Online Polarization in Times of War: a Reddit Network Analysis of the Israel-Palestine Conflict

Journal of Computational Social Science

Luca D'Ambrosio, Filippo De Min†, Caterina Ruzzi†, Federico Ungarelli†

Contributing authors: luca.dambrosio@studbocconi.it; filippo.demin@studbocconi.it; caterina.ruzzi@studbocconi.it; federico.ungarelli@studbocconi.it;
†These authors contributed equally to this work.

Abstract

The Israeli-Palestinian conflict is one of the longest-running territorial conflicts still ongoing in the world. Tensions in the area date back to the first half of the 20^{th} century with the migrations of Jews from Europe and Yemen to Palestine and the first clashes between the two groups after Israel's declaration of statehood. This culminated in the First Intifada and the massive protests that erupted in the Gaza Strip, the West Bank, and Israel among Palestinian-Israelis. On October 7^{th}, a violent attack by Hamas triggered the start of a harrowing conflict that persists to the present day. As the war evolves, controversies and polarization seem to increase among supporters of each faction, and we investigate whether online platforms might exacerbate these differences through network analysis and natural language processing techniques. By studying online interactions in the r/IsraelPalestine subreddit, we find that polarization seems to actually decrease in the months following the start of the conflict, suggesting that moderated online forums might actually facilitate dialogue and reduce divisiveness.

1 Introduction

Online polarization has emerged as a critical issue in the digital age, significantly influencing public discourse. This study delves into the online polarization surrounding the Israel-Palestine conflict, utilizing Reddit data from the r/IsraelPalestine subreddit to explore the dynamics of user interactions and community formation. By leveraging a comprehensive network analysis approach, this research aims to uncover the intricate patterns of partisanship and topic discussion that shape online debates in this contentious context.

The primary objective of this study is to understand how online polarization manifests and evolves in response to significant events in the Israel-Palestine conflict. To achieve this, we employed several methodological steps. First, we collected Reddit interactions from the r/IsraelPalestine subreddit to create a detailed network of user interactions. We then analyzed changes in network characteristics before and after key conflict events to understand how these events impact online discourse. Following this analysis, we used a Natural Language Processing classifier to determine the political leanings of users, which allowed us to better identify possible changes in polarization within the network. Furthermore, communities within the network were identified over time to study their features. Lastly, we analyzed the topics discussed within each community to gain insights into the central themes and to understand how these themes influenced interactions between the communities.

The primary research question guiding this study is whether online platforms exacerbate polarizing dynamics after the events that led to the Israel-Palestine conflict. This is particularly important because it addresses the role of social media in shaping public discourse during periods of tension.

Previous studies have highlighted the formation of echo chambers and the spread of misinformation on social media platforms, contributing to online polarization. However, there is a gap regarding how these dynamics work in the context of prolonged and high-stakes conflicts like the Israel-Palestine situation.

2 Literature Review

Online polarization has become a critical focus in social media research. Polarization online often manifests through the formation of echo chambers, where individuals predominantly engage with information that reinforces their existing beliefs, thus deepening ideological divides (Conover et al (2021)). Social media enables individuals to connect with like-minded peers, facilitating the exchange of ideas and the building of solidarity. The ease with which people can select their information sources and form segmented groups has been widely criticized in the political arena, with some suggesting it directly contributes to polarization (Bennett and Iyengar (2008)). However, the impact of online polarization is debated, with some arguing that its effects are overstated (Boxell et al (2017)) or even suggesting that it can reduce polarization (Barberá (2014)).

Polarization is particularly evident in conflict and contentious politics, as social media platforms like Reddit serve as arenas for intense ideological battles, facilitating the rapid spread of misinformation and the entrenchment of polarized communities (Zeitzoff (2017)). Network analysis, a methodological tool used to map interactions and identify clusters within online discourse, has proven effective in examining these dynamics (Himelboim et al (2017)). For instance, research on Twitter topic networks has demonstrated how polarized crowds and community clusters form around contentious issues, providing insights applicable to other platforms such as Reddit (Smith

et al (2014)). Specifically, studies on political controversy in Israel have shown that discussions on Twitter become highly polarized, with distinct communities emerging (Yarchi et al (2020)).
To detect communities within the network, we utilized the Louvain algorithm, which is known for its efficiency in uncovering hierarchical community structures in large networks (Blondel et al (2008)).
Furthermore, LDA has been successfully applied to social media text analysis in the past, as seen in a study by Melton et al (2021), where LDA was used to examine public sentiment and opinions regarding COVID-19 vaccines on Reddit.

3 Data Collection

3.1 Data Collection and Pre-Processing

Concerning our data source, we acquired every monthly Reddit dump from the Academic Torrents platform, from July 1st 2023 to April 2024 (totaling 10 months of data). Each downloaded dataset comprises the majority of Reddit submission posted by users on large subreddits during the respective month, along with post metadata (username of the redditor, submission id etc.). The data has then been filtered and decompressed to allow the extraction of data pertaining to our subreddits of interest ("r/Palestine, r/Israel, r/IsraelPalestine), through an ad-hoc python script sourced from the PushshiftDumps Github repository (Watchful1 (2024)). These monthly batches were then merged vertically into a single dataset, from which we removed posts that were either deleted by the user, or removed by moderators. Moreover, we removed posts that contained either a single URL or an image/poll, as they would not be very useful when analyzing online discourse. After the above procedure, the resulting dataset encompasses 14,039 total posts across the three subreddits, with 6667 posts for r/IsraelPalestine, 5960 posts for r/Israel, and 1412 posts for r/Palestine. As a final step, we separate submissions to r/IsraelPalestine from the other two subreddits; the latter will constitute an independent dataset, that will allow us to train the political classifier reported in section (4.2.1).

