

Cross-Hierarchical Communication in Twitter Conflicts

Zhe Liu*

College of Information Sciences and Technology
The Pennsylvania State University
University Park, Pennsylvania 16802
zul112@ist.psu.edu

Ingmar Weber

Qatar Computing Research Institute
PO Box 5825, Doha, Qatar
iweber@qf.org.qa

ABSTRACT

Social hierarchy plays an important role in shaping the way individuals interacting with each other. In this study, we propose three metrics: equality, diversity, and reciprocity to evaluate the social hierarchical differences in cross-ideological communication on Twitter. We do this within the context of three diverse conflicts: Israel-Palestine, US Democrats-Republicans, and FC Barcelona-Real Madrid. In all cases, we collect data around a central pair of Twitter accounts representing the two main parties. Our results show in a quantitative manner that social hierarchy can be considered a factor that impacts individual's communication in Twitter conflicts. As one of the first literatures in this area, we demonstrate social hierarchy's effect in online environments.

Categories and Subject Descriptors

J.4 [Social and Behavioral Sciences]: Sociology

Keywords

Conflict, social hierarchy, Twitter, social network

1. INTRODUCTION

Social stratification has been proved to play a substantial role in shaping the way people interact and perceive others, and has been investigated in sociological studies. The recent advent of social networking sites has provided great data sources for researches on social stratification. Besides the advantages of large data size, social networking sites also contain rich interactions between cross-ideological individuals [1]. Studying such kind of cross-hierarchical communication, especially in conflicts, can be very important, as lately "digital hate" on Twitter grows at an alarming speed [2]. In order to empirically understand how cross-hierarchical communication happens in Twitter conflicts, in this study we asked questions like: How much does the social hierarchy matter in cross-ideological communication? And how universal are such patterns across different types of polarized conflicts?

For our study, we choose three conflicts of very different natures, and identify likely supporters of either camp by using retweet signals of certain seed node in combination with additional rules. We validate our ideological labels, as well as all other labels used for this project, through extensive crowd-sourced evaluation. We evaluated the cross-hierarchical communication in all conflicts via three assessment metrics namely, (i) equality, (ii) diversity, and (iii) reciprocity. We find that individuals from the bottom class

*The work was done while the author was an intern at QCRI

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. Copyright is held by the owner/author(s).

HT'14, Sep 01-04, 2014, Santiago, Chile

ACM 978-1-4503-2954-5/14/09.

<http://dx.doi.org/10.1145/2631775.2631788>

initiate the most communication toward their foes in the topmost social hierarchical level. Although those bottom people tend to mention a diverse audience inter-ideologically, under most cases they get ignored and received no response back. We believe our results advances individual's understanding of the opportunities and limitations provided by Twitter for facilitating cross-hierarchical interactions.

2. DATA COLLECTION AND LABELING

To automatically detect users with similar or different ideologies, we started with three pairs of opposing seed users: @AlqassamBrigade and @IDFSpokesperson, @TheDemocrats and @GOP, and @FCBarcelona_es and @realmadrid. We intentionally chose these Twitter accounts as seed nodes considering their key roles in well-known Twitter conflicts. For each of the seed nodes, we obtained up to 3,200 of its latest tweets using the Twitter API. As retweet often indicates endorsement and preference of a message, then for each tweet we identified up to 10 retweeters and labeled them as the supporters of the corresponding seed node. To remove mediators and neutral intervenors, such as peace movement organizations and journalists, we first identified the top retweeted accounts from each side. After manually removing those with neutral leanings (e.g. @BBCBreaking, @Reuters, etc.), we kept only the top 10 influential ones from each camp. We next labeled users who retweeted the top 10 influential accounts from one side 4 times less than from the other side as with neutral or unclear leanings in the conflict, and eliminated them from our datasets. Here we chose 4 as the threshold because it provided the most reasonable balance between the classification accuracy and the size of data for later analysis, compared to other values ranging from 2 to 5. We validated our classification results via CrowdFlower (crowdfunder.com) by assigning 100 random users in each ideology to the HIT workers. By comparing user's pre-assigned ideology to the majority-voted label obtained from CrowdFlower, we found that our classification method worked well over all three datasets, yielding an average accuracy of 96.2%. With the classified users, we extracted all mentions between them as interactions between ideological-friends and foes. The descriptions of our collected data sets were shown in Table 1.

Table 1. Properties of the three collected datasets

Conflict	#Users	#Inter-Interaction	#Intra-Interaction
PA-IL	9,937	4,829	178,255
DEM-REP	17,869	20,257	576,848
FCB-RMFC	28,218	21,089	152,799

Next, we split the classified users into four social hierarchical groups according to their number of followers, including: the top 1%, the top 1%-10%, 10%-70%, and 70%-100% users. The division is arbitrary, but we think that given the specialty of the top and bottom users in Twitter, our grouping scheme can be most useful in comparing the communication patterns across social hierarchies. Also, we believe, that the number of followers at least partly indicates a person's degree of influence on the social

network [3], even though it may not fully represent the social “status” of an individual in real world.

3. ANALYSIS OF CROSS-HIERARCHICAL COMMUNICATION

We adopted three assessment metrics, namely equality, diversity, and reciprocity, to explore the cross-hierarchical communication flows. One-way ANOVA tests with post-hoc Tamhane’s T2 analyses were utilized to determine statistical significance at a p value of 0.05 or less.

Equality requires all individuals to engage and contribute equally in communication. People in lower hierarchies are supposed to participate less, as they are expected to have little power or control over outcomes. We quantify a user’s engagement by the total number of mentions each individual makes to their ideological-foes.

