Analysis of Moral Divide in Online Communities

Du Baofa* Kazutoshi Sasahara†

Abstract

In the modern society, there are many communities in the online social network. We study online communities related to three moral issues – #LGBT, #Metoo, and #Kutoo – in terms of using retweet networks, moral and emotion loading, topics and popular words using Twitter data. Our results revealed commonness and differences in these communities, providing insights into online moral divide.

1 Introduction

In the modern society, there are many communities in the online social network. How should we define them? Is there any connections? Although inferring communities features from social data is important in research in the social sciences as well as application in social media, complex features in different communities remain unstudied. Here we focus three moral topics of debate about human rights: #LGBT, #Metoo and #Kutoo. The analysis of the three hot topics can help us seek common characteristics and differences in different communities. The brief explanations about these topics are written below.

- #LGBT: LGBT (or GLBT) is an initialism that stands for lesbian, gay, bisexual, and transgender [1].
- #Kutoo: The #KuToo movement is an ongoing movement in Japan against the high heel policy in workplaces [2].
- #Metoo: The Me Too movement (or #MeToo movement), with a large variety of local and international related names, is a movement against sexual harassment and sexual assault [3].

The framework for our research draws from one important concept to quantify the morality. Moral Foundations Theory developed by Jonathan Haidt [4], holds that every person innately has five moral foundations that allow him or her to intuitively judge moral situations. They are:

- Harm (i.e., disliking the pain of others and feeling responsible for protecting vulnerable people);
- Fairness (i.e., taking the right or just action based on accepted rules);
- Ingroup loyalty (i.e., being loyal to social ingroups, including family and nation);
- Authority (i.e., respecting and obeying tradition and legitimating authority);
- Purity (i.e., feeling antipathy for disgusting things and contamination).

We measured those five moral foundations in English texts with reference to the original Moral Foundations Dictionary (MFD and Japanese texts with reference to the Japanese Moral Foundations Dictionary (J-MFD) [5].

*Graduate School of Informatics, Nagoya University, du.baofa@h.mbox.nagoya-u.ac.jp

†Graduate School of Informatics, Nagoya University; Center for Computational Social Science, Kobe University; JST PRESTO, sasahara@nagoya-u.jp

2 Data and Methods

2.1 Data

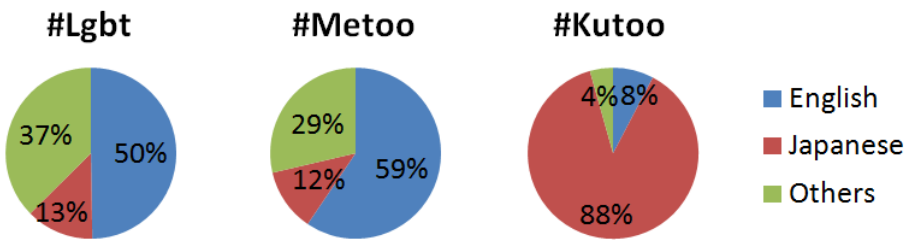

Figure 1: Distribution of tweets in different languages

For the data of different communities, we collected tweets that include the hastags #LGBT, #Metoo, and #Kutoo using Twitter Standard search API. There were 383,060 #LGBT tweets from July 23 to October 1. There were 219,629 #Metoo tweets and 1,503,173 #Kutoo tweets from June 4 to November 13.

Figure 1 shows the language percentage results in different hashtags. In the #LGBT hashtag, half of the tweets were sent by English, 13% in Japanese. English and Japanese were 59% and 12% for the #Metoo hashtag. As for the #Kutoo hashtag, 88% of the tweets were in Japanese because it is a theme from Japan. We would put main focus on the Japanese tweets analysis in this topic based on so high percentage of Japanese tweets.

2.2 Data Analysis

We analyzed the relationships of different communities from Twitter social data with the methods of morality analysis, emotion analysis, retweet network analysis, LDA topic modeling and word cloud.

At first, we used Mecab to separate Japanese tweets into words, and then extracted the moral loading like the study 1 based on the J-MFD to get the morality results. Emotion analysis was based on the "Emotion Expression Dictionary", and ML-Ask would seek expressions of specific emotion types in the emotive utterances. The emotion loading was the percentage of words in different emotion types.