3.2 Network Data

In order to perform a network based analysis in the r/IsraelPalestine subreddit, we needed to gather additional interaction data for the users in our dataset, hence requiring to obtain all comments under each submission, together with the users who posted them. To achieve this, we referred to the Python Reddit API wrapper (PRAW) which facilitates interaction with the Reddit platform and accelerates data acquisition. Using the post-ids we extracted in section (3.1), we were able to collect the Comment Forest object pertaining to each submission in our dataset; this step demanded that the submissions were split into 6 smaller batches in order for the API requests to succeed. The data structure we retrieved contains all comments under a post, while preserving the tree-structure of the Reddit comment section, thus allowing for distinctions between different layers of the latter. In other words, this enabled us to know whether a comment was in response to the original post, or if it responded

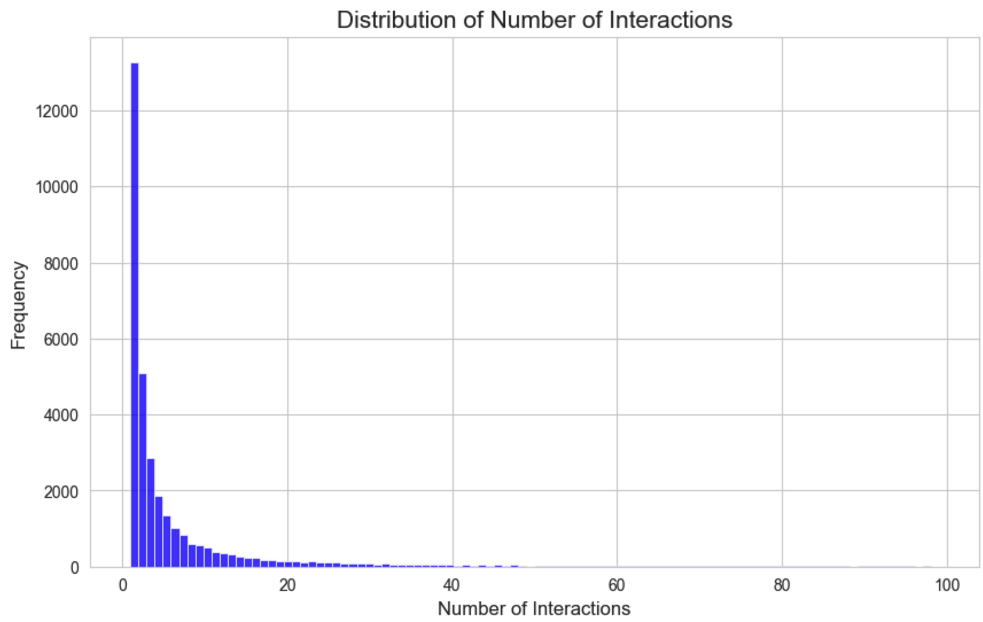

Fig. 1 Distribution of the number of interactions across users. We do not show users who had more than 100 interactions, who are sparse and isolated along the tail of the distribution

to other comments under a post, and if so, how deep in the forest it was located. This feature is fundamental as we model each interaction between user A and user B as a direct comment by user B under a comment/post by user A, so if a third user C replied to user B there will be no interaction between user C and user A.
By recursively iterating through each comment forest, we were able to construct an ordered list of comments labeled with their depth level, which was then processed in order to obtain every relevant interaction that occurred between July 2023 and April 2024 on r/IsraelPalestine (following the logic described above). The resulting dataset contains 599,199 total interactions, each of them constituting a statistical unit. Lastly, for each interaction (in addition to the depth level) we collect the corpus, the username and the posting date both for the comment/post and the reply. Note that some posts/comments that were collected in 2023, have been later deleted or removed (roughly 80,000 interactions contain a missing comment/post), thus causing some missing values in our dataset. We managed to impute some of the missing posts by retrieving the text and author username from the original dataset (3.1), while the remaining interactions have been dropped from the final dataset.

Finally, the users will be the nodes of the graphs we build in subsequent sections, with their interactions representing the edges/connections between nodes. Each edge is then weighted by the number of interactions between the two nodes (users). As expected, the number of interactions (degree) of each user follows a power-law distribution.

4 Methodology

4.1 Evolution of the Network of Top Authors

To provide a complete analysis of the network of top authors and its evolution over time, it is essential to first define what is meant by a "top author." In this study, we define a "top author" as an author who has the highest number of interactions within the dataset. Authors are ranked based on how often their content appeared in the data within a specific period, both in the set of authors who replied and the one containing the authors' comments. The results are combined, leading to a list of approximately 200 top authors for each network (this methodology applies to every section apart from section 4.2).

Once the top authors are identified, we aggregate them to build a network for each period in an effort of assessing how the dynamics among top authors evolved over time. In particular, this involves mapping out the connections between these top authors, based on their interactions. To understand the structural properties of these networks and how they evolve over time, we analyze them using average degree, density, average clustering, average path length, and diameter measures. On top of that, centrality measures, such as degree centrality, betweenness centrality, and closeness centrality, are calculated for each top author to assess their importance and influence within the network.

It is worth mentioning that to investigate how the network of top authors has changed over time, we divide the analysis into one-month segments beginning from July 1, 2023, to April 30, 2024, extending the period of September to October 6, to separate the networks creating from the start of the war.

4.2 Polarization of Discourse Among Most Central Users

Attempting to unveil patterns of polarization within our data, we consider the 600 top users by interaction count over the entire period (most active in the community) and construct a sub-population. These authors will then be classified according to their political beliefs (either Pro-Palestine or Pro-israel) by fine-tuning an appropriate classifier. This choice of methodology stems from constrained computational resources and an attempt to simplify the interpretation of the results.

4.2.1 Political Classifier

We finetuned a roBERTa transformer model (Conneau et al (2019)), aimed at classifying text as either pro-Israel or pro-Palestine, by using posts from the r/Israel and r/Palestine subreddits as our training samples, interpreting the subreddit on which a post was submitted as its label. One inherent limitation of this approach is the clear class imbalance that emerges from the data (5960 posts for r/Israel versus 1412 posts for r/Palestine). This concern could not be addressed due to the lack of additional labeled pro-Palestine content on the platform. Although the classifier exhibited a clear bias towards classifying text as pro-Israel, attributable to the aforementioned class imbalance, it achieved an f-1 score of roughly 70% on the evaluation dataset. The final version of the model was trained on the entire dataset.

