What makes Facebook groups resilient to misinformation?

Melissa Mokhtari[1], Leila Hamdad[2], and Oana Goga[1]

[1]CNRS, Inria, Institut Polytechnique de Paris
[2]ESI Alger

Abstract

1 Introduction

Previous works have shown that users with particular characteristics are more vulnerable to misinformation [AJJ19, MNA22]. For example, users that are over 65 yo and republican [AJJ19] are more prone to share misinformation that appears in their social media feeds. In this work, we take a related but complementary perspective in understanding group-level vulnerability to misinformation. More precisely, our goal is to study *how misinformation is disseminated in Facebook groups* and understand *what characteristics make some Facebook groups more vulnerable to disinformation than others?* In this abstract, we present our work's earliest stages and focus on the data collection strategy and *the hypothesis* we plan to test.

Context The user's vulnerability to misinformation has been addressed by many researchers. They worked on understanding the characteristics that differentiate robust users from vulnerable ones. Guess et al. [AJJ19] did a statistic study using data collected from a self-reported survey given to 3500 users in addition to collected data about shared links on these users' Facebook profiles. Guess had concluded that the age of the user is the main influent factor. Users over 65 years share more than 2-time disinformation content than users within 45 to 65 years. He also did deduce that political leaning influences the vulnerability to misinformation where republicans and those that define themselves as independent share more disinformation than democrats.

Mu et al. [MNA22] investigated whether users that share misinformation have a different language style. For this, they collected two datasets, the first one from Weibo and the second from Twitter using the Weibo and Twitter APIs respectively. Mu et al. searched for profiles sharing links to true and false information. For each found Twitter user they collected the 3200 most recent tweets and for each Weibo user the 2000 most recent posts. The writings of each user are concatenated and used as input to linguistic analysis models like N-grams and LIWC[1]. As a result of the study, they deducted that vulnerable people to misinformation tend to use a lot of political hashtags and causal terms (for example: because, that's why, etc) on their social media accounts. While resistant users use uncertainty emojis more, adverbs as well as words referring to false information (for example: disinformation, misleading, etc). In addition, Mu et al. [MNA22] proposed a hierarchical transformer to predict whether the users are robust or not to misinformation. The model gives an F1 score of 0.851 with the Weibo dataset and 0.802 with the Twitter one.

Teng et al. [TLC+22] showed that the sequence of the user's shared links domains is determinative of whether the individual will share misinformation or not. To do this task, they adapted a medical model predicting if a patient will visit his doctor the following week using transfer learning. They constructed a merged list from multiple lists of information sources links from Media Bias/Fact Checkers and used Twitter API to collect the users tweeting them and their friends in addition to each one's recent tweets. They constructed for each user an exposer sequence representing the links shared by their friends and a share sequence representing the links they did share. These sequences are the inputs of the model which gave an accuracy of 0.72 and an F1 score of 0.737.

Regarding group-level characteristics, Chang et al. [ISO+21] used CrowdTangle to collect data about Facebook groups sharing misinformation links and build different regressors and classifiers. The

[1]Linguistic Inquiry and Word Count

best model gave an F1 score of 0.86107. Chang deducted that the more the interactions (love, sad, like, unlike, etc) of a group are diverse, the more the group is robust to misinformation. Since Facebook groups have become one of the main ways to spread news/information [Ong18]. We believe it is important to understand more deeply what makes some Facebook groups resilient to misinformation and not others.

Measurement methodology The measurement methodology consists in (1) aggregating a list of sources of misinformation. That is, either domains that post misinformation repeatedly or precise URLs with misinformation content; and (2) based on these lists, aggregate a list of Facebook groups in which misinformation links were posted.