Diversity measures the degree of variation in an individual’s conversational network. We adopt a metrics called E-I ratio to measure the diversity of one’s communication. The E-I ratio captures the extent to which one’s communication network is composed of individuals that are across ideologies as compared to within ideologies. The E-I ratio is calculated as:

$$EI_i = \frac{E_i - I_i}{E_i + I_i}$$

where E_i = number of ideological-foes user i has mentioned, I_i = number of ideological-friends user i has mentioned. The E-I ratio ranges from -1 to 1. The closer the E-I index is to -1, the more likely an individual only talk to ideological-friends.

Reciprocity refers to dyadic information exchanges between individuals. It is assumed that individuals in lower social hierarchy engage in more dyadic interactions than individuals in higher hierarchical level since communication constructed from the higher-ranked individuals are more often one-to-many interactions. To evaluate the reciprocity levels across ideologies and hierarchies, we adopted the maximum length of inter-ideological conversations as the measurement.

4. RESULTS

Figure 1 showed the conditional probability of a communicator interacting with another, given their social hierarchies. Here we used the PA - IL conflict for illustration purpose and only reported findings that can be generalized to all three datasets. The results were visualized using parallel sets, with the width of the bar denoted the number of interactions existed cross-hierarchical levels. We saw that except the bottom-most level, users from the other three hierarchies have about the same probabilities of being mentioned by their friends. However, under an inter-ideological context, we noticed that users in the top-most hierarchical level have the highest chance of receiving a mention initiated by their foes, which is even higher than the sum of the probabilities derived from the remaining three levels. This indicated that people are more willing to attack or challenge “authorities”. Besides, under both conditions, there is very little chance that the bottom 30% of users will be mentioned by either their friends or foes. In addition, from viewing the width of all “ribbons”, we found that users from the bottom-most hierarchical level maintain the highest probability of initiating a mention of the top 1% of users.

The ANOVA results on equality demonstrated no significant social hierarchical difference, except on the DEM-REP dataset ($F_{(3, 6977)} = 27.45, p = 0.00$). Tamhane’s T2 tests further revealed that only users in the bottom group are involved in significantly less

with their ideological-foes, whereas, the other three social hierarchical groups did not differ on the total number of interactions that they initiated. In that sense, we claim that in general, social hierarchy has little or no effect on user’s participation in cross-ideological communication.

The one way ANOVA tests on E-I ratio also showed significant differences for all three datasets (PA-IL: $F_{(3, 7374)} = 25.29, p = 0.00$; DEM-REP: $F_{(3, 15963)} = 24.06, p = 0.00$; and FCB-RMCF: $F_{(3, 22004)} = 62.34, p = 0.00$). Tamhane’s T2 further revealed that the second hierarchical group of both political datasets had the significantly lowest E-I ratio, indicating that people in that social hierarchy are more insular toward their ideological-foes. In contrast, the bottom hierarchy exhibited the highest tendency towards inter-ideological communications.

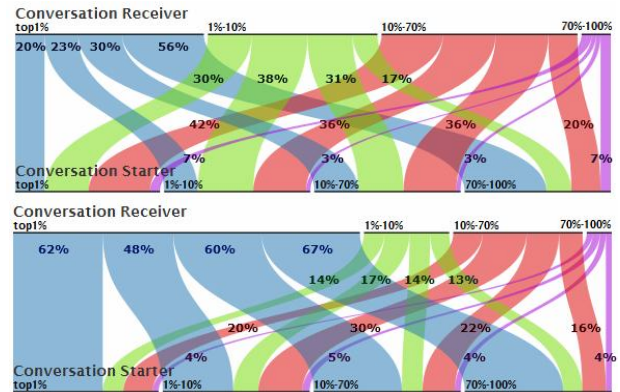


Figure 1: Intra (above) and inter-ideological (below) communication across hierarchies.

Social hierarchy also demonstrated significant impact on the continuity of inter-party discussions on PA-IL ($F_{(3, 3768)} = 4.60, p = 0.00$) and FCB-RMCF ($F_{(3, 7593)} = 10.24, p = 0.00$) datasets, whereas the DEM-REP data failed to demonstrate such difference on inter-ideological conversation length ($F_{(3, 104774)} = 27.56, p = 0.06$). Post hoc analyses further revealed that in general conversation starters from the bottom hierarchy had significantly less back and forth exchanges in cross-ideological conversations.

5. CONCLUSION

We present an initial analysis on the effect of social hierarchy on cross-ideological communication in Twitter conflicts. Our results provided evidence that people in the bottom social hierarchy tend to be aggressive toward their ideological-foes in the upper classes, however, more than often they tend to be ignored by their adversaries. In addition, our results showed that the two political conflicts revealed approximately the same cross-hierarchical communication characteristics, whereas the sports-related contradiction exhibited quite different patterns. Future work will involve the investigation of other aspects of inter-ideological communications, such as tweet politeness and informative.

6. REFERENCES

- [1] Liu, Z., and Weber, I. Predicting ideological friends and foes in Twitter conflicts. *WWW companion 2014*
- [2] D. MacMillan. Twitter aids rise of web-based hate forums. <http://www.bloomberg.com/news/2013-05-07/twitter-aids-rise-of-web-based-hate-forums-report-finds.html>. May 2013.
- [3] Kwak, C. Lee, H. Park, and S. Moon. What is twitter, a social network or a news media? *WWW2010*