We then selected our population of interest and collected a random sample of 20 posts for each user, classified it and obtained a soft-max score for each post (which we interpret as a level of confidence in the prediction). This score attributes a value between 0 and 1, p to one predicted label and value $1 - p$ to the other, which we interpret as a confidence level in the prediction or the probability that the post/comment's ideas align with the respective faction. These scores (20 per user) are averaged across users, thus obtaining a "probability" of each user standing for Israel/Palestine.
As a last step, due to the previously mentioned bias in Pro-israel classification, we set a threshold of $Pro - Palestine - Probability = 1 - p = 0.25$, above which we classify a user as Pro-Palestine; this value was validated through a manual annotation of a portion of the data.

4.2.2 Evolution of Interactions Between Opposite Factions

Once our sub-population has been defined and classified, we build a bi-weekly undirected network connecting the users according to their respective interactions, with each interaction representing an edge. For each user, we then define an attribute "faction" which contains the information regarding their beliefs (Palestine or Israel). Moreover, we count the weighted number of interactions that each node had with a neighbor belonging to the opposite faction and divide it by the weighted node degree, thus obtaining a measure of cross-faction interaction percentage for each user. The latter will constitute a metric that we interpret as dialogue and integration between the two communities, with lower values corresponding to increasing polarization, as the node will mainly interact with people who share the same views. It is important however, to interpret the evolution/change of such a metric rather than its absolute value, as Pro-israel users embody a large portion of the network which makes interactions among them more likely purely due to their prevalence. Finally, we aggregate this score across all users with a bi-weekly time step, and its evolution is studied and reported throughout the entire time span of our study.

4.3 Community Analysis

In this section, as a way to uncover polarization and social division within our network, we detect underlying community structures through the Louvain algorithm, which was chosen due to its documented performance when dealing with densely connected groups of nodes. Once these communities are identified, we apply topic modeling techniques to analyze the content and themes discussed within each group.

4.3.1 Community Detection Algorithm

In order to successfully detect communities we implement the Louvain algorithm in our dataset. This is a popular method for community detection due to its efficiency (close to linear time complexity) and better partitioning capabilities compared to competing strategies. The algorithm will identify clusters or groups of nodes that are more densely connected internally than with the rest of the network, through modularity optimization (a metric that measures the strength of the division). High modularity indicates dense connections within communities and sparse connections between them. In practice the algorithm will create a sequence of sub-graphs of "super-nodes" formed by joining nodes whose union will increase. For our analysis, we create two datasets starting from the one described in section 3:

- The first subset will contain interactions between the top users (union of the top 100 authors from the child comments and the top 100 authors from the parent comments) based on their interaction counts, spanning the months of July, August, September, and the first week of October 2023. (Before the start of the war)
- The second subset will contain the top users (with the same rationale of the previous point) in the remaining days of October, November and December 2023. (Immediate post-war).

For each dataset we run the Louvain algorithm and partition the graph into communities, hoping to uncover different patterns in the two periods, and study the polarization of the two networks. Finally, we plot the resulting partitions (storing them). As a last step, we repeat the process (using the same logic) creating communities each month for the 10 months of data in our analysis, and report the modularity evolution over this period.

4.3.2 Topic Modeling within Communities

Topic Modeling is a machine learning technique used to uncover the underlying thematic structure in a collection of documents. It assumes that documents are mixtures of topics, where each topic is a distribution over words.
Latent Dirichlet Allocation (LDA) starts by initializing the number of topics K and randomly assigning words in documents to topics. Then, it iteratively refines these assignments through Gibbs Sampling, where for each word in each document, probabilities of assigning the word to each topic are calculated based on the current assignments. These probabilities consider the proportions of words in the document assigned to each topic and the proportions of times each word has been assigned to each topic across all documents. New topic assignments for words are sampled based on these probabilities. This process is repeated multiple times until a steady state is reached, resulting in stable topic assignments. After convergence, LDA outputs the distribution of topics in each document along with the distribution of words in each topic.
The tuning parameter in a Topic Modeling is the number of topics present generating the corpus of the document, chosen with Pointwise Mutual Information (PMI) which computes the statistical significance of word co-occurrences within topics. We leveraged the BERTopic Python library to uncover topics. BERTopic (Grootendorst (2022)) is a topic modeling technique that offers significant advantages over standard approaches like LDA. It uses BERT embeddings to capture semantic relationships. In the purpose of keeping only words with strong semantic meaning, we apply basic preprocessing and keep only nouns, verbs, adjectives, and proper nouns.
We run a separate topic modeling for each of the detected communities as explained in section 4.3. For each community we extract the 5 most relevant topics together with their associated words.

Topic Modeling has two major limitations. First, the final label describing each topic must be interpreted by human intelligence from the words composing each of them. Second, this method requires the assumption that the corpus of texts formed by the posts and comments from communities users belong to a distribution of topics and such topics are described by a distribution of words composing them.