• *Misinformation sources* - For this, we relied on information provided by fact-checkers. There are multiple fact-checking organizations across the world such as Politifact, AfricaCheck, and Leadstories. Our idea was to collect the most recent fact-checking articles and extract, using an HTML parser, the links to the sources of misinformation. We collected 20 052 fact-checked articles from Politifact. A different method is collecting data from the Google Fact Checking API [GFC] that proposes a programmatic way to get access to a large pool of fact-checked articles. The API does not provide an easy way to gather all fact-checking articles in the database, however, it allows users to retrieve all the fact-checking articles of a particular fact-checker. Hence, we query the API with all the fact-checking organizations we grabbed from the IFCN site [ifc]. In total, we collected information from 69 830 fact-checked claims. So far, we extracted 4 301 misinformation links.

• *Facebook Groups* - Meta provides an API named CrowdTangle through which one can retrieve public groups and pages in which a particular link was posted. The purpose of CrowdTangle is to help organizations monitor the non-private content being shared on social media. Access to CrowdTangle is limited and invitation based, and our group obtained such invitation. The API also allows to find popular posts in real-time and collect data about the reactions and comments of users on publications. Hence, using the API endpoint *link* we collected data about the publication of Facebook groups sharing the misinformation links collected. For example, we collected the identifier of each group posting each link, the number of reactions to the publication, the date of publication, and much other information. Next, using another endpoint, *leaderboard*, we collected data about the groups sharing these links: the number of followers, the number of each type of publications shared within the group, the total number of reactions to videos/images, etc. So far, we collected information about 8 207 unique public groups that had a post with a misinformation link.

Note that this data collection is ongoing as CrowdTangle has severe rate limits.

Hypothesis Given the problem, we would like to test the following hypothesis, and we hope to refine these ideas with feedback at the workshop.

• Similar to Teng et al. [TLC+22], we would like to test whether the sharing sequence of information sources links may determine the group's resilience to misinformation.

• [AJJ19] showed that the age of the individual impacts his resistance to misinformation. We plan to investigate if this factor is also important at group-level. The challenge here is to obtain the approximate average age of users in a particular group.

• Groups on Facebook can vary from communities for mothers to advocacy groups against the current government. Hence, we plan to study to which extent the purpose of the group has an impact on the resilience to misinformation.

• We plan to study what is the impact of the political leaning of the group. The challenge here is how to collect such data. Our idea is to join several of these groups and propose to them to participate in a survey. Another possibility is to analyze the links shared in the group, and check how they are classified in the Media Bias Fact Check database. This database is maintained by journalists and contains information about the political leaning of newspapers and the degree of factualness of their content.

While the number of hypotheses we can test is easily extendable, one more fundamental question this research brings is *how to measure the resilience of a group to misinformation?* We, of course, would welcome feedback from workshop participants.

References

[AJJ19] Andrew Guess A, Nagler J, and Tucker J. Less than you think: Prevalence and predictors of fake news dissemination on facebook. *Science advances, 5(1), eaau4586*, 2019.

[GFC] Fact check tools api from: https://developers.google.com/fact-check/tools/api?hl=fr.

[ifc] The international fact-checking network from: https://www.poynter.org/ifcn/.

[ISO+21] CHANG I, SUN, Orion, AHN, Jasper Sang, and et al. Is social diversity related to misinformation resistance? an empirical study on social communities. *IEEE Global Humanitarian Technology Conference*, 2021.

[MNA22] Y Mu, P Niu, and N Aletras. Identifying and characterizing active citizens who refute misinformation in social media. *arXiv preprint arXiv:2204.10080*, 2022.

[Ong18] Cabañes J. V. A. Ong, J. C. Architects of networked disinformation: Behind the scenes of troll accounts and fake news production in the philippines. architects of networked disinformation: Behind the scenes of troll accounts and fake news production in the philippines. 2018.

[TLC+22] X Teng, Y. R Lin, W. T Chung, A Li, and A. Kovashka. Characterizing user susceptibility to covid-19 misinformation on twitter. *In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 16, pp. 1005-1016).*, 2022.