Seeking Like-Minded People: Twitter Users' Homophily based on Ideological Hashtags

Sifan Xu
University of Tennessee Knoxville

Abstract:

- Twitter generates vibrant discussions related to key sociopolitical issues and has great ability to project various discourses into public arena. Yet, these discourses can be overwhelming and heated, in particular when controversial events happen. In this study, based on 100,000 tweets about Gillette's controversial cause-related marketing campaign in 2019, exponential random graph models (ERGM) were used to investigate the homophily tendency of users who used certain hashtags. Results show the mention network of these users showed homophily tendency based on hashtags. Homophily in this study was distinguished based on attraction of common users (i.e. increased chance of ties for users who both engage the hashtag) and alienation of nonusers (i.e. decreased chance of ties for users neither of whom engages the hashtag), and the comparison between ideological hashtags and conceptual hashtags revealed that homophily only manifested through ideological hashtags

Data Collection and Method:

- Tweets were collected using R package *rtweet* over the span of 11 days following Gillette's release of its advertisement "The Best Man Can Be" on twitter from January 18, 2019 to January 28, 2019. The keyword for search is "Gillette." Only original tweets (i.e. tweets that have any original content, including tweets that quoted other tweets) were collected.
- The final dataset contained 107,641. The number of words of all the tweets in the dataset ranged from 1 to 123, with the median of 17 words and average of 21.94 words
- First, since some users tweeted more than once, we combined all the tweets based on user ids. Next all the hashtags used by each user were extracted. Based on the definition of ideological markers and conceptual markers provided by Blevins et al. (2019), top hashtags were categorized as either ideological markers or conceptual markers, and then users were tagged accordingly as whether the user used a particular hashtag. Then the mention (@user in each tweet) network was extracted and the tags were passed along as nodal attributes for statistical modeling[1]. For users who were mentioned but did not have any actual content in the dataset to judge whether the nodes used certain hashtags or not, their nodal attributes were coded as missing

Results:

- For ideological hashtags #MeToo and #MAGA, homophily effects did exist for both, but they worked in a different pattern, with #MeToo driving homophily by attraction of common users (increased chance of ties for users who both engage the hashtag) and #MAGA driving homophily by alienating nonusers (decreased chance of ties for users neither of whom engages the hashtag).

	Estimate	SE	P
Constant (edges)	-10.484	0.0043	< .001***
Neither engaged #MeToo	-0.035	0.0463	0.450
Both engaged #MeToo	2.138	0.1561	< .001 ***
Neither engaged #MAGA	-0.359	0.0505	< .001 ***
Both engaged #MAGA	-0.051	0.5808	0.930
Neithr engaged #ToxicMasculinity	-0.069	0.0441	0.117
Both engaged #ToxicMasculinity	0.307	0.2818	0.275
Neither engaged #TheBestMenCanBe	0.409	0.0276	< .001***
Both engaged #TheBestMenCanBe	-0.878	0.2908	0.003**
Indegree main effect of #MeToo	-0.390	0.0535	< .001 ***
Outdegree main effect of #MeToo	0.229	0.0440	< .001 ***
Indegree main effect of #MAGA	-0.047	0.0583	0.415
Outdegree main effect of #MAGA	0.282	0.0525	< .001***
Indegree main effect of #ToxicMasculinity	-0.173	0.0484	<.001***
Outdegree main effect of #ToxicMasculinity	0.386	0.0410	< .001***
Indegree main effect of #TheBestMenCanBe	-1.553	0.0442	< .001***
Outdegree main effect of #TheBestMenCanBe	0.492	0.0220	< .001***

Table 1. Coefficients of the ERGM model.

(Results continue)

- In comparison, for conceptual hashtags #ToxicMasculinity and #TheBestMenCanBe (ad slogans), neither of the homophily patterns existed. In fact, for users who both engaged #TheBestMenCanBe, their chance of forming a tie significantly decreased; and for users neither of whom engaged the hashtag, their chance of forming a tie significantly increased. This is opposite to homophily.
- A bar graph representing coefficients including confidence intervals is available here

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Reference: Blevins, J. L., Lee, J. J., McCabe, E. E., & Edgerton, E. (2019). Tweeting for social justice in #Ferguson: Affective discourse in Twitter hashtags. *New Media & Society, 21*(7), 1636-1653.

[1] Mention networks and retweet networks are two most common types of conversational networks formed by Twitter users. However, it should be noted here that these two types of networks also differ from each other, with the former ones more significant in ways how users interact (Conover et al., 2011). The mention network was used because that is how conversations emerge, through the actual exchange of information and active contributions to dialogues, not simply relaying what others have said.