5 Results

5.1 Evolution of the Network

From July 7, 2023, to April 30, 2024, the measures describing the network of top authors evolved as depicted by Fig. 2. For clarity, the networks in this section are indexed by the periods they represent:

- *Network 1*: July 1, 2023, to July 31, 2023
- *Network 2*: August 1, 2023, to August 31, 2023
- *Network 3*: September 1, 2023, to October 6, 2023
- *Network 4*: October 7, 2023, to October 31, 2023
- *Network 5*: November 1, 2023, to November 30, 2023
- *Network 6*: December 1, 2023, to December 31, 2023
- *Network 7*: January 1, 2024, to January 31, 2024
- *Network 8*: February 1, 2024, to February 29, 2024
- *Network 9*: March 1, 2024, to March 31, 2024
- *Network 10*: April 1, 2024, to April 30, 2024

Average Degree Centrality: over time, the average degree of the network under analysis changes significantly. It increased significantly from 18.23 in Network 1 to a peak of 29.56 in Network 5, indicating periods of high connectivity among authors. The degree then starts to decline until it reaches stable levels ranging from 18.35 to 21.24, which are, however, higher compared to pre-war levels.
Density: density measures the ratio between the actual connections in the network in ratio to the maximum number of possible connections. Starting with a value of 0.145 in Network 1, it peaked at 0.226 in Network 5. However, following the month of November, the density of the networks is characterized by a sharp decline, reaching a value close to pre-war level, 0.138, in Network 9, suggesting a reduction in interaction intensity.
Average Clustering Coefficient: the clustering coefficient represents the degree to which nodes in a graph tend to cluster together. Interestingly, this measure never reached the first Network level (0.31). However, analyzing only the post-war period, we would find, once again, a peak at Network 5 (0.29), followed by a constant decline in the next periods, indicating that in the underlying period discussions are not only more frequent but also more focused within specific groups or communities compared to the rest of the moments.
Average Path Length and Diameter: average path length measures the average number of steps along the shortest paths for all possible pairs of nodes, while the diameter represents the longest shortest path in the network. Hence, shorter average path lengths and smaller diameters during significant events imply that information spreads more

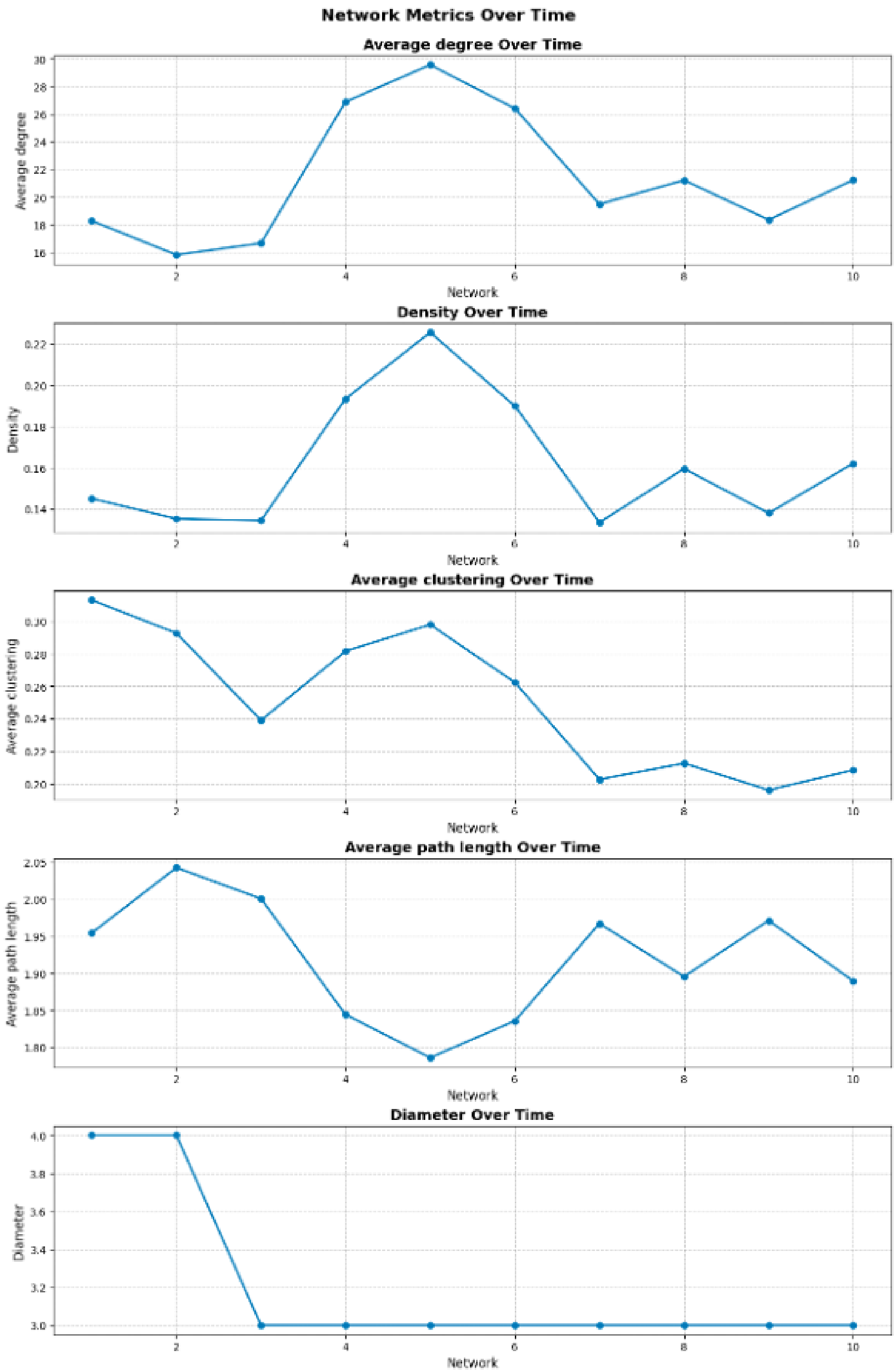

Fig. 2 Description Measures of Network Over Time

quickly and efficiently across the network. Both metrics were lowest in Network 5 (1.79 for average path length and 3 for diameter), suggesting efficient communication and rapid information spread during periods of intense activity.
In summary, there is compelling evidence of increased cooperation among ”top authors” in Network 5. All of the network's descriptive measures peaked during this period, likely due to high levels of connection among them. In fact, a higher average degree implies greater density and aggregation among units. This results in a decrease in the graph's diameter and the maximum distance between any two nodes.

