# Analysis of the (Mis)information Dissemination Network in WhatsApp

Gabriel P. Nobre
Jussara M. Almeida
{gabrielnobre,jussara}@dcc.ufmg.br
Universidade Federal de Minas Gerais
Belo Horizonte, Brazil

## Abstract

In this thesis, we aim at investigating the spread of information, notably misinformation, on WhatsApp. We show how WhatsApp enables the formation of a rich underlying network that, while crossing the boundaries of existing groups, have structural properties that favor information dissemination at large. Next, we delve deeper into this network to identify user communities, i.e., groupings of users who, intentionally or not, share the same content disproportionately often and characterize how such components could affect information dissemination. By analyzing content sharing and network topological properties focusing on three perspectives (individuals, WhatsApp groups, user communities), our study offers valuable insights into how WhatsApp users leverage the underlying network connecting different groups to gain large reach in the spread of misinformation in the platform.

***Keywords:*** Misinformation, WhatsApp, User Behavior, Network Analysis

## 1 Introduction

A topic that has attracted large interest in recent years is the study of (mis)information dissemination, especially in social media applications, exploring the underlying networks that connect their users [15, 16, 18]. Indeed, it has been well argued by [6] that online social networks are widely used as tools for misinformation spread at large, with important negative impact on the real world [1, 12].

In such context, WhatsApp emerged as an important platform for information dissemination, reaching more than 2 billion users worldwide in 2020 [1]. Though originally designed for exchanging messages, researchers have shown that it enables the formation of rich underlying networks with structural properties similar to those observed in popular online social networks [16]. These networks are fostered especially by the so called WhatsApp groups, which though limited to only 256 simultaneous members, may have members in common who act as bridges through which information can cross group borders and spread through the environment at large.

A few recent studies have hinted at the potential for information virality by taking a bird's eye view of the network that emerges from (mis)information exchange in different groups [3, 11, 12, 15, 16]. However, an investigation of the properties of this network, which may reveal an underlying structure that facilitates information dissemination is still lacking. In this thesis, we take a step towards filling this gap by analyzing (mis)information dissemination in WhatsApp *from a network perspective*. In particular, we are interested in investigating to which extent users in different groups, intentionally or not, build *communities* of content spread, which consistently help speeding up information dissemination within the platform. In order to achieve it, we build a sequence of *media-centric networks* so as to capture the relationships established among users who shared the same piece of content (text, image, audio or video) in one or more groups during pre-defined time windows.

Specifically, we tackle the following research questions:

- **RQ1**: Are there underlying structural user connections that favor information dissemination under the WhatsApp platform?
- **RQ2**: What is the role that different user groupings have in the spread of misinformation in the platform?

For **RQ1**, we will investigate the existence of underlying user communities and if such structures could act as bridges that spread content beyond the WhatsApp group boundaries - facilitating information dissemination in the platform. In its turn, for **RQ2**, we will investigate and characterize the role that hierarchical components of the media-centric WhatsApp networks may play in content dissemination, notably misinformation.

Towards tackling such goals we have performed several analyses, summarized in the following key steps. First, we constructed subsequent media-centric networks, extracted the backbones and then we investigated underlying user communities emerged and we assessed if such user groupings extrapolate the WhatsApp groups' boundaries. We highlight

[1] https://www.whatsapp.com/about/

In: IV Concurso de Teses e Dissertações (CTD 2022), Curitiba, Brasil. Anais Estendidos do Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia). Porto Alegre: Sociedade Brasileira de Computação, 2022.

ISSN 2596-1683

that, in our study, this underlying structure, namely the *backbone*, represents a subset of the network users with higher indication of coordinated efforts for content dissemination.

Next, we performed a hierarchical network-oriented characterization of user engagement in misinformation spread from three perspectives: individual users, WhatsApp groups and user communities. Furthermore, we characterized the roles of the network periphery and backbone in terms of misinformation content sharing and finally, we executed a high level analysis of the temporal dynamics of the network.

To perform our investigation we used a dataset of WhatsApp messages gathered by [16]. For the misinformation analysis, we used a previously built labeled dataset for images [16], text [15] and audios [7]. In summary, the dataset covers six weeks around the 2018 general elections in Brazil ($1^{st}$ and $2^{nd}$ rounds in October $7^{th}$ and $28^{th}$, respectively), ranging from September $17^{th}$ until October $28^{th}$ in 2018. It contains more than 100 thousand messages shared by around 5 thousand users per week on a selection of 155 groups.

The remainder of this paper is organized as follows. We present our main contributions in Section 2, our related work in Section 3 and the methodology followed during the analysis in Section 4. Next, we present our main results in Section 5, followed by our overall conclusion in Section 6.