Second, Twitter has a retweet function that allows one to spread posts displayed on the timeline to one's followers. The entire picture of such information diffusion can be represented by a network. A retweet event on Twitter can be represented by a directed link as original user to re-tweet user, using "the original user" and the "re-tweet user" as nodes and the directed link as edge. By listing such links for all retweets, it is possible to construct an information diffusion network in which nodes are users and directed links are retweet flows. This is called the "retweet network." According to the tweets data, we built a retweet network of different communities by Gephi and sought community features via LDA model and word cloud.

3 Results

To measure moral loadings, we separated tweet texts into words and computed the frequencies of moral words in Japanese and English defined by the J-MFD and MFD.

Figure 2 shows that users in the #Metoo community had higher fairness loading and harm loading than other communities. At the same time, a lower ingroup loading and authority loading were found in #Metoo community. The #LGBT and #Kutoo communities had nearly the same proportion in the ingroup, authority and purity part. In particular, the #Kutoo had the lowest moral loading in fairness compared with other communities, the #LGBT had the lowest moral loading in harm.

We used ML-ask to calculate the emotion loading for different communities. Figure 3 shows that the users in the #LGBT community have a high emotion expression in iya, yorokobi, suki, and yasu. Both the #Metoo and #Kutoo communities had very high emotion loading in iya. The iya emotion in the #Metoo community was even over 0.5. Most of users in these three topics mainly released the iya emotion.

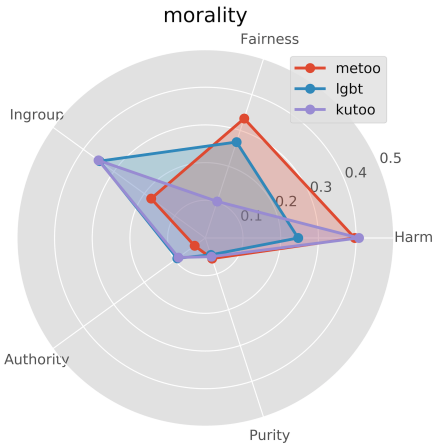

Figure 2: Moral loading in Japanese

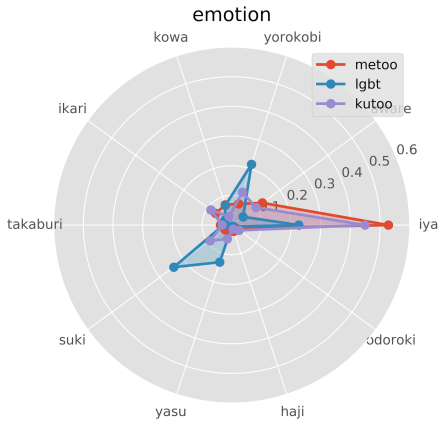

Figure 3: Emotion loading in Japanese

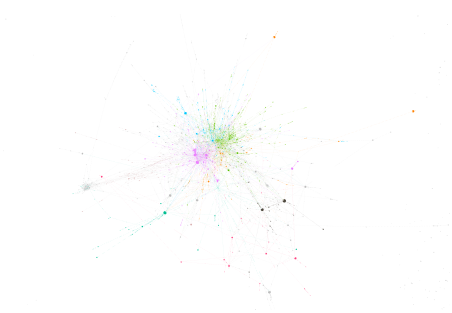

Figure 4: #LGBT network in English

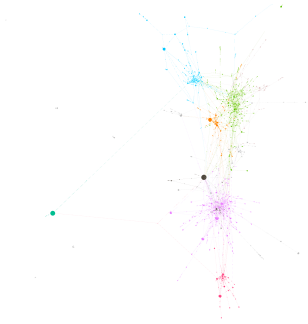

Figure 5: #LGBT network in Japanese

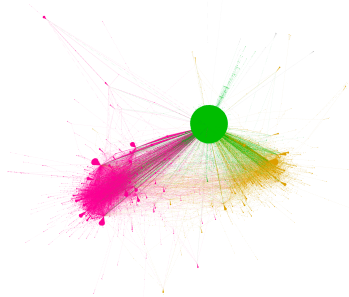

Figure 6: #Metoo network in Japanese (green: G_{mj1} red: G_{mj2} yellow: G_{mj3})