5.2 Polarization of Discourse

By implementing the methodology described in section 4.2.2, we manage to extract some valuable insights on the levels of polarization (or lack thereof) in the network. In particular, we find that the top 100 users in the network reflect larger levels of interaction with authors supporting the opposite faction after the events occured on October 7th. In fact, the average percentage of cross-faction interactions increases by roughly 10 percentage points, compared to pre-war levels. This growth persists in following months, as reported in Fig. 3. This finding solidifies the idea that online platforms such as Reddit are indeed promoting dialogue instead of polarization.

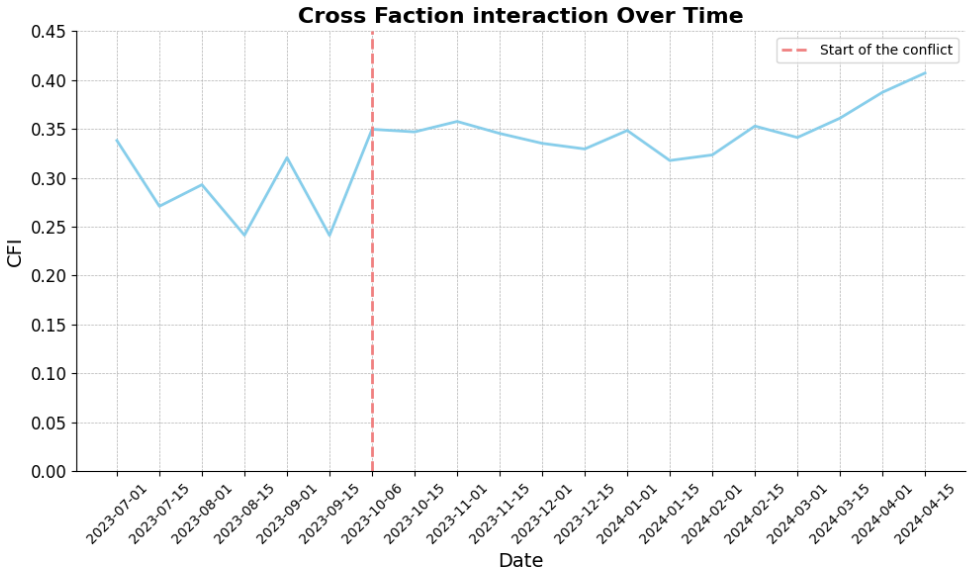

Fig. 3 Evolution of the average percentage of interactions among users with opposite views about the conflict.

5.3 Community Analysis

By analizying the modularity in the network each month, we can observe a small, albeit evident pattern. As, shown in Fig. 4, the levels of modularity seem to decrease (23% decrease compared to pre-war levels) during the immediate aftermath of the war, suggesting a higher concentration and connectedness of the network, likely related to increased levels of dialogue and communication. We interpret this trend as a signal of decreasing polarization in the subreddit. Communication between users involves two parallel processes: first, two users need to interact with each other, and second, when they do interact, there must be mutual willingness to listen to the other's arguments and reply in a constructive manner, which adds valuable insights to the dialogue. Due to the subreddit being highly moderated, we assume that the detected interactions must reflect these features, which leads us to interpret the decrease in modularity as an increase in dialogue and discussion. In the following months the modularity seems to regress to the old value, suggesting that the top authors became more divided as the conflict progressed.

Table 1 and Table 2 display the topics detected from the algorithm for each of the communities pre- and post-October 7th events. From the series of words that are best characterizing each topic (Representation), we arbitrarily decide its label (interpretation), ex-ante crossing the results with the network analysis. The order of importance of each topic together with the number of documents belonging to each topic are also