## 2 Contributions

We summarize our main findings as follow. We observed strongly connected user communities that cross WhatsApp group boundaries. Ultimately, these communities may serve as an indication of coordinated effort to spread the same pieces of content. Next, we noted that the restrictions imposed by the platform on group size offer little constraints in terms of content propagation throughout the platform, reaching multiple groups and users. Furthermore, we performed a hierarchical characterization of misinformation dissemination, whereas we observed that a few of the users are responsible for a large fraction of misinformation shared per week and also for introducing more fresh content. This was also observed at the group and community levels, while the high misinformation sharing groupings tend to have more members and thus reach a larger potential audience. Another finding is that the network backbone often includes the top misinformation sources while the network periphery demonstrated a significant impact on the total volume shared in the network. Finally, by analyzing the temporal evolution of the network, we identified that the subset of users who share misinformation is highly dynamic, varying a lot over time. Yet, we noted that the subset of WhatsApp groups in where the larger volume of the misinformation content was shared tended to remain the same from week to week.

Our results are summarized in two papers [9, 10].

## 3 Related Work

WhatsApp attracted a large number of users in many countries, naturally followed by a number of studies [2, 4, 7, 15] that rely on the existence of a large number of *publicly accessible WhatsApp groups*. Some researchers developed automatic tools to expose, in an anonymized fashion, the content being shared in WhatsApp groups [2, 4, 7, 13, 17].

Moreover, a number of studies analyzed the content and propagation properties of messages containing content previously checked as fake by fact checking agencies, comparing them against the properties of other (unchecked) messages [3, 15, 16]. For example, the authors in [16] characterized the dissemination of images containing information priorly checked as fake, noticing that fake images spread much faster, reaching a much larger user population, compared to other content. More recent studies followed a similar path but focused on textual [15] and audio [7] content. In a complementary direction, other work focused their efforts to study and experiment with possible strategies to mitigate the spread of misinformation on WhatsApp [8, 14].

The spread of misinformation has also raised concerned on other social media platforms. For instance, Facebook and Twitter have reportedly being abused for misinformation spread since Unites States 2016 presidential elections [5]. Since then, both platforms applied distinct action plans to try to mitigate the problem [1, 5].

In this thesis we complement prior analyses by performing a network-driven analysis information dissemination in WhatsApp while considering a multimedia content sharing perspective. To achieve it we carried out a hierarchical characterization of the role of different network components, notably user communities and backbone-periphery, with a posterior focus in misinformation content spread. With this novel perspective, we identified key aspects of the underlying content dissemination network in WhatsApp that, in the end, suggest a facilitated environment for misinformation dissemination.

## 4 Methodology

Towards characterizing user participation in misinformation spread on WhatsApp at different levels of aggregation, we adopted a methodology consisting of 6 steps. Our starting point (Step 1) is a dataset of messages shared in a number of publicly accessible WhatsApp groups. Second, the dataset is expanded with the labeling of a sample of messages containing text, audio and image content, as *misinformation* (Step 2). Third, we defined and constructed a sequence of media-centric networks by connecting users who shared similar content during given time intervals. We built multiple networks, enabling the study of (mis)information spread during different (successive) *snapshots* of the platform (Step 3). Then, for each network, we identified important components to the study of information dissemination, notably the network

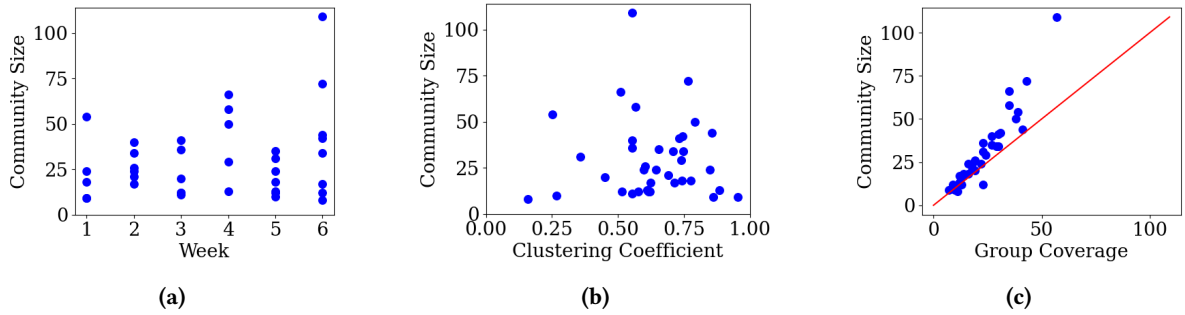



**Figure 1.** Communities (a) Size, (b) Average Clustering Coefficient, (c) Group Coverage.

*backbone* and *periphery* (Step 4), as well as the *communities* of nodes composing the backbone (Step 5). Finally, we analyzed the participation of users, groups and communities in misinformation spread, focusing on both sharing properties and topological characteristics of these entities with respect to the defined networks and their main components (Step 6).

