manning, C. D. 2003. Accurate Unlexicalized Parsing. In *ACL 2003*.
- Kouloumpis, E.; Wilson, T.; and Moore, J. 2011. Twitter sentiment analysis: The good the bad and the omg! In *ICWSM 2011*.
- Kwak, H.; Lee, C.; Park, H.; and Moon, S. 2010. What is Twitter, a Social Network or a News Media? In *WWW 2010*.
- Lin, Y.-R.; Margolin, D.; Keegan, B.; and Lazer, D. 2013. Voices of Victory: A Computational Focus Group Framework for Tracking Opinion Shift in Real Time. In *WWW 2013*.
- Liu, B. 2012. Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies* 5(1):1–167.
- Maynard, D.; Bontcheva, K.; and Rout, D. 2012. Challenges in Developing Opinion Mining Tools for Social Media. In *Proceedings of @NLP can u tag #usergeneratedcontent?! Workshop at LREC 2012*.
- Mustafaraj, E.; Finn, S.; Whitlock, C.; and Metaxas, P. T. 2011. Vocal Minority Versus Silent Majority: Discovering the Opinions of the Long Tail. In *Privacy, security, risk and trust (passat), SocialCom 2011*.
- Pak, A., and Paroubek, P. 2010. Twitter as a Corpus for Sentiment Analysis and Opinion Mining. In *LREC 2010*.
- Pang, B., and Lee, L. 2008. Opinion Mining and Sentiment Analysis. *Found. Trends Inf. Retr.* 2(1-2):1–135.
- Pang, B.; Lee, L.; and Vaithyanathan, S. 2002. Thumbs Up?: Sentiment Classification Using Machine Learning Techniques. In *EMNLP 2002*.
- Popescu, A.-M., and Etzioni, O. 2005. Extracting Product Features and Opinions from Reviews. In *HLT 2005*.
- Preece, J.; Nonnecke, B.; and Andrews, D. 2004. The Top Five Reasons for Lurking: Improving Community Experiences for Everyone. *Computers in Human Behavior* 20(2).
- Sang, E. T. K., and Bos, J. 2012. Predicting the 2011 Dutch Senate Election Results with Twitter. In *Proceedings of Workshop on Semantic Analysis in Social Media 2012*.
- Skoric, M.; Poor, N.; Achananuparp, P.; Lim, E.-P.; and Jiang, J. 2012. Tweets and Votes: A Study of the 2011 Singapore General Election. In *HICSS 2012*.
- Sleeper, M.; Balebako, R.; Das, S.; McConahy, A. L.; Wiese, J.; and Cranor, L. F. 2013. The Post That Wasn’t: Exploring Self-censorship on Facebook. In *CSCW 2013*.
- Socher, R.; Perelygin, A.; Wu, J.; Chuang, J.; Manning, C. D.; Ng, A.; and Potts, C. 2013. Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. In *EMNLP 2013*.
- Tagarelli, A., and Interdonato, R. 2013. “Who’s out There?”: Identifying and Ranking Lurkers in Social Networks. In *ASONAM 2013*.
- The Unicode Consortium. 2015. The Unicode Standard, Version 8.0.0. <http://www.unicode.org/versions/Unicode8.0.0/>.
- Tumasjan, A.; Sprenger, T. O.; Sandner, P. G.; and Welpe, I. M. 2010. Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. In *ICWSM 2010*.
- Wang, Y.-C.; Burke, M.; and Kraut, R. E. 2013. Gender, topic, and audience response: An analysis of user-generated content on facebook. In *CHI 2013*.
- Wilson, T.; Wiebe, J.; and Hoffmann, P. 2005. Recognizing Contextual Polarity in Phrase-level Sentiment Analysis. In *HLT 2005*.