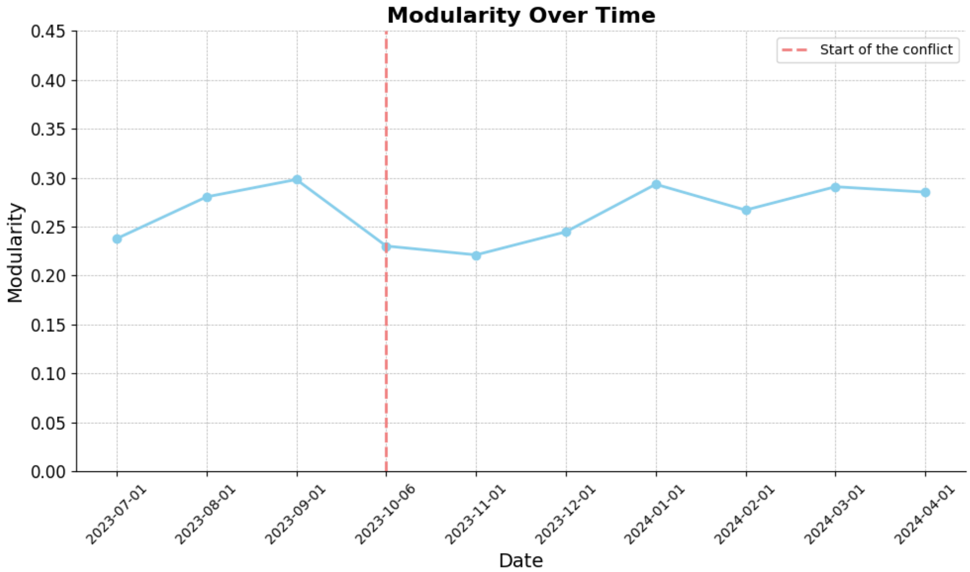

Fig. 4 Monthly evolution of modularity among the top authors of that month

community	authors	posts	interpretation	Topic	Count	Representation
0	12	863	State Legitimacy	0	617	['israel', 'have', 'palestinians', 'people', 'state', 'jews', 'country', 'jewish', 'say', 'israeli']
1	16	1326	State Legitimacy	0	1275	['israel', 'have', 'state', 'jews', 'palestinians', 'people', 'arab', 'palestinian', 'jewish', 'right']
2	21	1540	(Anti) Zionism	0	175	['anti', 'jews', 'jewish', 'zionism', 'religion', 'judaism', 'semitism', 'zionist', 'people', 'have']
2	21	1540	Reddit Rules	1	135	['rule', 'post', 'reddit', 'user', 'comment', 'mod', 'content', 'attack', 'remove', 'sub']
2	21	1540	Americans, Reparation	2	97	['black', 'white', 'americans', 'people', 'reparation', 'native', 'jews', 'indigenous', 'american', 'america']
2	21	1540	State Legitimicy	3	84	['jews', 'have', 'israel', 'palestinians', 'palestine', 'people', 'state', 'palestinian', 'country', 'right']
3	15	882	State Legitimacy, Peace	0	65	['state', 'israel', 'palestinians', 'palestine', 'palestinian', 'solution', 'want', 'peace', 'have', 'support']
3	15	882	Palestinians Jews Terrorism	1	56	['palestinians', 'jews', 'palestinian', 'palestine', 'disaster', 'say', 'pro', 'see', 'terrorism', 'correct']
3	15	882	Politeness	2	48	['thank', 'remindme', 'point', 'hour', 'leave', 'interesting', 'reply', 'good', 'read', 'sure']
3	15	882	(Anti) Zionism	3	47	['anti', 'zionism', 'zionist', 'antisemitism', 'jews', 'racist', 'people', 'israel', 'think', 'antisemitic']
4	19	1265	Civilian, Target, IDF, Terrorist, Settler	0	156	['civilian', 'target', 'idf', 'terrorist', 'settler', 'kill', 'attack', 'terrorism', 'group', 'israel']
4	19	1265	Israel-Palestine Peace	1	85	['israel', 'palestinians', 'conflict', 'peace', 'palestinian', 'state', 'palestine', 'war', 'right', 'israeli']
4	19	1265	State Legitimacy, Religion	2	80	['jewish', 'jews', 'state', 'law', 'islam', 'religion', 'muslim', 'christians', 'people', 'muslims']
4	19	1265	Reddit Rules	3	80	['rule', 'post', 'comment', 'sub', 'violate', 'mod', 'moderation', 'metaposte', 'ban', 'respond']
4	19	1265	Politeness	4	70	['answer', 'thank', 'question', 'talk', 'tone', 'think', 'light', 'kid', 'have', 'explain']
5	15	1577	State Legitimacy	0	1557	['israel', 'have', 'palestinians', 'people', 'say', 'jews', 'palestinian', 'state', 'israeli', 'right']
6	21	2292	State Legitimacy	0	2246	['israel', 'have', 'jews', 'people', 'arab', 'palestinian', 'palestinians', 'jewish', 'state', 'say']

Table 1 Topic Modeling for detected communities pre- October 7th

displayed. To avoid noise, we chose not to display the category that plays the role of placeholder for data points that do not clearly belong to any of the identified topics along with topics that account for less than 5% of the observations of the topic.

Fig. 5 displays the 7 optimal communities as detected by the Louvain algorithm, together with its low modularity score of 0.2, indicating a weak communities' subdivision in this network over this period. Spring layout uses distance between the nodes as a measure of the strength of interactions among these communities. Communities 2 and 3 (C2, C3) seem to be well connected to all the others while community 6 (C6) is marginalized.

When crossing the results from topic modeling (Table 1), one can see that for C2 and C3, the algorithm wasn't able to detect a clear topic for the majority of the posts. In fact, the most important topic for C2 was "(Anti) Zionism" with (175/1540 observations) and for C3 "State Legitimacy, Peace" with (65/882). Interestingly, the inability

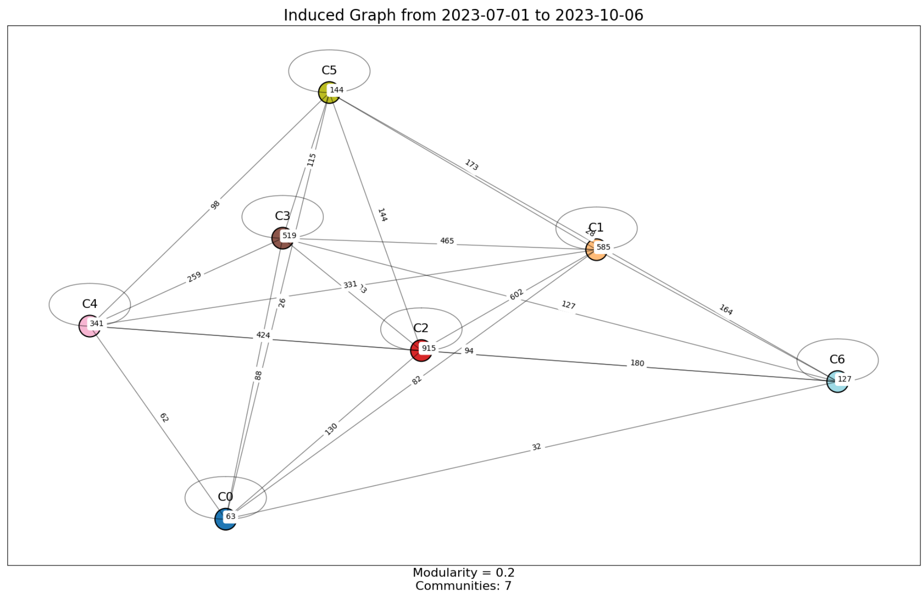

Fig. 5 Detected Communities pre- October 7th

of the algorithm to detect any topic gives credit to the finding that C2 and C3 are more central in the network, possibly due to the wide range of topics discussed. C4 is the third and last community which classified the majority of its posts as noise. Among the topics discussed by these communities, we can find "(Anti) Zionism", "Peace", but also talks about "Terrorism". While for C3, it is unclear whether Israel or Palestine is being blamed for terrorism, for C4 the association of "terrorism" with "IDF (Israel Defence Forces)", "settler" but not "hamas" hints that Israel is being blamed. The latter interpretation should be taken with care since for these communities, the majority of the posts are considered to be noise. Concerning the rest of the communities (C0, C1, C5, C6), they are all composed of the single topic "State Legitimacy", indicating that these groups only interacted about that topic. Unfortunately, the algorithm is not able to detect the position in the way "State Legitimacy" is discussed among these groups.