## 5 Results

Now, we summarize our results in two sections. First, we describe our results regarding the observation and characterization of WhatsApp communities in Section 5.1. Next, we focus on the hierarchical analysis of misinformation content spread in Section 5.2.

### 5.1 Unveiling and Characterizing Communities

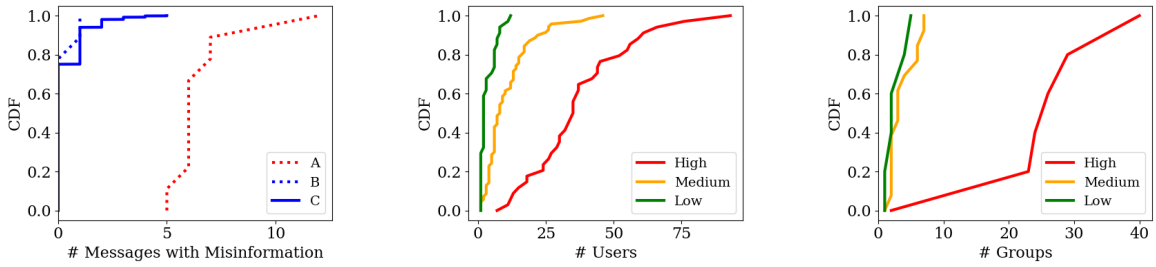
We first focus on the properties of the media co-sharing networks. We observed a fairly stable number of around 2,000 of users over the weeks. Nevertheless, average node degree was quite high (from 27 to 54), indicating that, on average, a user shares similar content with many other users. After that, we extracted the network backbones that consisted on from 114 to 338 users, and we noted that they have a reasonably strong average clustering coefficient (ranging from 0.49 to 0.63) as well as large average degree (6.07 to 20.7). These measures suggest well connected topologies and also hint at the formation of communities.

Figure 1 provides an overview of different community properties for all communities in all snapshots. Figure 1a shows the sizes (in number of nodes) of the communities identified in each snapshot (week), with each point representing a different community. We observe a great diversity of community sizes, suggesting that the communities are dynamically built over time with variable number of users. We also correlate community size with average clustering coefficient, which is a measure of internal connectivity. In Figure 1b we observe that most communities are strongly connected (even the larger ones), as the vast majority of them have average clustering coefficient above 0.50. Thus, in essence, the identified user communities are well structured, and, offer clear indications of consistent user co-sharing activity. In its turn, Figure 1c shows a scatter plot with community size versus group coverage. There is a strong one-to-one relationship, in which the community size is strongly correlated with the number of groups it reaches. For larger communities, the sizes often are greater than the number of groups covered. These results suggest a broad reach in the ability to disseminate information, since communities often have members participating in multiple groups during the same time period.

### 5.2 Hierarchical Analysis in Misinformation Spread

Now we outline the results for the hierarchical characterization of misinformation spread on WhatsApp in different level of user aggregation for a single snapshot that refers to one week in our analysis. We briefly summarize this characterization in Figure 2. For the user level, we display in Figure 2a that the top-10 misinformation sharing users (curve A) tend to share more misinformation content then the top-10 message sharing users (curve B) and the the other users (curve C). Furthermore, we illustrate in Figures 2b and 2c that the high misinformation sharing groups and communities (top 25% with highest level of misinformation sharing in red lines) indeed holds a potentially larger audience.

Focusing on the backbone-periphery characterization of the media co-sharing networks, we noted that often around half of all users in the backbone shared some misinformation. Furthermore, despite covering only around 15% of all users in the network, the backbone nodes, who include most top misinformation sources, are responsible for a great part of all misinformation spread on the platform, with a clear dominance in terms of content diversity. The strong co-sharing behavior linking these users in the backbone suggests the possibility of coordinated effort to gain audience scale. Regarding the periphery characterization, we noted a large presence of users engaged in misinformation spread, though with a lesser degree. Indeed, representing the other 85% of the users in the network, they are responsible for a large fraction (up to 68%) of all misinformation shared in the period. Considering that the periphery mainly consists on users that are less likely to be engaged in coordinating efforts, we




(a) (b) (c)

**Figure 2.** Hierarchical misinformation spread as of (a) users, (b) WhatsApp groups, (c) communities.

highlight a considerable probability of having many of them sharing misinformation while being unaware of it.

As of our temporal analysis, we observed that the individual users mostly engaged in misinformation spread change drastically over time. Yet, they are frequently members of the same groups.

## 6 Conclusion

We presented a multi-level characterization of (mis) information dissemination in WhatsApp, with a particular focus on the potential role of the media-centric network as a means to gain content visibility and scale of dissemination. Our analysis relied on dataset of multi-media messages shared in publicly accessible WhatsApp groups in Brazil.