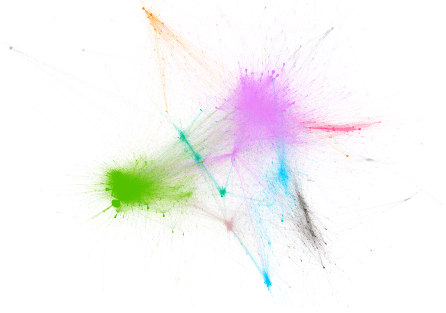

Figure 7: #Metoo network in English

The node and edge of network in English and Japanese is 32,269, 32,437 and 9,980, 10,563 respectively for #LGBT. The English retweet network(Figure 4) and the Japanese retweet network(Figure 5) have many communities structural features in common. This suggests that each community is sharing their tweets on a different topic. However, these two communities are based on the different situations. In the network of English, we can separate the core users into two types. The first type is the high correlation accounts about LGBT like “LGBT Nation,” “LGBTQ Shrink,” and “LGBTPlusProud.” They sent a lot of tweets about LGBT

Table 1: The LDA topic model result of #LGBT

Topic 1	'ゲイ','男装','女装','音楽','セット','ゲイバー', '今日','音楽サイト','本日','バイ'
Topic 2	'同性愛','映画','開催','家族','セクマイ','エイズ', '参加','ウガンダ','是非','日本'
Topic 3	'不妊','治療','精子提供','精子','セクマイ','セクマイカップル', 'ドナー','spermdonate','友達','フォロー'

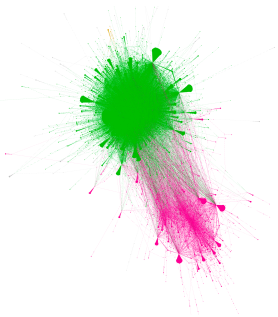

Figure 8: #Kutoo network in Japanese (green: G_{kj1} red: G_{kj2})

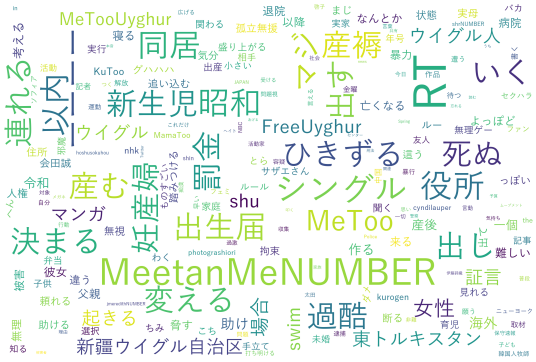

Figure 9: G_{mj1} #Metoo word cloud

in the daily life. The other type of accounts is some famous accounts who send the contents about LGBT topics by accident like "Zara Kay." On the other side, for Japanese Network, each user's attention to different topics forms different communities in the network.

The LDA model(Table 1) shows that in the Japanese LGBT network, the contents can be divided into three topics. The first topic is something about gay bar. The second topic is the contents about the account ''ウガンダの友人". This account sent the tweets to discuss the Uganda problem like AIDS for calling for some preventive measures against AIDS. The third topic is on fertility.