community	authors	posts	interpretation	Topic	Count	Representation
1	18	10737	War	0	10660	['israel', 'hamas', 'have', 'people', 'palestinians', 'jews', 'gaza', 'war', 'say', 'want']
2	13	5244	War, State Legitimacy	0	5166	['israel', 'hamas', 'have', 'people', 'palestinians', 'jews', 'say', 'think', 'gaza', 'state']
3	27	6749	(Anti) Zionism	0	493	['anti', 'zionism', 'jews', 'antisemitism', 'semitic', 'zionist', 'semitism', 'antisemitic', 'jewish', 'hate']
4	18	5369	Hamas, Terrorist, Massacre	0	468	['hamas', 'support', 'terrorist', 'oct', 'civilian', 'celebrate', 'thousand', 'isis', 'massacre', 'condemn']
5	11	5394	Hamas, Terrorist	0	300	['hamas', 'destroy', 'goal', 'eliminate', 'action', 'leadership', 'group', 'terrorist', 'power', 'remove']
6	8	1404	Propaganda, Lie, Source	0	341	['propaganda', 'side', 'say', 'provide', 'fact', 'read', 'lie', 'argument', 'source', 'see']
6	8	1404	State Legitimacy, Peace	1	102	['gaza', 'hamas', 'israel', 'want', 'people', 'murder', 'peace', 'palestinians', 'make', 'have']
6	8	1404	People, Murder, Peace	2	83	['israel', 'citizen', 'right', 'marriage', 'have', 'israeli', 'equal', 'israelis', 'class', 'fact']
6	8	1404	State Legitimacy	3	73	['land', 'palestine', 'jordan', 'jews', 'create', 'empire', 'buy', 'own', 'england', 'muslims']
7	16	3923	(Anti) Zionism	0	345	['islam', 'jews', 'muslims', 'jew', 'antisemitism', 'muslim', 'antisemitic', 'religion', 'hate', 'antisemite']

Table 2 Topic Modeling for detected communities post- October 7th

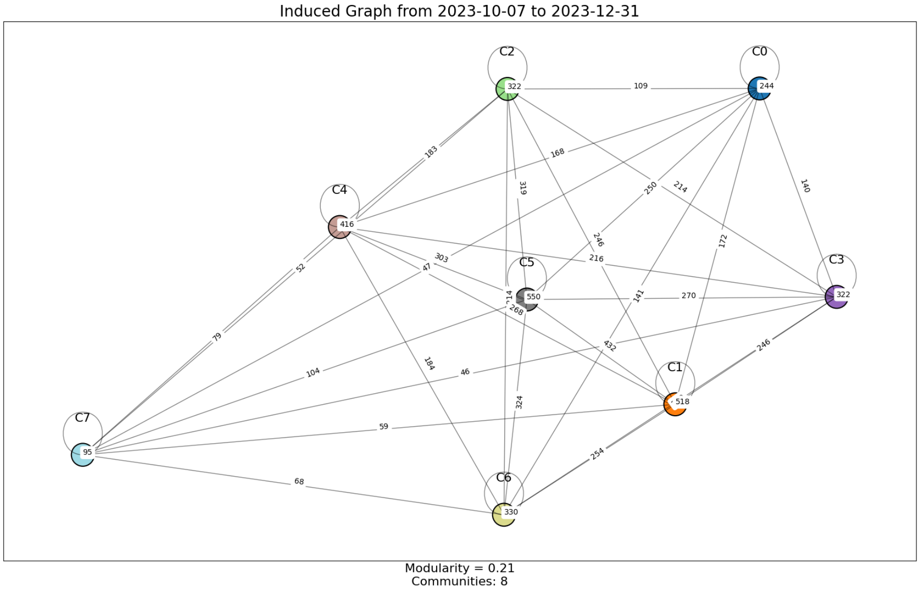

Fig. 6 Detected Communities post- October 7th

Since the start of the war until the 31st of December, there is a switch from "State Legitimacy" to "War" as the main topic of discussion. This can be inferred from having for communities C1 and C2 "War" as interpreted topic for the majority of the posts in these communities (10660/10737 and 5166/5244). For this period of analysis, the BERTopic algorithm was not able to distinguish well defined topics. If fact, except for C1 and C2, the count of documents for each found topic is very low, implying that the majority of the posts have been classified as noise. This means that while probably discussing about war as the main topic of discussion, the variety of topics it was linked to made it impossible for the algorithm to detect it as a single identifiable category. As expected, the topics discussed pre- and post- October 7th are different. "Hamas" and "Terrorist" appear as distinct topics with notably for C4, the terms "civilian", "massacre", "thousands" being detected and clearly referring to Hamas deadly attacks. C4 and C5, the two communities discussing this topic show some proximity in the induced graph of Fig. 5, pointing at frequent interactions among the two. Community 6 (C6) is interesting because of the relative variety of discussed topics compared to others. "Peace" was in fact detected together with "State Legitimacy", "People", "Murder" indicating that this community engaged in progressive and conflict resolution interactions. C6 also shows the surprising yet informative result of having "Propaganda", "Lie", "Source" as most discussed topic (341/1404) revealing that this community was somehow more rational when interacting, paying attention to detect fake news.

6 Conclusion

In this study, we have examined the dynamics of online polarization within the context of the Israel-Palestine conflict by analyzing user interactions on the r/IsraelPalestine subreddit. The analysis of network measures from July 7, 2023, to April 30, 2024, revealed significant fluctuations in author connectivity, with a notable peak in all descriptive measures during Network 5, indicating increased interaction and cooperation among top authors in the period covering the entire month of November 2023. These findings are solidified by lower levels of modularity (the network is less divided) in those months and an increased number of interactions between authors with opposite views. The topic modeling approach showed a switch in the main discussed topic from "State Legitimacy" to "War" after October 7th 2023. The analysis of the non-principal topics of discussion revealed some communities that showed partisanship for one of the parties, together with a community engaging in rational debating. Communities who discussed a wider range of topics took a central place in the network, possibly playing the role of coordinator among communities.