Our insights are novel and greatly complement the current literature on WhatsApp and closely similar platforms (e.g., Telegram). By looking at the problem from this perspective, we offer a new look into (mis)information spread on a widely used application that servers as main means of communication in many countries of the world.

## References

[1] Hunt Allcott, Matthew Gentzkow, and Chuan Yu. 2019. Trends in the diffusion of misinformation on social media. *Research & Politics* (2019).

[2] Victor S. Bursztyn and Larry Birnbaum. 2019. Thousands of Small, Constant Rallies: A Large-Scale Analysis of Partisan WhatsApp Groups. In *2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. 484–488.

[3] Josemar Alves Caetano, Gabriel Magno, Marcos André Gonçalves, Jussara M. Almeida, Humberto Torres Marques-Neto, and Virgílio A. F. Almeida. 2019. Characterizing Attention Cascades in WhatsApp Groups. In *Proceedings of the 10th ACM Conference on Web Science*. 27–36.

[4] Kiran Garimella and Gareth Tyson. 2018. WhatApp Doc? A First Look at WhatsApp Public Group Data. *Proceedings of the International AAAI Conference on Web and Social Media* (2018).

[5] Christine Geeng, Savanna Yee, and Franziska Roesner. 2020. *Fake News on Facebook and Twitter: Investigating How People (Don't) Investigate.* Association for Computing Machinery, 1–14.

[6] Srijan Kumar and Neil Shah. 2018. False Information on Web and Social Media: A Survey. *Social Media Analytics: Advances and Applications* (2018).

[7] Alexandre Maros, Jussara Almeida, Fabricio Benevenuto, and Marisa Vasconcelos. 2020. Analyzing the use of audio messages in Whatsapp groups. In *Proceedings of The Web Conference 2020*. 3005–3011.

[8] Philipe Melo, Carolina Coimbra Vieira, Kiran Garimella, Pedro OS Vaz de Melo, and Fabrício Benevenuto. 2019. Can WhatsApp counter misinformation by limiting message forwarding?. In *International conference on complex networks and their applications*. Springer, 372–384.

[9] Gabriel P Nobre, Carlos HG Ferreira, and Jussara M Almeida. 2020. Beyond Groups: Uncovering Dynamic Communities on the WhatsApp Network of Information Dissemination. In *International Conference on Social Informatics*. Springer, 252–266.

[10] Gabriel P Nobre, Carlos HG Ferreira, and Jussara M Almeida. 2022. A hierarchical network-oriented analysis of user participation in misinformation spread on WhatsApp. *Information Processing & Management* 59, 1 (2022).

[11] Elaine O Nsoesie and Olubusola Oladeji. 2020. Identifying patterns to prevent the spread of misinformation during epidemics. *The Harvard Kennedy School Misinformation Review* (2020).

[12] Gordon Pennycook, Jonathon McPhetres, Yunhao Zhang, Jackson G. Lu, and David G. Rand. 2020. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science* (2020), 770–780.

[13] Raquel Recuero, Felipe Soares, and Otávio Vinhas. 2020. Discursive strategies for disinformation on WhatsApp and Twitter during the 2018 Brazilian presidential election. *First Monday* (2020).

[14] J. C. S. Reis, P. Melo, K. Garimella, and F. Benevenuto. 2020. Can WhatsApp benefit from debunked fact-checked stories to reduce misinformation?. In *Harvard Kennedy School (HKS) Misinformation Review*.

[15] Gustavo Resende, Philipe Melo, Julio C. S. Reis, Marisa Vasconcelos, Jussara M. Almeida, and Fabrício Benevenuto. 2019. Analyzing Textual (Mis)Information Shared in WhatsApp Groups. In *Proceedings of the 10th ACM Conference on Web Science*. 225–234.

[16] Gustavo Resende, Philipe F. Melo, Hugo Sousa, Johnnatan Messias, Marisa Vasconcelos, Jussara M. Almeida, and Fabrício Benevenuto. 2019. (Mis)Information Dissemination in WhatsApp: Gathering, Analyzing and Countermeasures. In *The World Wide Web Conference (WWW '19)*. ACM, 818–828.

[17] Marisa Vasconcelos, Erica Pereira, Samuel Guimarães, Manoel Horta Ribeiro, Philipe Melo, and Fabrício Benevenuto. 2020. Analyzing YouTube Videos Shared on Whatsapp in the Early COVID-19 Crisis. In *Proceedings of the Brazilian Symposium on Multimedia and the Web (WebMedia '20)*. 25–28.

[18] Narisa Zhao and Hui Li. 2020. How can social commerce be boosted? The impact of consumer behaviors on the information dissemination mechanism in a social commerce network. *Electronic Commerce Research* (2020).