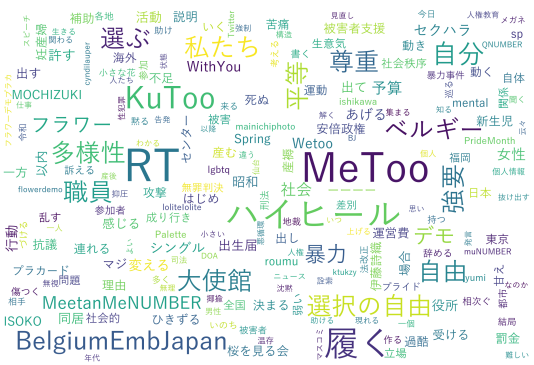

Figure 10: G_{mj2} #Metoo word cloud

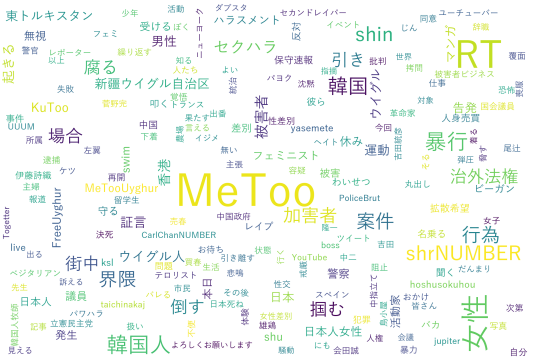

Figure 11: G_{mj3} #Metoo word cloud

The Japanese #Metoo retweet network (Figure 6) G_{mj} has three communities structure that are clearly separated from those of English (Figure 7). This suggests that each community is sharing a tweet on a different concentrated topic. Another notable feature of G_{mj} is that only the G_{mj1} community has a radial network and contains many users (63.21% of the total).It has a structure that is clearly different from other communities.

A closer look showed that the tweets of a particular user were affected by the spread caused by the spread of many users. The core account is ‘‘清水ともみ” which focuses on human rights of the Uygur ethnic group. In order to take advantage of the news of improper discourse on Metoo, which gained attention in the limelight in the country, “#Metoo” along with hashtags “# 東トルキスタン” and “# ウイグル”have been used.

A word cloud was created to know what words were used in communities G_{mj}. Figure 9 shows the G_{mj1} word cloud. In this figure, words are displayed up to the 200th appearance frequency (the size is proportional to the appearance frequency). This graph shows that along with Metoo, the words about the human rights of Uygur ethnic group frequently appeared. In the word cloud(Figure 10) of G_{mj2}, we can seek the words like ‘‘ベルギー,”‘‘大使館,” “KuToo” and ‘‘ハイヒール,” because the Belgian embassy tweeted about supporting free dress.The word cloud of G_{mj2} mainly contains the discussions about this topic. On the other hand, in the word cloud(Figure 11) of G_{mj3}, most of words in the word cloud are about news like ‘‘韓国,”‘‘韓国人,” ‘‘案件” and ‘‘暴行,”because of cases of sexual assault by Koreans.

The Japanese #Kutoo retweet network (Figure 8) G_{kj} has a two-community structure. This suggests that each community is sharing a different topic. G_{kj1} contains 76.12% users, G_{kj2} is 22.03%. The G_{kj1} community has nearly a radial network. In G_{kj1}, the core user’s content is based on the tweets about ‘‘田嶋陽子” and ‘‘石川優実”. They play a decisive role in building networks.

4 Discussion

Different communities showed totally different preferences. The LGBT community showed high ingroup, authority moral loading as well as “suki,” “yorokobi” emotion loading. According to the results of network and word clouds, the formation of English network in the #LGBT movement came from a large number of tweets by official accounts or several tweets by popular accounts. That means the users were influenced more by accounts for English network. On the other side, the Japanese network put more attention on the topics. The LDA model results show that the users in Japanese network focuses gay bar, Uganda AIDS problem and fertility.

As for Metoo community, it has a high fairness, harm moral loading and “iya” emotion loading. They want to defend themselves and show hatred feeling against the persecution. Compared with the network of English in #Metoo, the Japanese network has three clear communities structure. The aggregation of three groups means significant differences on topic focus. Based on the results of the network, word clouds and LDA model, the topics were mainly about Uygur ethnic group’s right, Belgian embassy’s support for women free dress and Koreans sexual assault.

The Kutoo community showed high harm, ingroup moral loading and “iya” emotion loading. This means defense for the women free dress right with high hate emotion toward the present situation. The network and word clouds show that the formation of the network is highly related to influential accounts like ‘‘田嶋陽子” and ‘‘石川優実”.

In summary, although these are all topics about safeguarding rights, different groups show different moral and emotional commonness and differences. They can also form different social networks with the attention to different topics. These results provide insights into how moral divde affect the emergence and evolution of online communities.

References

[1] https://en.wikipedia.org/wiki/LGBT.

[2] https://ja.wikipedia.org/wiki/KuToo.

[3] https://en.wikipedia.org/wiki/Me$_{T}oo_{m}ovement$.

[4] Jesse Graham, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean Wojcik, and Peter Ditto. Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, 47, 12 2012.

[5] Akiko Matsuo, Kazutoshi Sasahara, Yasuhiro Taguchi, and Minoru Karasawa. Development of the japanese moral foundations dictionary: Procedures and applications. *CoRR*, abs/1804.00871, 2018.