However, some limitations must be acknowledged. Firstly, accurately measuring polarization in a network presents some intrinsic difficulties, and our chosen methods may not capture the full complexity of these dynamics. Additionally, our focus on a single subreddit, r/IsraelPalestine, with its specific regulation and user demographics, limits the generalizability of our findings. Similar studies on other social media platforms, such as Twitter, or even other subreddits may reveal different polarization effects.

Moreover, the relatively small sample size of our observations (1328 overall top users) restricts the extent to which we can generalize our results to the broader population. The classifier used to determine user partisanship is not infallible and may misclassify some users, reflecting the broader challenge of assessing user beliefs based only on online posts. Lastly, by treating all interactions equally, we may have oversimplified the network, potentially obscuring more complex community structures and the presence of polarization.

Despite these limitations, our research contributes to the understanding of how online discourse evolves in the context of prolonged conflictual events. Future research should aim to address these limitations by incorporating multiple social media platforms, enhancing classification accuracy, and developing more sophisticated measures of interaction quality and network dynamics.

References

Barberá P (2014) How social media reduces mass political polarization. evidence from germany, spain, and the u.s. URL https://api.semanticscholar.org/CorpusID:1345712

Bennett WL, Iyengar S (2008) A New Era of Minimal Effects? the Changing Foundations of Political Communication. Journal of Communication 58(4):707–731. https://doi.org/10.1111/j.1460-2466.2008.00410.x, URL https://doi.org/10.1111/j.1460-2466.2008.00410.x, https://academic.oup.com/joc/article-pdf/58/4/707/22323968/jjnlcom0707.pdf

Blondel VD, Guillaume JL, Lambiotte R, et al (2008) Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment 2008(10):P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Bodrunova S, Blekanov I, Smoliarova A, et al (2019) Beyond left and right: Real-world political polarization in twitter discussions on inter-ethnic conflicts. Media and Communication 7(3):119–132. https://doi.org/10.17645/mac.v7i3.1934, URL https://www.cogitatiopress.com/mediaandcommunication/article/view/1934

Boxell L, Gentzkow M, Shapiro JM (2017) Is the internet causing political polarization? evidence from demographics. Nber working paper no. w23258, National Bureau of Economic Research, URL http://www.nber.org/papers/w23258, accessed on April 14, 2017

Clauset A, Newman MEJ, Moore C (2004) Finding community structure in very large networks. Phys Rev E 70:066111. https://doi.org/10.1103/PhysRevE.70.066111, URL https://link.aps.org/doi/10.1103/PhysRevE.70.066111

Colleoni E, Rozza A, Arvidsson A (2014) Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data. Journal of Communication 64(2):317–332. https://doi.org/10.1111/jcom.12084, URL https://doi.org/10.1111/jcom.12084, https://academic.oup.com/joc/article-pdf/64/2/317/22322361/jjnlcom0317.pdf

Conneau A, Khandelwal K, Goyal N, et al (2019) Unsupervised cross-lingual representation learning at scale. CoRR abs/1911.02116. URL http://arxiv.org/abs/1911.02116, 1911.02116

Conover M, Ratkiewicz J, Francisco M, et al (2021) Political polarization on Twitter. Proceedings of the International AAAI Conference on Web and Social Media 5(1):89–96. https://doi.org/10.1609/icwsm.v5i1.14126, URL https://doi.org/10.1609/icwsm.v5i1.14126

Grootendorst MR (2022) Bertopic: Neural topic modeling with a class-based tf-idf procedure. ArXiv abs/2203.05794. URL https://api.semanticscholar.org/CorpusID:

247411231

Himelboim I, Smith M, Rainie L, et al (2017) Classifying twitter topic-networks using social network analysis. Social Media + Society 3:205630511769154. https://doi.org/10.1177/2056305117691545

Imtiaz A, Khan D, Lyu H, et al (2022) Taking sides: Public opinion over the israel-palestine conflict in 2021. 2201.05961

Itai Himelboim MS, Shneiderman B (2013) Tweeting apart: Applying network analysis to detect selective exposure clusters in twitter. Communication Methods and Measures 7(3-4):195–223. https://doi.org/10.1080/19312458.2013.813922, URL https://doi.org/10.1080/19312458.2013.813922

Jungherr A (2014) Twitter in politics: A comprehensive literature review. SSRN URL https://ssrn.com/abstract=2402443, february 27, 2014

Melton CA, Olusanya OA, Ammar N, et al (2021) Public sentiment analysis and topic modeling regarding covid-19 vaccines on the reddit social media platform: A call to action for strengthening vaccine confidence. Journal of Infection and Public Health 14(10):1505–1512. https://doi.org/https://doi.org/10.1016/j.jiph.2021.08.010, URL https://www.sciencedirect.com/science/article/pii/S1876034121002288, special Issue on COVID-19 – Vaccine, Variants and New Waves

Smith MA, Rainie L, Shneiderman B, et al (2014) Mapping Twitter topic networks: From polarized crowds to community clusters. Tech. rep., Pew Research Center, URL https://www.pewresearch.org/internet/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/

Watchful1 (2024) Pushshift dump utils. URL https://github.com/Watchful1/PushshiftDumps

Yarchi M, Baden C, Kligler-Vilenchik N (2020) Political polarization on the digital sphere: A cross-platform, over-time analysis of interactional, positional, and affective polarization on social media. Political Communication 38:1–42. https://doi.org/10.1080/10584609.2020.1785067

Zeitzoff T (2017) How social media is changing conflict. Journal of Conflict Resolution 61(9):1970–1991. https://doi.org/10.1177/0022002717721392, URL https://doi.org/10.1177/0022002717721392